A PROOF-OF-CONCEPT STUDY IN THE DEVELOPMENT OF AN OBSERVATION PROTOCOL BASED ON SCIENCE AND ENGINEERING PRACTICES

By

Jennifer L. Maguire

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Curriculum and Instruction

Nancy E. Bodenhorn, Co-Chair

Natalie L. K. Ferand, Co-Chair

Brenda R. Brand

Takumi C. Sato

Gary E. Skaggs

April 18, 2024

Blacksburg, Virginia

Keywords: inquiry, instructional leadership, observation protocol, science and engineering practices, science teaching, three-dimensional teaching

A Proof-of-Concept Study in the Development of an Observation Protocol

Based on Science and Engineering Practices

Jennifer L. Maguire

Abstract

Inquiry-based science education has been a topic of extensive study and discussion. In 2018, the National Science Teaching Association revised their recommendations about inquiry-based teaching, endorsing three-dimensional learning as a more comprehensive approach for students in science classes. This model, known as three-dimensional teaching and learning (3DT), integrates three key dimensions: 1) science and engineering practices, 2) cross-cutting concepts, and 3) disciplinary core ideas. Accordingly, this study investigated the feasibility of developing an observation protocol based on the science and engineering practice "Planning and Carrying Out Investigations." Participants included both science teachers and instructional leaders. Two research questions guided this investigation: (1) Is it feasible to utilize the Science and Engineering Practices as a framework for the development of an observation protocol intended for use by instructional leaders in secondary science classrooms? (2) How do instructional leaders and science teachers assess a science lesson differently, and what are the implications of these potential similarities or differences on the feasibility of the observation protocol outlined in Research Question 1?

Data was collected over the course of three sequential phases: 1) soliciting feedback from expert reviewers, 2) conducting a focus group with secondary school science teachers and secondary school administrators, and 3) administering an online pilot study of secondary science teachers and secondary school administrators from Virginia, who tested the observation protocol

using video recordings of science lessons. The findings that emerged from this investigation support the idea of using a Science and Engineering Practices-based observation protocol like the one tested in this study. Additionally, the data suggest potential avenues for future research, such as the degree to which administrators would benefit from having subject-specific observation forms.

A Proof-of-Concept Study in the Development of an Observation Protocol

Based on Science and Engineering Practices

Jennifer L. Maguire

General Audience Abstract

Inquiry-based science education has been a topic of extensive study and discussion. In 2018, the National Science Teaching Association revised their recommendations about inquiry-based teaching, endorsing three-dimensional learning as a more comprehensive approach for students in science classes. This model, known as three-dimensional teaching and learning (3DT), integrates three key dimensions: 1) science and engineering practices, 2) cross-cutting concepts, and 3) disciplinary core ideas. Accordingly, this study investigated the feasibility of developing an observation protocol based on the science and engineering practice "Planning and Carrying Out Investigations." Participants included both science teachers and instructional leaders. Two research questions guided this investigation: (1) Is it feasible to utilize the Science and Engineering Practices as a framework for the development of an observation protocol intended for use by instructional leaders in secondary science classrooms? (2) How do instructional leaders and science teachers assess a science lesson differently, and what are the implications of these potential similarities or differences on the feasibility of the observation protocol outlined in Research Question 1?

Data was collected from expert reviewers, secondary school science teachers, and secondary school administrators in Virginia, and involved a focus group as well as rating video recordings of science lessons. Results suggest that using tools based on ideas that are specific to

science could be helpful to both school administrators and teachers. There are many potential follow-up research studies that can be done in the future based on the results of this study.

Dedication

For all of my teachers, starting with Mom (the best teacher I know), whose lessons on Ten-Town, the spelling of "tomato," how to cross-stitch and crochet continue to ring true in my mind- but most importantly her emphasis on the skills of taking pride in my work and doing the best I could do; and Dad, who didn't get too frustrated when I stole all of the ball bearings from his workshop, reset all of the pins in his cylinder locks (I just couldn't stand not knowing how they worked), and took all of the pliers from his tool box to my room to play with- he inspired my love of science way more than he probably realizes; and the amazing schoolteachers that I had through my childhood- the ones who went tremendously out of their way to provide me with opportunities because they saw something in me, and even the few who had dubious faith in my abilities because I couldn't keep my shoes tied and could never organize my notebook to save my life.

And for all of my students, who I love way more than they could ever realize; who seriously don't know the pure joy I feel when I overhear them explaining part of a lesson to each other, when they're whispering to each other about which sticker they got on their quiz this week (you're never too old to get a sticker from your teacher), when I read the beautiful and well-thought-out work they've done, and when they stop at my desk on some days after class to say a simple, "Thank you, Ms. Maguire." The ones who have grown into beautiful adults and become friends to me, the ones who have been tragically lost way too soon, and even the ones who rolled their eyes at me more times than I can remember. They have totally been the best part of so many days as a teacher.

Acknowledgements

Thanking all of the people who have helped me on this journey is nearly as overwhelming as the journey itself. If I forgot you, I am so sincerely sorry. I owe my heartfelt gratitude:

To George Glasson, Brenda Brand, and Takumi Sato. Thank you for allowing me to begin this journey in the first place.

To some truly amazing professors I've learned from during my years at Virginia Tech: Amy Azano, Bonnie Billingsley, Jennifer Bondy, Jen Brill, Peter Doolittle, Jim Garrison, Brett Jones, Betti Kreye, Takumi Sato, Gary Skaggs. A paltry thank you for all you taught me seems so inadequate. I love learning and you are incredible teachers.

To Nancy Bodenhorn, Natalie Ferand, Brenda Brand, Gary Skaggs, and Takumi Sato, my invaluable committee members. My endless gratitude for sticking with me during this process, your endless patience, and your valuable feedback; I am extremely grateful for all of you. Especially to Natalie and Nancy, my advisors, I can't even express how blessed I am to have worked with you both. A double extra special thanks to Nancy, for staying by my side through *so much* and having the ability to always make everything seem so much better.

To Takumi Sato. Your support on this entire journey has been invaluable on so many levels: the consistent attention to my best interests; the mountains of feedback on various projects; finding people to introduce me to at conferences who have similar goals to me; countless cups of coffee, lunches, dinners, and a few adult beverages; and for knowing when I needed to be pushed out of my comfort zone.

To Stephen Biscotte and everyone in Undergraduate Education. I am immensely grateful for the opportunity to work in your office for the past few years. The chance to practice and develop skills I've learned as a student was invaluable, and the dedication you showed to me as a person and to my professional development was humbling. Furthermore, I've made some lifelong friendships with you all and I feel so much joy at having found people who are passionate about some of the same things that I am.

To the teachers, administrators, and other educators who have contributed to the work I did for this dissertation. Thank you for taking the time out of your very busy lives to serve as experts, to attend focus groups, to take surveys, and to kick ideas around with me.

To the fellow doc students that I've met along the way: Pamela, Erika, Justin, Mae, Suzanne, Caryn, Rebecca, Lezly, Larry, Michelle, Marie, Mary, Rachelle, Jim, Tanji, Clint, Kregg, Ricky, Brooke, Azka, Jenni, Donna, Kasey, and Kevin (may he rest in peace). Thank you. So many of you have made this journey as fun as it could be, and on the days when it wasn't fun you were there to cheer me on and pull me along beside you. An extra thanks to Kregg, who has seriously redefined the role of friendship in my life.

To all the pre-service teachers I mentored while here. I delight in seeing the teachers you've become, and I thoroughly enjoyed my small part in your own journeys. It was the best job ever.

To the new colleagues and friends that I made at Christiansburg High School. It was such a joy to be reminded of how much I love and am good at teaching. I didn't expect to work at a school again. But look at all I would have missed out on! I can't imagine. The encouragement, the happy hours, the lake trips! My life is so much richer with you as my friends.

To the colleagues and friends I left behind at Colonial Forge High School. Thank you for allowing me to become the teacher that I am. For nurturing my dreams and not thinking I was insane for wanting this. For the bonds of friendship that have stayed with me these past years.

To Church on Main. I’m beyond blessed to have found a place that values all people and enacts love for our community on a regular basis. The support and encouragement have been especially meaningful from all of you.

To my editor Laurie Good. To say you are a lifesaver is not an overstatement!

And most certainly not least, to my family. I have zero doubts that my parents have done more work to get me to this point in life than I have. The sacrifices they’ve made and endless work they’ve done to make sure I would have more opportunities in life than they did are staggering. To my siblings from birth—Charlie, Lee, and Stephanie; and the ones that joined us later by marriage—Amanda and Patrick: thank you for your love, support, and friendship. And to the extra awesome gifts of my sweet niece and nephews—Maddie, Philip, and Henry: you three make my life sparkle.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Over the past 60 years, one of the most prevalent paradigms in science education is inquiry-based teaching and learning (inquiry; DeBoer, 1991). An early and ardent proponent of inquiry was F. James Rutherford, whose long and distinguished career as a science educator includes time at Harvard University, the National Science Foundation, the United States Department of Education, and the American Association for the Advancement of Science (Peters & Woolley, n.d.). In an often-quoted article from a 1964 volume of the *Journal for Research in Science Teaching,* Rutherford wrote:

> When it comes to the teaching of science it is perfectly clear where we, as science teachers, science educators, or scientists stand: we are unalterably opposed to the rote memorization of the mere facts and minutiae of science. By contrast, we stand foursquare for the teaching of the scientific method, critical thinking, the scientific attitude, the problem-solving approach, the discovery method, and, of special interest here, the inquiry method. (p. 20)

Rutherford distinguished between inquiry as a tool of scientists and inquiry as a pedagogical technique. He argued for teaching methodology that incorporated the use of inquiry-based techniques within the teaching of science content (Rutherford, 1964). From the time of Rutherford's writing up through today, a continuous struggle has existed to see inquiry as a pedagogical technique implemented consistently in K-12 science classrooms (Barrow, 2006; Crawford, 2014; Gray, 2014; Lustick, 2009; Roehrig & Kruse, 2005; Roehrig & Luft, 2004; Sandoval, 2005; Silm et al., 2017; Weiss et al., 2003).

Problem

Inquiry has been one of the most promoted pedagogies in science education (DeBoer, 1991). However, research suggests that science teachers have historically struggled to implement inquiry in K-12 schools (for example, Sandoval, 2005; Weiss et al., 2003). A large volume of research focuses on the implementation of inquiry. One extensive study commissioned by Horizon Research (Weiss et al., 2003) conducted significant research in K-12 science classrooms. One component of their study involved a review of lesson plans; the researchers concluded that in grades K-5, only 17% of the lesson plans they analyzed were inquiry based (Weiss et al., 2003). In grades 9-12, that number dropped to just 2% (Weiss et al., 2003). While some studies do suggest that variations of inquiry are being used sometimes (for example, Capps & Crawford, 2013), others reflect the same observations made by Weiss et al. (2003), that inquiry has not been widely or consistently implemented in K-12 schools in the US (Barrow, 2006; Crawford, 2014; Gray, 2014; Lustick, 2009; Roehrig & Kruse, 2005; Roehrig & Luft, 2004; Sandoval, 2005; Weiss et al., 2003).

Several problems arise if inquiry-based teaching is, as reported, not being implemented in the schools. First, this creates a disconnect for teacher educators in universities. Teacher preparation programs are charged with educating future teachers with skills to teach their content area (science) according to best practice. If universities educate teachers to teach in ways that are not being used or accepted in K-12 schools, these future teachers are being placed in precarious positions of being not able to perform their teaching as it will be evaluated for their performance reviews. Secondly, the students in K-12 are not being taught science in a way that is consistent with scientific practice in the professional community, and thus may enter the workforce or higher education without an understanding of what they will be expected to perform in higher level classes or employment.

Defining Inquiry

When describing the concept of inquiry, in a seminal work on the history of science education, DeBoer states, "If a single word had to be chosen to describe the goals of science educators during the 30-year period that began in the late 1950s, it would have to be inquiry" (1991, p. 206). He described some struggles with the concept of inquiry, namely that multiple definitions of the concept almost immediately evolved and that the field could not come to a consensus on what it did and did not mean in the context of science teaching and learning. Just five years after DeBoer's writing, one of the first large-scale, formalized definitions of inquiry came, when the National Research Council (NRC) published the National Science Education Standards (NSES; NRC, 1996). Inquiry was defined in the NSES as:

> Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations. (National Research Council, 1996, p. 23)

While this definition of inquiry became something of a standard definition that some researchers and practitioners adopted, consistent agreement did not exist within the field about what inquiry meant (Barrow, 2006; Binns & Popp, 2013; DeBoer, 1991). In 2004, the National Science Teaching Association (NSTA) expressed their support for inquiry in a position statement that referred to and elaborated on the NSES definition given above (NSTA, 2004). However, research continued to suggest inquiry was not taking place in classrooms (Barrow, 2006; Crawford, 2014; Gray, 2014; Lustick, 2009; Roehrig & Kruse, 2005; Sandoval, 2005; Silm et al., 2017). In 2018, the NSTA addressed the problem, saying "an uneven implementation of scientific inquiry has occurred in science classrooms" (NSTA, 2018). In response to this problem, the NSTA changed their recommendations from supporting inquiry to supporting a newly coined concept in science education, "Three-Dimensional Teaching" (NGSS, 2012; NSTA, 2018). The NSTA clarified their thoughts about this shift in their recommendations by saying, "It's important to note that this transition is not a rejection of scientific inquiry, but represents further evolution of our understanding about what is essential to promote student learning" (NSTA, 2018).

Appendix A has been included for ease of reference, providing readers with a comprehensive list of acronyms and corresponding definitions encountered throughout the study. Also included in the appendix items are the IRB approval letters for Phase 1 and Phases 2/3 of this study (Appendix L).

Three-Dimensional Teaching and The Next Generation Science Standards

Three-dimensional teaching (3DT) is a new idea that came about with the introduction of new science standards, called the Next Generation Science Standards (NGSS; NRC, 2012; NRC, 2013). The development of the new set of standards began with the publication of *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (2012; Framework). The Framework is a National Research Council (NRC) report that laid the groundwork for the publication of the NGSS and was written by a national committee of science educators and researchers. In contrast to the NSES, which was primarily composed of content standards, the NGSS provides a foundation of science education through the presentation of three dimensions of science learning (NRC, 2012; NRC, 2013). Three-dimensional teaching (3DT) rests fundamentally on the idea of integrating these three dimensions in the classroom (NRC, 2012; NRC, 2013).

Dimension 1 consists of eight science and engineering practices (SEP), shown in Figure 1, that can be incorporated into what students do in their science classes.

Figure 1.

Science and Engineering Practices

1. Asking questions (for science) and defining problems (for engineering)
2. Developing and using models
3. Planning and carrying out investigations
4. Analyzing and interpreting data
5. Using mathematics and computational thinking
6. Constructing explanations (for science) and designing solutions (for engineering)
7. Engaging in argument from evidence
8. Obtaining, evaluating, and communicating information

Note. NRC, 2012, p. 48

When discussing the rationale for focusing on these practices as opposed to a more singular focus on the idea of inquiry, the Framework says,

> Similarly, because the term 'inquiry,' extensively referred to in previous documents, has been interpreted over time in many different ways throughout the science education community, part of our intent in articulating the practices in Dimension 1 is to better

specify what is meant by inquiry in science and the range of cognitive, social, and physical practices that it requires. (NRC, 2013, p. 30)

Dimension 2 is focused on the incorporation of seven cross-cutting concepts (CCC) that describe commonalities in all science topics and that transcend grade level. CCC are generalized statements about science content that apply to all science fields, such as biology and physics. For example, one CCC is Patterns: “Observed patterns in nature guide organization and classification and prompt questions about relationships and causes underlying them” (NRC, 2013, p. 92).

Lastly, Dimension 3 is Disciplinary Core Ideas (DCI), which detailed the specific science content students should know or be able to do by certain times in their schooling. For example, one disciplinary core idea for fourth grade science (number 4-ESS1-1) in the core idea “Earth’s Place in the Universe” and the sub-idea “The History of Planet Earth” is: “Local, regional, and global patterns of rock formations reveal changes over time due to Earth’s forces, such as earthquakes. The presence and location of certain fossil types indicate the order in which rock layers were formed” (NRC, 2013, p. 39). According to the NSTA, 20 states and Washington, D.C. have adopted the NGSS; another 24 states have their own standards that are influenced in some way by the Framework (NSTA, n.d.).

Educator Feedback and Support for Inquiry-based Teaching

There are many potential reasons why teachers might struggle to implement inquiry. Some potential reasons fall under the category of availability of support and resources (Lustick, 2009). Examples of support might include professional development programs and instructional coaching (Luft, 2001), and examples of resources might include physical resources such as lab supplies. One type of support that may be lacking is the feedback teachers receive from local instructional leaders (for example, principals and department chairpersons) on their teaching and lesson construction. Feedback that aligns with inquiry-based pedagogy would be helpful for teachers (NRC, 2000).

Traditionally, feedback from school administrators and other local leaders has not included a strong content-specific component (Hill & Grossman, 2013). Some research has advocated for such and identified some of the barriers to this practice (Hill & Grossman, 2013). For example, a stronger content-specific component would necessitate different practices of instructional leaders by grade level and subject area, which is more difficult for school administrators to use (Hill & Grossman, 2013). However, some research has indicated that local instructional leaders, such as school administrators and district supervisors, have struggled to work with science teachers (Braaten et al., 2017; Lochmiller, 2016).

The Changing Role of School Administrators in Instructional Leadership

Over the past several decades, school administrators have been increasingly tasked with responsibility for instructional leadership that encourages teacher improvement, as opposed to only evaluating teacher performance. *SuperVision and Instructional Leadership* (Glickman et al., 2014), a staple text often used in administrator training programs, stated, "the history of instructional supervision is viewed most often as an instrument for controlling teachers" (2014, p. 8). The authors advocated for a departure from that traditional model of instructional leadership, which emphasized authoritarian control of teacher practice, by providing an argument for a more collegial method of instructional leadership where administrators work collaboratively with teachers. While the authors only discuss subject-specific leadership briefly, the collegial model of instructional leadership complements the idea of exploring feedback mechanisms for science teachers (Glickman et al., 2014). One way for administrators to get involved with inquiry is by collaboration with experts in science education.

Teacher Evaluation and Instructional Leadership in Science

In 2000, science educators made a call for more attention to be shed on the role of teacher evaluation and improvement as it specifically applies to science teachers (NRC, 2000). This publication suggested science teachers need more intense support focused specifically on science education, if they are to implement inquiry in their classrooms (NRC, 2000). This publication was prior to the advent of the NGSS, so the suggestions in this specific publication do not align perfectly with 3DT. However, as 3DT is an evolution of what science education leaders understand inquiry to be, the recommendations remain relevant. Among those recommendations to instructional leaders were: acquiring an understanding of what it means to teach and learn by inquiry, cultivating an understanding of the research-based advantages of such teaching methods, advancing their understanding of the changes that must take place for teachers and students when transitioning to inquiry from a more traditional non-inquiry based classroom, and the development and use of a comprehensive support system that will encourage staff to grow and develop in their ability to use inquiry. The NRC (2000) went on to describe seven concrete suggestions for actions that could result in the support they describe, one of which is "teacher evaluation consistent with inquiry teaching" (p. 144). However, one issue with evaluation that is consistent with inquiry is that "an evaluator must understand inquiry to know what to observe in the classroom" (NRC, 2000, p. 151). The authors were particularly concerned about this because they thought evaluators might see inquiry-based teaching and perceive it as disorganization (NRC, 2000). Additionally, they cited teacher concerns over poor administrative support for any "new" teaching methods (NRC, 2000, p. 147). The advice of the NRC (2000) to administrators is to take teachers' concerns seriously and encourage dialogue among faculty and staff. To summarize the need for their robust recommendations, the authors stated that if changes are not made to teacher evaluation to reflect what classrooms that implement inquiry look like, it will ultimately cause problems somewhere in the future (NRC, 2000).

When the Framework introduced NGSS to the science education community, the NRC also took the time to describe the current state of science education and identified some barriers to the improvement of conditions for students and teachers (NRC, 2012). They revisited the idea of instructional leadership as it relates to science education (NRC, 2012). They suggest that we will know progress is being made in removing barriers to improved teaching and learning when we begin to see changes to teacher evaluation and administrator training as it related to science teacher performance (NRC, 2012). As previously discussed, the NSTA began to push for the implementation of 3DT, rather than simply inquiry, in 2018. Presumably recognizing the role that instructional leaders play in classroom practice, the NSTA offered specific professional development (PD) for them in 2023, a full day Leaders Institute that took place concurrently with the NSTA annual conference. In advertising the event, the NSTA said "education leaders will explore practical, research-based approaches that transform teaching and learning in the classroom" (NSTA, n.d. a). One breakout session offered during the Leaders Institute was titled "Next Gen Standards for Next Gen Students: Supporting the Teacher as the Designer of Science Instruction," which was described as being a session where participants would "learn about "look for's" for high quality instruction, their implications for evaluation and the value of giving teachers both confidence and inspiration to innovate," and that the session takeaway would be "Administrators will learn practical ways to support teachers in shifting instructional practices that support the demands of the NGSS" (NSTA, n.d. a).

Research Proposal

Numerous actions and interventions might result in increased inquiry taking place in classrooms. Considering the above-discussed concerns about instructional support and leadership that aligns with science teaching and learning, this study will focus on the role that instructional leaders could play in supporting teachers implementing inquiry-based instruction by investigating the ways that an observation protocol might be structured around the SEP.

Research Questions

1. Is it feasible to utilize the Science and Engineering Practices as a framework for the development of an observation protocol intended for use by instructional leaders in secondary science classrooms?
2. How do instructional leaders and science teachers assess a science lesson differently, and what are the implications of these potential similarities or differences on the feasibility of the observation protocol outlined in Research Question 1?

Purpose

The purpose of this study was to determine the potential for developing an observation protocol tool based on observable features of the SEP. SEP were chosen as the framework for this observation protocol because they are the most recent iteration of what it means to teach by inquiry. The Framework discussed their choice to emphasize practices over inquiry by saying, "Our view is that this perspective is an improvement over previous approaches in several ways" (NRC, 2012, p. 43). The authors listed three ways:

- "First, it minimizes the tendency to reduce scientific practice to a single set of procedures…" (p. 43).

- “Second, a focus on practices (in the plural) avoids the mistaken impression that there is one distinctive approach common to all science—a single ‘scientific method’…” (p. 44).
- “Third, attempts to develop the idea that science should be taught through a process of inquiry have been hampered by the lack of a commonly accepted definition of its constituent elements. Such ambiguity results in widely divergent pedagogic objectives—an outcome that is counterproductive to the goal of common standards” (p. 44).

For these three reasons, the SEP provided a solid framework for the preliminary design of an observation protocol.

Study Outline

Since this study only determined the feasibility of using SEP as a framework for an observation protocol, only one SEP, Planning and Carrying Out Investigations, was piloted in this study. The Framework summarizes the importance of Planning and Carrying Out Investigations by saying, “Scientists and engineers investigate and observe the world with essentially two goals: (1) to systematically describe the world and (2) to develop and test theories and explanations of how the world works” (2012, p. 59). While it would be overly restrictive and misrepresentative to suggest that any SEP is more important than the others, Planning and Carrying Out Investigations (PCI) is of paramount importance in the development of science learners, and incorporates some aspects of many other SEP. The Framework (NRC, 2012), NGSS (NRC, 2013), and Virginia Standards of Learning (SOLs) each define and describe PCI with a finer grain size at different grade levels and with different topics. For example, the Virginia Biology SOLs includes the statement “select and use appropriate tools and technology to collect, record, analyze, and evaluate data” as a feature of PCI (VDOE, 2018, p. 35), while the NGSS (2013) includes the statement “collect data to produce data to serve as the basis for evidence to answer scientific questions or test design solutions under a range of conditions” when describing what PCI looks like for middle school students (p. 239). These finer grain-sized statements clarifying PCI were expanded and adapted into individual observable items in the development of this prototype observation protocol, which this study tested with instructional leaders and teachers.

Teachers and instructional leaders acted as participants. The participants used the PCI-based prototype observation protocol to identify elements of inquiry in previously recorded video segments of a secondary science class. Comparisons were made between how the teachers and instructional leaders rated the class. Data analyses common to the process of instrument development were performed. Participants were asked to reflect on the experience of using the observation protocol and reported their thoughts about the experience. These findings were used to discuss the potential for using SEP as a framework to develop a full observation protocol. Additionally, future use of an inquiry-focused observational protocol facilitating the implementation of inquiry in secondary science classes was discussed.

Chapter Summary

Inquiry has been a frequently promoted and debated topic in science education for more than 60 years. Two major issues with inquiry were that 1) there was no standard definition of what it means to teach or learn by inquiry, and 2) evidence suggests science teachers are not regularly using inquiry-based methods to teach. The NSTA previously recommended inquiry-based teaching to science teachers, but more recently updated this recommendation to promoting 3DT, which involves the incorporation of three dimensions: cross-cutting concepts, science and engineering practices, and disciplinary core ideas. One reason for this change was to streamline and simplify the meaning of inquiry-based teaching to support teachers in switching over to using inquiry-based methods with more frequency and fidelity.

Instructional leaders are potentially an untapped support mechanism for science teachers regarding the implementation of inquiry-based teaching. Trends in the field of educational administration support and encourage instructional leadership that could be leveraged to provide science teachers with greater support for inquiry-based teaching. However, support and training for instructional leaders themselves may be helpful and even necessary to this process. This research involved the participation of both science teachers and instructional leaders in the Commonwealth of Virginia to develop and test an observation protocol based around the SEP Planning and Carrying Out Investigations.

Chapter 2. Literature Review

In Chapter 1, I discussed the concept of inquiry and that it is not happening universally in science classrooms. In Chapter 2, I reviewed literature related to how inquiry is identified in practice. Discussions of how inquiry is identified are important components within the larger issue of why it may not be occurring. Additionally, I reviewed literature regarding support and leadership available to science teachers.

The first section of this chapter examined existing instruments that are used to identify or measure the use of inquiry-based teaching practices. The second section of this chapter examined work that has been done that describes the roles of various support personnel for science teachers, such as department chairpersons. The third and last section examined work related to content-area-specific instructional leadership.

Conceptual Framework: Leadership Content Knowledge

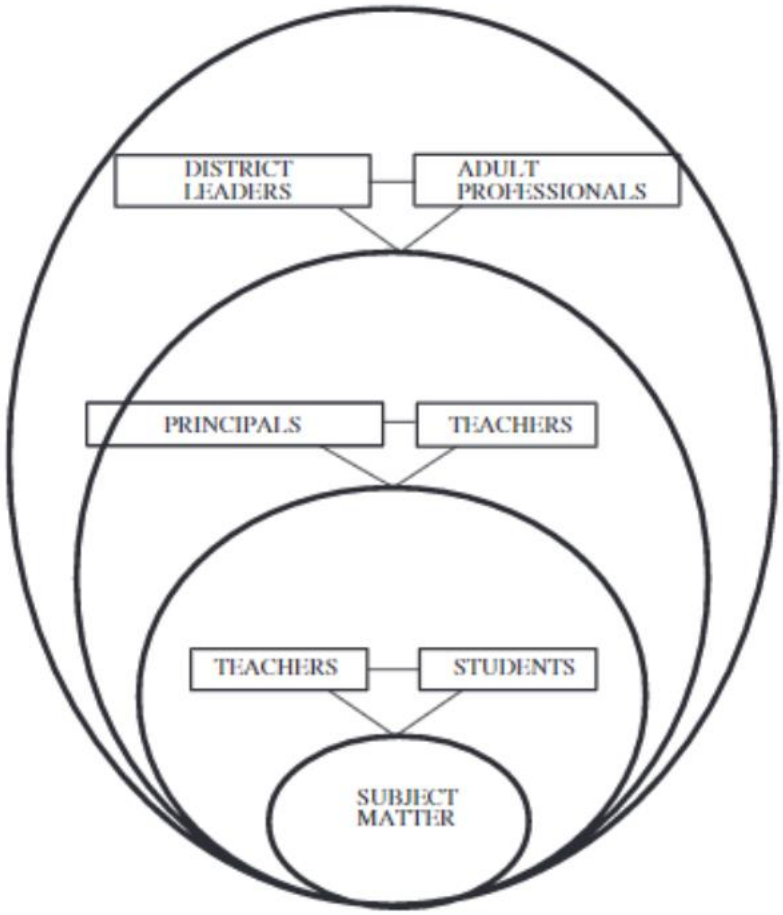
Stein and Nelson (2003) conducted a cross case analysis to construct a detailed model for what they term a "missing paradigm" for viewing instructional leadership (p. 445). They found that, in the same way that teachers require pedagogical content knowledge to best teach a subject (Shulman, 1986), administrators require a unique type of knowledge to be effective instructional leaders. They termed this new construct "leadership content knowledge" (LCK; Stein & Nelson, 2003, p. 424) and suggested that it has been missing in the past and is a vital component for strengthening instructional leadership.

In their first case, Stein and Nelson (2003) analyzed the use of "classroom observation and teacher supervision to identify professional development needs" (p. 427). They followed the case of Claudia, an elementary school principal and analyzed her instructional leadership over math lessons. In the second case, they followed an assistant superintendent tasked with selecting a new curriculum for a small school district. In a third case they followed a three-person team from a central office tasked with leading instructional improvement efforts over 25 elementary schools (Stein & Nelson, 2003).

Stemming from their work with these cases, Stein and Nelson (2003) proposed a four-tiered model to represent their interpretation of the construct of leadership content knowledge (see Figure 2). In the first tier, we find the subject matter itself. Stein and Nelson's research (2003) focused on math as a subject area. In all the cases, the leaders understood the nuances of math as a content area, though it was to varying levels and for different purposes. The second tier is where teachers teach, and students learn. This tier involves an understanding of how students learn math and how teachers can help facilitate this process. In the third tier, the teachers become learners and principals act as teachers to them. This tier involves knowing the mechanisms by which teachers learn to teach math, and the ways in which their administrators can help them improve the learning of teaching skills. In addition, it involves what principals should know about guiding the learning of all the teachers in their school as a group. The fourth tier is where district leaders, such as superintendents, become the instructors, and all other professional adults (that is, principals, teachers, and other school personnel) within a school district act as learners (Stein & Nelson, 2003).

Figure 2.

Conceptual Framework: The Construction of Leadership Content Knowledge



Note. Stein & Nelson, 2003, p. 425

The fourth tier involves a relatively unknown idea in schools and among educational leaders about "how principals learn to lead subject-matter reform," (Stein & Nelson, 2003, p. 441) and what type of support the principals need to do this effectively. Stein and Nelson (2003) concluded their work by proposing that leadership which connects more closely to subject matter may be much more effective than generic leadership because it will align more closely to the core of education: what students and teachers are doing in classrooms.

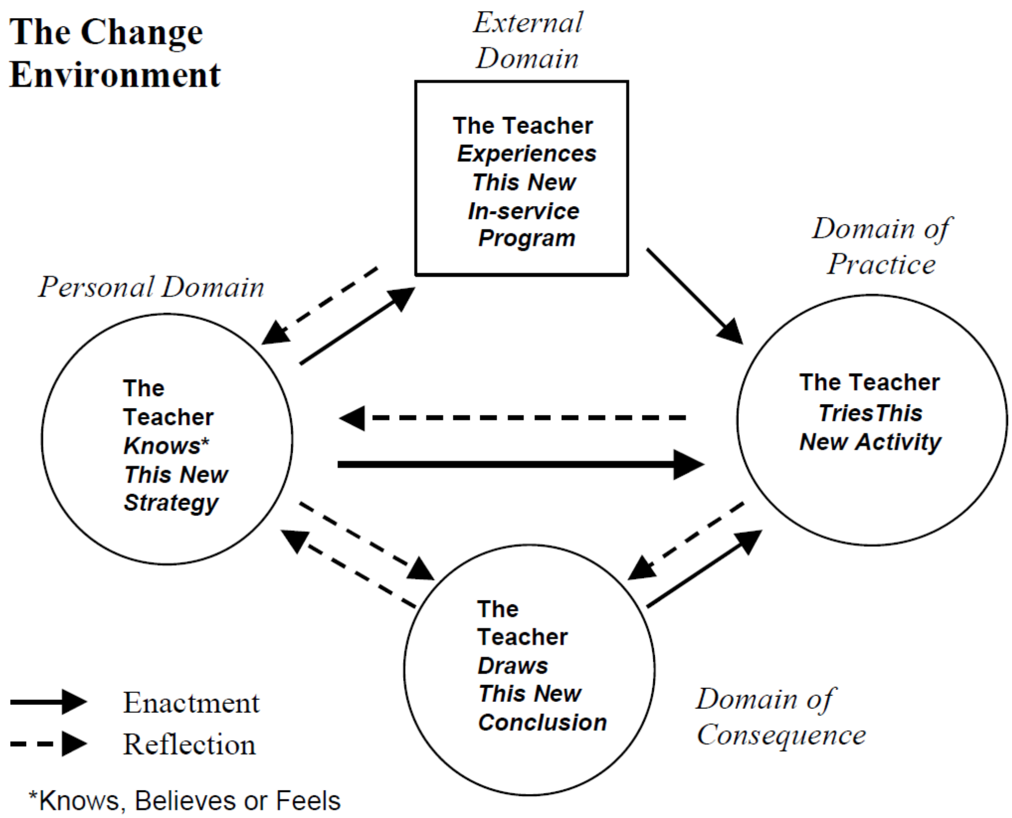
The research in this dissertation study primarily aligns with the third tier of Stein and Nelson's (2003) proposed model, though it has implications for the fourth tier as well. The process of developing and using the observation protocol, termed the Planning and Carrying Out Investigations Observation Protocol (PCI-OP), was viewed through the lens of leadership content knowledge as a construct. The idea that the contextual features of science as a unique subject is an important consideration for successful instructional leadership was paramount to this research.

Conceptual Framework: Teacher Professional Growth

Clarke and Hollingsworth (2002) proposed the Interconnected Model for Teacher Professional Growth, which focused on complex and nonlinear interactions between four primary domains: personal domain, external domain, domain of practice, and domain of consequence. The personal domain includes knowledge, beliefs, and attitudes. The external domain is an external source of information or stimulus. The domain of practice is where professional experimentation takes place. The domain of consequence is where salient outcomes are observed (Clarke & Hollingsworth, 2002). The interactions between domains take the form of either reflection or enactment, or sometimes both reflection and enactment (Clarke & Hollingsworth, 2002). Figure 3 illustrates this model with specific examples of events that might occur during teacher professional learning and change. Figure 3 illustrated a situation where a teacher might experience something in an in-service, causing them to try a new activity or to know, believe, or feel a certain way about a strategy. Either or both of those interactions might result in the teacher drawing a new conclusion. The cycle may continue in further interactions between these events.

Figure 3.

Conceptual Framework: Teacher Professional Growth



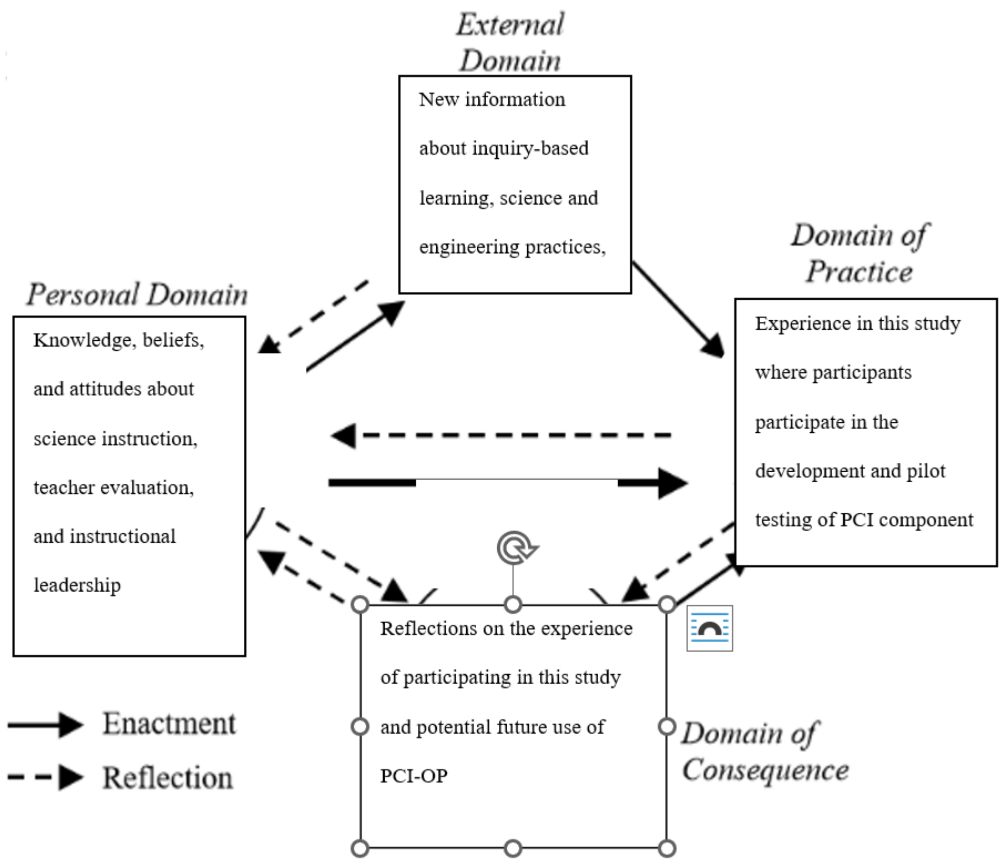
Note. Clarke & Hollingsworth, 2002, p. 957

Clarke and Hollingsworth (2002) described three functions of their model- use as an analytical tool, use as a predictive tool, and use as an interrogatory tool. As an analytical tool, the model might be used to categorize data about teachers who make changes to their practices, identifying structural locations where further teacher professional development (PD) might have the most impact. For instance, data might suggest that focusing on what teachers believe rather than their experimenting of strategies might be more likely to result in a permanent change in practice. As a predictive tool, the model is used more to identify specific sequences between the domains and potentially leverage those in professional development. In that case, the model might find that teachers typically bounce back and forth between changing their beliefs and trying new ideas several times before drawing a conclusion, and this might provide a structure that can be used in certain PD events. As an interrogatory tool, the model provides support for answering specific questions about interactions between the domains. For example, researchers might want to know more about how many times a teacher might experiment with a new strategy before concluding they do not believe in its effectiveness with students.

Clarke and Hollingsworth's (2002) model of teacher professional growth (TPG) provided an additional theoretical lens for this study. In this study, information about inquiry and 3DT occupies the External Domain of the model, subsequently interacting with the participants' Personal Domain- pre-existing knowledge, beliefs, and attitudes about science teaching, teacher evaluation, and instructional leadership. As the model shows, these two domains interact with one another, and both affect the Domain of Practice. The professional experimentation element in the Domain of Practice was the actions that participants engaged in as part of study, where participants (teachers and instructional leaders) use an instrument to rate a science lesson. Following their use of the instrument, participants were asked to reflect on their experience, and to discuss foreseeable salient outcomes in the Domain of Consequence. Figure 4 uses Clarke & Hollingsworth's (2002) diagram, adapted to show applications of this study.

Figure 4.

Conceptual Framework for Implementing Three-Dimensional Teaching



Note. Adapted from Clarke & Hollingsworth, 2002, p. 951

Clarke & Hollingsworth's (2002) model of TPG is particularly relevant to this study because, as discussed in chapter 1, science education has historically had a problem with engaging teachers with inquiry-based teaching and learning (e.g. Silm et al., 2017). In 2012 and 2013 the NRC addressed this issue with the creation of the Framework and NGSS, stating that 3DT better represents what is meant by the term inquiry. In 2018, the NSTA reiterated the NRC's position by discussing problems with inquiry never quite catching on and explained that they were transitioning from using the phrase inquiry to using the phrase 3DT in their events and materials to support teacher change. The NSTA bookstore, website, and conferences are rife with teacher resources that support 3DT, and just recently they have begun offering science education-based PD to instructional leaders. Science education as a field is deeply entrenched in the throes of change, with professional organizations such as the NSTA and the NRC developing robust support mechanisms to support these changes. While teaching is already a complex endeavor, the adoption of inquiry-based teaching by any definition is a large-scale and even more complicated process for science teachers to undergo. Clarke & Hollingsworth (2002) argued for the need for a complex model that represents such difficult changes by saying:

> The optimization of the outcomes of a process is facilitated by the understanding of that process. If we are to facilitate the professional development of teachers, we must understand the process by which teachers grow professionally and the conditions that support and promote that growth (p. 947).

Applying Clarke & Hollingsworth's (2002) model to the pilot testing of the PCI-OP instrument allowed the researcher to consider the experiences of participants as they use PCI-OP in order to evaluate whether it is a viable support tool for teachers and instructional leaders.

The Interaction of Two Lenses

Stein and Nelson's (2003) concept of LCK is relevant as a lens through which to view this work because it describes the relationship of an instructional leader to science classrooms and the potential changes that might take place in becoming a support mechanism for science teachers. Clarke and Hollingsworth's (2002) model for TPG is relevant as a lens through which to view this work because it describes the relationship of a science teacher to their curriculum and instruction and the potential changes they might make if they were to adopt pedagogical methods that are more inquiry-based. For instructional leaders, this work is about finding the tools they need to be well-versed in inquiry-based teaching and other trends in science education so that they may become strong supports for teachers. For teachers, this work is about making potential inquiry-based changes once they have strong support in place. The problem studied in this work rests on the relationship between the unique but complementary roles of instructional leaders and teachers.

Identifying and Measuring Inquiry-Based Teaching

If school administrators, department chairpersons, mentors, and others are to serve as vital supports in the efforts for science teachers to receive feedback on their practices, these stakeholders must have tools available that will help them know inquiry when they see it and provide feedback appropriately (NRC, 2000; NRC, 2012; NRC, 2013). Observation protocols are instruments that a respondent completes while observing a class (Bodzin & Beerer, 1993; Hayes et al., 2016; Luft, 1999; Marshall et al., 2010; Sampson, 2004; Smolleck et al., 2006; Wheeler et al., 2019). However, most instruments reviewed below are primarily designed for science education researchers and other science education experts.

I searched for publications that described the development or use of instruments for evaluating science teaching. I specifically looked for information regarding the purpose and context for the use of these tools. I noted who the respondent was intended to be, such as a teacher, researcher, or principal. I looked specifically for publications that discussed the tools from an instrument development viewpoint, as opposed to studies that used the instruments and reported their findings about measuring inquiry in a specific context.

I considered instruments that focused on attitudes and beliefs of science teachers, such as the Self-Efficacy Beliefs about Equitable Science Teaching and Learning instrument (Ritter et al., 2017) and the Science Teaching Efficacy Beliefs Instrument (Enochs & Riggs, 1990). I considered instruments that focused on a narrow aspect of science teaching, such as the Classroom Observation Protocol for Engineering Design (Wheeler et al., 2019), which focuses only on the implementation of engineering design in science classrooms.

Utilizing the methods previously described, I identified six instruments that specifically measured inquiry-based teaching. I described these six below. Three are observation protocols, one is a combination of an observation protocol and a self-report, and two are only self-reports. After discussing those six instruments, I also discussed the current parameters of teacher evaluation in Virginia.

Extended Inquiry Observation Rubric

The Extended Inquiry Observation Rubric (EIOR; Luft, 1999) was designed specifically to align with a professional development (PD) program for secondary science teachers that took place over the course of a year and a half. The purpose of EIOR was to measure the degree to which teachers were implementing the science as inquiry standard as defined by the NSES, which was a component of the PD program. The PD program involved summer courses, collaboration with fellow teachers, and observation of fellow teachers (Luft, 1999). Teachers were observed by researchers and PD leaders using the clinical supervision model as described by Acheson and Gall (2003), which involved pre- and post- conferences in conjunction with a classroom visit. Teachers were assessed throughout the course of the PD program and were involved with the design of the rubrics used in EIOR. The researchers found this process to be very useful and made numerous robust suggestions for how teachers can become more involved with the process of getting feedback (Luft, 1999).

The argument that Luft (1999) made for observation protocols that align well with both purpose and instruction resonates deeply within the creation of any new observation tools.

Further, the involvement of teachers within the process of developing observation tools is a critical step that is important to success. However, EIOR was developed to be used by science education researchers. This would not lend well to use by instructional leaders in a secondary school.

The Science Teacher Inquiry Rubric

The Science Teacher Inquiry Rubric (STIR; Bodzin & Beerer, 2003) was developed initially as a self-reflection tool for elementary teachers. The purpose was to increase reflection on elements of inquiry-based teaching in order to lead teachers to implement those elements in their classrooms. During the development of STIR, the researchers also decided to pilot test the instrument as an observation protocol (in addition to testing the instrument as a survey for teachers).

The authors developed a rubric based on the NSES definition of inquiry (NRC, 1996). They used an expert panel to gather feedback and made the appropriate revisions. For the pilot study the observers were a researcher and a district level supervisor; both had prior experience as principals. They observed ten middle and high school teachers for the pilot study (even though they previously said the instrument was for elementary teachers). Following the observations, teachers reflected on the lesson they presented and completed STIR. In their discussion, the authors specifically addressed principals as being the intended respondents and the purpose as being teacher improvement (Bodzin & Beerer, 2003).

As an observation tool, the authors suggest high inter-rater reliability based on the results of the pilot test (both researchers had the exact same ratings for every observation and every indicator). However, data led the researchers to conclude that the STIR was not effective enough when used as self-assessment. The teachers consistently rated themselves differently (typically higher) than the outside observers on many of the indicators. The researchers concluded that the instrument can effectively be used as an observation tool but not a self-report survey for teachers. The STIR rubric might be helpful in informing the use or creation of future tools, and the lessons its authors learned about the difference between self-reports and observer reports (Bodzin & Beerer, 2003).

The idea that teachers and raters who used STIR rated lessons so differently bears further inspection. In this case, both outside raters had been both elementary teachers and principals, but neither had specific training in science teaching. When viewed through the theoretical lenses of LCK and TPG, the implications are that it is possible the ratings were so different, not because the teachers had inflated ratings of themselves, but because the outside raters lacked training on recognizing inquiry-based teaching. The development and testing of PCI-OP involved cooperative work between the researcher, science teachers, and instructional leaders in an effort to mitigate some of these issues. In addition, RQ2 pertains to how science teachers and instructional leaders rate a science lesson differently, the results of which might support or refute the feasibility of using an observation protocol based on SEP.

Science Management Observation Protocol

The Science Management Observation Protocol (SMOP; Sampson, 2004) was designed to help teachers implement inquiry by considering aspects of their classroom management. Sampson (2004) argued that classroom management was a major barrier to the implementation of inquiry. The purpose of the instrument was to help teachers realize aspects of their management that an outside observer might be better positioned to notice than teachers would notice about themselves. The piece was published in *The Science Teacher* (2004), which is primarily a practitioners' journal published for secondary science teachers by NSTA. The author did not give details about how he developed the rubric. The instrument was composed of 25 statements that an observer would rate in Likert fashion. Sampson (2004) briefly discussed the potential for the instrument to be used by teacher colleagues and other leaders. The idea of classroom management as a major factor on inquiry is unique among the other tools reviewed (Bodzin & Beerer, 1993; Hayes et al., 2016; Luft, 1999; Marshall et al., 2010; Smolleck et al., 2006; Wheeler et al., 2019).

While Sampson (2004) did not provide details on either development or validation of SMOP, the idea that SMOP was to be used between colleagues, and potentially by instructional leaders, shows promise. While classroom management is certainly a component of all teaching, the role of classroom management's *specific* relationship to inquiry-based teaching is outside the scope of this project, which involves implementing components of SEP as a way on teaching in a more inquiry-based way. Furthermore, instructional leaders focus a fair amount on classroom management in teacher evaluation already, though it is possible the specific effects of classroom management on inquiry-based teaching have been outside of the scope of that focus. Regardless, Stein and Nelson's (2003) third tier of interaction between principals and teachers is relevant here. Whether considering SEP or classroom management of science teachers, leadership content knowledge related to inquiry-based teaching would be a valuable skill for instructional leaders to develop.

Teaching Science as Inquiry

The Teaching Science as Inquiry instrument (TSI; Smolleck et al., 2006) was developed to measure inquiry, but through the lens of teacher beliefs and self-efficacy. The researchers composed an initial set of 31 items for elementary pre-service teachers (PSTs) to use in a self-report format. Smolleck et al. (2006) went through a total of six rounds of expert panel feedback, followed by significant revisions. For different phases of the review process, the panel of experts differed. Eventually the instrument was pilot tested with 190 elementary PSTs.

Electronic Quality of Inquiry Protocol

The Electronic Quality of Inquiry Protocol (EQUIP; Marshall et al., 2010) was designed to measure both the quantity and quality of inquiry-based teaching that takes place in K-12 science classrooms. The work was conducted in response to the authors' perceptions of inquiry being poorly defined. In addition, the authors were motivated by what they believed was poor guidance provided to both pre- and in-service teachers for how to accomplish inquiry-based lessons. Marshall et al. (2010) stated "we hope that the protocol can be used to guide pre- and in-service teachers' discussions and analyses of inquiry-based instruction" (p. 300). The project that prompted their need for an instrument to measure inquiry-based education involved a science and math PD partnership between a university and a public school district. They sought to understand not only *if* science and math teachers were implementing inquiry-based methods of teaching, but to what degree it was happening in terms of both frequency and fidelity. In addition to inquiry-based teaching and learning, their project also focused on the inclusion of formative assessment and teacher reflection.

Marshall et al. (2010) conducted multiple stages of pilot testing and revising. Ultimately, the team used what they learned during the pilot testing of EQUIP to move from using a simple Likert scale to using a more descriptive rubric format. They felt this was more likely to give better formative feedback on which teachers could reflect and that this allowed observers to rate teachers more objectively and systematically. Additionally, the rubric was designed to be completed directly *after* observing a lesson. The final section of the instrument was designed to give a summative, overall analysis of inquiry-based teaching (Marshall et al., 2010).

EQUIP was well grounded in inquiry-based teaching. However, like other protocols (Luft, 1999; Bodzin & Beerer, 2003), it was meant to be used by science education researchers. This misalignment of purposes would make it unideal for instructional leaders to use.

Science Instructional Practices

The Science Instructional Practices instrument (SIPS; Hayes et al., 2016) incorporated elements of NGSS-defined best practices for science teaching. Specifically, the SIPS instrument gave special consideration to the SEP as defined in the NGSS. This instrument was developed by researchers as an element of an NSF funded project in California that studied PD. The instrument measured the range of science instructional practices that third through tenth grade science teachers implemented as part of a changing landscape of science education (Hayes et al., 2016). When it came to choosing items for inclusion on SIPS, Hayes et al. (2016) defined four priorities: 1) conceptual relevance to the literature they previously reviewed, 2) consistency across multiple existing survey instruments, 3) mid-level in specificity (neither too specific nor too general in reference to instruction), and 4) broadly applicable (Hayes et al., 2016, p. 144).

The final version of the SIPS instrument contained 31 items in two sections. Both sections are self-reported by teachers and are scored on a Likert scale from 1-5, 1 meaning never, 2 rarely (a few times a year), 3 sometimes (once or twice a month), 4 often (once or twice a week), and 5 daily or almost daily. The first section is composed of 21 items that answer the question “How often do your students do each of the following in your science classes?” The second section includes 10 items that answer the question “How often do you do each of the following in your science instruction” (Hayes et al., 2016, pp. 160-161)? This was a unique feature, to focus specifically on student actions in one section and teacher actions in another.

SIPS is promising in that it is a more recent publication written after the introduction of 3DT as a construct. Additionally, the idea that observation protocols might include items that rate teacher behaviors and that separate items might rate student behaviors might be useful. Ultimately, SIPS was meant to be used as a teacher self-report, making it invalid for use by instructional leaders in its current state, though features of it influenced the construction of PCI-OP.

Analysis of Existing Instruments

None of the instruments reviewed above was used for the purpose of instructional leadership, or teacher professional change. Only one (SIPS, Hayes et al., 2016) used elements of NGSS, but SEP was not one of those elements. Not all authors conducted or reported validation of their instruments (for example, SMOP, Sampson, 2004; EIOR, Luft, 1999). While some instruments might be valid in achieving their stated purpose (for example, STIR, Bodzin & Beerer, 2003; EQUIP, Marshall et al., 2010) none of the stated purposes was directly related to instructional leadership and therefore would not be valid for that purpose.

Teacher Observation and Evaluation in Virginia

The Virginia Department of Education (VDOE) does not mandate the use of any one observation tool for administrators to use when evaluating teachers. Rather, the optional use of validated observation protocols is suggested as just one part of a holistic evaluation system. The evaluation system in Virginia lists eight "unified performance standards" for all teachers, listed and described below.

> **Performance Standard 1**: *Professional Knowledge*: The teacher demonstrates an understanding of the curriculum, subject content, and the developmental needs of students by providing relevant learning experiences.

Performance Standard 2: *Instructional Planning*: The teacher plans using the Virginia Standards of Learning, the school's curriculum, student data, and engaging and research-based strategies and resources to meet the needs of all students.

Performance Standard 3: *Instructional Delivery*: The teacher uses a variety of research-based instructional strategies appropriate for the content area to engage students in active learning, to promote key skills, and to meet individual learning needs.

Performance Standard 4: *Assessment of/for Student Learning*: The teacher systematically gathers, analyzes, and uses all relevant data to measure student progress, guide instructional content and delivery methods, and provide timely feedback to students, parents/caregivers, and other educators, as needed.

Performance Standard 5: *Learning Environment*: The teacher uses resources, routines, and procedures to provide a respectful, positive, safe, student-centered environment that is conducive to learning.

Performance Standard 6: *Culturally Responsive Teaching and Equitable Practices*: The teacher demonstrates a commitment to equity and provides instruction and classroom strategies that result in culturally inclusive and responsive learning environments and academic achievement for all students.

Performance Standard 7: *Professionalism*: The teacher demonstrates a commitment to professional ethics, collaborates and communicates appropriately, and takes responsibility for personal professional growth that results in the enhancement of student learning.

Performance Standard 8: *Student Academic Progress*: The work of the teacher results in acceptable, measurable, and appropriate student academic progress. (VDOE, 2021, pp. 7-8)

VDOE acknowledged that not all standards above are elements of an observed lesson. The *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* (VDOE, 2021) includes a feedback form for administrators to share with teachers following a formal observation, which lists the eight standards and elaborates on each with eight-ten bulleted statements that might be observed during a lesson.

Virginia's New Science-Specific Protocol. During the spring of 2023, the Office of STEM Instruction at VDOE developed a tool for administrators in the Commonwealth of Virginia to use specifically with science teachers (A. Petersen, personal communication, June 17, 2023). The Office of STEM Instruction's protocol (*The Science Classroom Observation Tool for School Administrators*) has two sections: one for observed student actions and one for observed teacher actions. The section for student actions has four primary items and the section for teacher actions has five items. Each section contains a question based on the VA SEP. The student section contains the item, "Students use the scientific and engineering practices (SEPs) to gather or critique evidence to explain phenomena or solve a problem" and then listed the six practices in a checklist format. The instructor section contains the item, "Teacher supports students as they engage in the SEPS to gather, make sense of, and/or critique evidence to figure out phenomena." Responses to a pilot test of this instrument during the spring of 2023 were in a checklist format where participants checked off the item if observed. See Appendix K to see the Science Classroom Observation Tool for School Administrators.

While it is understandable that Virginia's evaluation system for teachers does not wish to be overly prescriptive by dictating use of a specific observation protocol, it is helpful for observers who might not be familiar with certain instructional strategies, such as those that are inquiry-based, to have a tool to guide them through observing a teacher's class. VDOE's effort to create a science-specific tool is a step in the right direction. Listing the SEP individually and in greater detail might be a more advantageous format for such a tool, which was a potential format explored during the development and testing of PCI-OP.

Supervising and Supporting Science Teachers

In this section, I review literature related to inquiry that also discusses people or programs that provide help to science teachers. Some school districts employ administrators who supervise all science teachers in a district (Montgomery County Public Schools, n.d.). I will refer to these professionals as science supervisors. In addition to science supervisors, many individual schools have science department chairpersons who provide leadership to science teachers. Apart from individuals employed by individual schools and school districts, support for science teachers might come in the form of professional development from outside sources, such as teacher induction programs (Heredia & Yu, 2017; Navy et al., 2019).

Personnel

Aoki et al. (2005) defined science supervisors as both science department chairs and district science supervisors for a study that focused on what science supervisors know about inquiry. They administered a survey to the science supervisors. Findings indicated that the participating science supervisors had an incomplete conception of inquiry as defined in publications like the NSES and *Science for All Americans* (Aoki et al, 2005; Rutherford & Ahlgren, 1990). Trends in the data suggested that the participants adequately described and identified some components of inquiry but struggled with others. Aoki et al. (2005) suggested that the incomplete understandings of inquiry and the influence of science supervisors on the teachers they supervise may negatively affect classroom practices of science teachers. Aoki et al. (2005) concluded their work by stating that they believed the research done about science supervisors prior to the publication was inadequate for understanding their perceptions of science teaching. They suggested more work needs to be done to increase understanding of science supervisors.

Melville et al. (2016) conducted a longitudinal narrative study of a science department chair over the course of 15 years (other studies were also embedded within that time period). Melville at el.'s (2016) work with Doug (third author and the department chair participant) followed his efforts to be a leader in his high school science department. The research team considered Doug's department to be a successful one in terms of adopting "reform-based" practices. Melville et al. (2016) concluded that there were three primary implications for science department chairs. The first was that department chairs need to know about more than just science content in order to make connections to other pertinent areas, such as assessment. The second was that department chairs should be actively involved with recent developments within the field of science education. Melville et al.'s (2016) reason for this was that it would help the department chair adequately gauge the professional development needs of their department considering new initiatives such as NGSS. The final implication was that department chairs should have a focus on continuous improvement, which involves knowing where each teacher is developmentally and what work can reasonably be done to collaborate with one another and make improvements to practice.

Aoki et al., 2005 and Melville et al., 2016 suggested that it is not only teachers who need support when it comes to inquiry and science instruction. Personnel such as science supervisors and department chairpersons need specific training aimed at staying up to date on research in the field and initiatives and resources available to teachers.

Science Teacher Induction Programs

Navy et al. (2019) studied the experiences of beginning secondary science teachers who attended a state science teachers' conference. All participants were involved in a two-year, science-specific induction program. While inquiry was not the sole focus of the article, one of the research questions Navy et al. investigated had to do with the alignment of the PD the teachers experienced at the conference with the teaching instruction they received in both their induction program and their preservice programs. Inquiry was included as one of the concepts that the participants experienced in their prior PD experiences. The authors found that the teachers did find the PD they received from parts of the conference to align with pedagogy and strategies they had received instruction on previously. Furthermore, subsequent classroom observations of the teachers showed that some of them were implementing lessons and strategies they learned at the conference into their classrooms. Navy et al. theorized that the success of the conference for these participants was largely due to their ability to choose their own workshops and sessions, which allowed for a differentiated experience based on factors like their own interests and recommendations of mentors (Navy et al., 2019). Only one less positive experience was discussed, "The only misalignment teachers identified was with their school departments or contexts. The teachers felt some of the ideas would be difficult to implement without the support of other teachers in their departments" (Navy et al., 2019, p. 424).

Heredia and Yu (2017) described a teacher induction program offered by a science museum. The researchers found that informal science institutions may be uniquely situated to offer support to teachers to implement inquiry, as well as other science specific professional development needs that teachers may have. Participants in the teacher induction program offered by the science museum reported that they benefitted from the access to like-minded science teachers as well as museum scientists, support in learning science content they might not already know, and use of science instructional strategies. Heredia and Yu (2017) discussed inquiry at length, and cited concerns about beginning teachers not having learned science by inquiry-based methods as one reason they might need more robust science-specific induction programs. Follow-up research with former participants of the induction program showed that 73% of them were still teaching, and that 81% of them reported taking on leadership roles at their schools (Heredia & Yu, 2017). As part of their discussion, Heredia and Yu (2017) suggested that further collaboration between school systems and researchers is needed to continue to find ways to fill out teacher induction experiences for science teachers.

Navy et al. (2019) and Heredia and Yu (2017) indicated that teachers can make connections across the field of science education to find support for inquiry, such as at local conferences and informal science institutions. However, Navy et al. (2019) also suggested teachers need to feel as though their colleagues in their schools are on board with new initiatives in order for those initiatives to be successful.

School Administrators and Instructional Leadership

The Framework identified some barriers to the improvement of conditions for students and teachers (NRC, 2012). They suggested that one sign of positive change in science education will be more comprehensive training for administrators to work specifically with science teachers (NRC, 2012). A previous NRC publication written by science educators called for more attention to be shed on the role of teacher evaluation as it specifically applied to science teachers (NRC, 2000). The authors posited that teacher evaluation practices had the potential to be a mechanism for support for teachers to implement inquiry in their classrooms (NRC, 2000). Some elements of this support were 1) an understanding of what it means to teach and learn by inquiry, 2) an understanding of the research-based advantages of such teaching methods, 3) an understanding of the changes that must take place for teachers and students when transitioning to inquiry from a more traditional non-inquiry based classroom, and 4) the development and use of a comprehensive support system that will encourage staff to grow and develop in their ability to use inquiry. The authors described ways to offer the support they describe, one of which is "teacher evaluation consistent with inquiry teaching" (2000, p. 144).

The NRC (2000) stated that teachers often feel personal concerns about new teaching strategies, among which is the concern "about whether the new teaching strategies will be acceptable to the principal" (p. 147). The report did not advocate for a specific comprehensive support system that will provide teachers with the encouragement and support necessary for the use of inquiry, and they also acknowledged the difficulties involved with any type of institutional change. Advice to administrators was to take these concerns seriously and encourage dialogue among faculty and staff.

The manuscript concludes by describing in more detail why teacher evaluation that aligns with inquiry is so crucial for its implementation and endurance. The authors stated that "an evaluator must understand inquiry to know what to observe in the classroom" (NRC, 2000, p. 151), and suggested that without this understanding, evaluators might see inquiry and perceive it as disorganization. Additionally, the authors warn that if changes are not made to teacher evaluation to reflect what classrooms that implement inquiry look like, this will only cause problems in the future. Various publications from the NRC (2000; 2012; 2013) indicated that there is a need in science education for greater involvement from administrators and other instructional leaders. There is a gap for them to step in to provide support to science teachers.

Content Specificity and Instructional Leadership

Research involving instructional leadership within the context of science education is needed to better understand the role of instructional leaders and the influence they exert on science teachers. The existing literature that addresses the relationship between instructional leadership and content primarily addresses how content specificity is missing from teacher evaluation and instructional leadership.

Hill and Grossman (2013) argued that teacher evaluation has been somewhat ineffective for distinguishing teacher quality in the past and that administrators should resist the urge to continue using streamlined and oversimplified teacher evaluation processes, since teaching is a very complex act that depends on context such as grade level and content area. They discussed the idea of grain size, which is the level of specificity in an observation instrument. For example, if an instrument had indicators specific to content area, that instrument would have a finer grain size than instruments that can be used in any content area. They argued that grain size is a major determining factor in the usefulness of the feedback teachers receive during teacher evaluation practices, and they believe this will in turn result in improvement of teaching practices. While acknowledging that major changes to teacher evaluation would be a difficult and slow process, they described in detail several reforms that they believe are necessary for improving the quality of teacher evaluation and instructional leadership. In addition to specificity of content area and finer grain size, they called for: a) the involvement of content area experts to be part of the observation instrument writing process; b) robust training for evaluators to use new teacher evaluation instruments; c) increased frequency of teacher observations; d) alignment of narrative feedback with numerical scores that might result from observation instruments; and e) making other methods of content area support available to teachers, such as content specific professional development and instructional coaching (Hill & Grossman, 2013).

Lochmiller et al. (2012) reviewed pertinent research with the intent of proposing recommendations for educational leadership programs to train principals who wish to become specialized leaders in math and science. They discussed at length how there has been virtually no empirical research related to subject-specific leadership skills in principals. Lochmiller et al. argued that “one of the keys to improving math and science performance in the nation’s schools involves developing principals who understand how school conditions influence efforts to improve math and science instruction that increase student achievement” (Lochmiller et al., 2012, p. 199). The authors summarized existing research by stating five characteristics of effective leadership in math and science education. Among those five things was an emphasis on “the principal’s role in encouraging the adoption and use of project-based or inquiry-based student learning” (Lochmiller et al., 2012, p. 204). They described this as helping to foster classrooms that encourage exploration and making connections to the world. Lochmiller et al. also mentioned that this will involve a principal encouraging and supporting a teacher to take pedagogical risks.

After summarizing the characteristics of effective math and science leadership, the authors used the research they reviewed to propose three suggestions for how to design educational leadership programs that will foster better math and science leadership skills. Lochmiller et al. (2012) described the features of each suggestion in detail, and they analyzed the strengths and weaknesses of each approach. Their suggestions involve tailoring educational leadership curriculum to include math and science content, using disciplinary faculty members in educational leadership courses (such as a science education specialist in addition to leadership faculty), and offering content- specific programs where all aspects of educational leadership training would center around math or science. Lochmiller et al. (2012) concluded their piece by acknowledging the legitimacy of some of the skepticism that may meet their proposals, but also by countering it with research that suggests that current programs that train principals and other administrators are not meeting the needs of schools, teachers, and students.

These calls for instructional leaders to consider the content area of teachers (Hill & Grossman, 2013; Lochmiller et al., 2012; Stein & Nelson, 2003) suggest that teacher evaluation practices that are specific are most helpful. Instructional leaders who understand the specific nuances of science content may be able to offer better instructional leadership to their teachers. The incorporation of content specificity when supplying feedback on instructional practice would be advantageous to teachers and students.

Specific Experiences of Administrators in Science Education

In this section I reviewed specific studies describing interactions between school administrators and science teachers. Searles and Kudeki (1987) conducted a quantitative study about science teaching in schools in Quebec, Canada. The researchers analyzed the results from a survey given to both teachers and administrators and drew comparisons between the teachers' definition of a good science teacher and the administrators' definition of a good science teacher. They found that there was not much difference between the responses of teachers and the responses of administrators. However, of the 38 questionnaire items that were rated the highest, only five have to do specifically with science teaching. Those few items refer to lab activities, generating hypotheses, and science processes, which are components of inquiry. The highest rated item that had to do with science teaching was rated thirteenth highest. Searles & Kudeki (1987) did not publish the items that rated outside of the top 38, so it is not known how many may be related specifically to inquiry or science teaching.

In a case study, Bradford and Braaten (2018) followed a middle school science and math teacher for a period of two years. In the beginning of the study, their participant, “Jenny,” had great self-efficacy, and was, by many measures, considered to be a great teacher. Her school district then started implementing new teaching strategies, such as Close Reading and text annotation. Jenny felt the strategies didn’t fit into math or science easily, therefore conflicting with her prior inquiry practices and the teaching philosophy that had been working for her. The new strategies Jenny was required to implement were generic to all subject areas, and an administrator said this was an intentional move, so that students start to be able to expect a routine in each of their classes. This concerned Jenny, because she felt strongly that science and math require certain teaching techniques that she began to feel she could not use anymore.

To complicate the introduction of the new strategies, the district simultaneously employed new teacher evaluation techniques where an evaluator did unscheduled walk-through observations for a ten-minute period every few weeks. Jenny’s biggest frustration was often that there was no follow-up to these walk-throughs. She felt that she had no recourse for explaining what she was doing and why she was doing it, and that a ten-minute period was hardly sufficient to get an accurate picture of her lesson for that day. (Bradford & Braaten, 2018)

Over time, Jenny’s morale and attitude towards teaching became markedly different. Bradford and Braaten (2018) attributed this change to a conflict between her previous ideas of what good science teaching was, and the apparent idea of good teaching adopted by her administrators and district. When confronted with frustration from teachers, Jenny's principal promised that the new program was going to be better for students and teachers. The authors believed that many teacher evaluation models such as the one used in Jenny's district encourage "performativity," which is when teachers only enact a certain practice because an evaluator is watching. Bradford and Braaten (2018) conclude in their case study of this initially excellent and motivated teacher, that some teacher evaluation systems don't encourage growth and instead contribute to low morale, thus having a reverse effect from their intended purpose. The authors make some recommendations for teacher evaluation models that would encourage meaningful professional growth, particularly those that encourage cooperation between instructional leaders and science teachers.

In a qualitative multi-case study in secondary math and science, Lochmiller (2016) conducted a thematic analysis of data collected from more than 50 participants, including both teachers and administrators. Participants were from five different high schools and represented a wide variety of characteristics, such as age and level of experience in their field. Lochmiller (2016) discussed a push within teacher evaluation for the use of data, but that it has been very broadly instituted on a school or district-based level, not so much at the individual classroom level (which would be more useful for teachers). Lochmiller found that administrators often reported that good teaching is not content area specific- that good teaching transcends all content areas. At the same time, the teachers in his study desperately wanted content area specific feedback and support. However, some of the teachers felt that even administrators with math or science teaching experience did not give great feedback, with one teacher suggesting that they felt like the administrator mindset had taken over and what they knew about good teaching within their discipline had been replaced with district-wide values instead. (Lochmiller, 2016)

Lochmiller (2016) ultimately concluded that most administrators see teacher evaluation almost exclusively through the lens of their own prior teaching experiences and content area training, and that this significantly limits the type of feedback teachers receive. He also made a strong argument regarding the need for more research in this area to know what leadership in different instructional areas should look like.

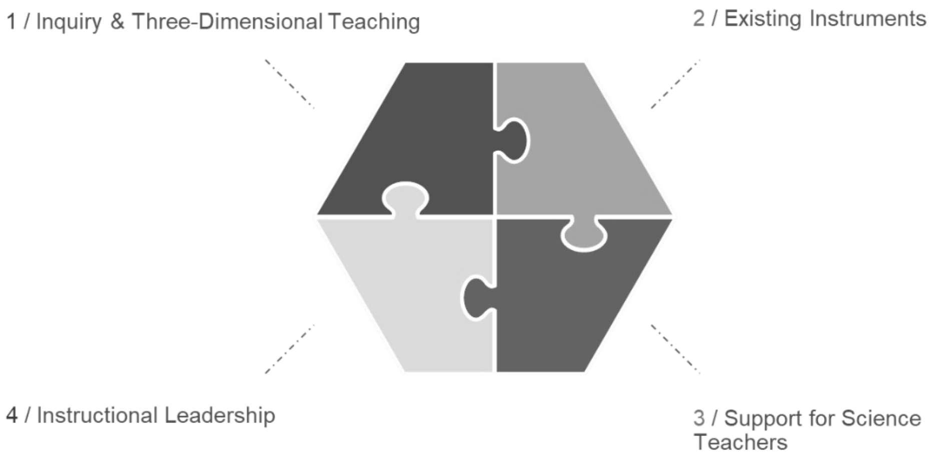
Teacher evaluation experiences as they related specifically to science education were present in a small segment of the literature, and some of their findings differed. Several pieces agree that there have been disconnects between administrators and science teachers and recommend that more research in this area is needed. Science teachers sometimes report that they do not feel supported by their administrators. Multiple studies suggested that sometimes administrators do not think content specific skills are important for a science teacher to be effective in their jobs.

Chapter Summary

The literature examined in this chapter is represented in Figure 5 in the form of a puzzle. When viewed through the lens of instructional change and growth, this is the backdrop for the study detailed in Chapter 3. Four categories of literature and background information were presented in Chapters 1 and 2: 1) inquiry and 3DT, 2) existing instruments that measure inquiry, 3) existing support for science teachers, and 4) the current landscape of instructional leadership.

Figure 5.

Interactions Between Literature



Interactions between some of these puzzle pieces already exist in the literature reviewed previously. For example, the EIOR instrument (Luft, 1999) discussed the use of EIOR to measure inquiry-based science teaching for teachers involved in professional development. This involved elements of the first three puzzle pieces in Figure 5, but did not touch on the fourth puzzle piece, Instructional Leadership. The research conducted for this dissertation sought to focus on the nexus of all four areas. Viewed through leadership content knowledge (Stein & Nelson, 2003) and teacher professional growth (Clarke & Hollingsworth, 2002), the information gleaned from the studies reviewed in Chapter 2 suggest that there is an existing gap at the nexus of those four areas. PCI-OP was developed to determine if an SEP focused observation protocol shows a potential way to bridge that gap.

Chapter 3. Methodology

This chapter describes the steps taken to develop and test a protype observation protocol that focuses on observation of the SEP Planning and Carrying Out Investigations. Additionally, the differences between SEP as stated by NGSS (2012) and how they are represented in Virginia are discussed within this chapter. The prototype observation protocol will focus on a single SEP as a way of providing proof of concept prior to developing and testing a full observation protocol that includes all six VA SEP. “Proof of concept research aims to establish that prototypes achieve functions or results desired of them and to argue for their continued development in settings beyond those in which they were established” (Eliott, 2022).

This research was a mixed methods development study (Creamer, 2018). I discussed the participants and their involvement in this study. I listed the data I collected and described the subsequent analyses that I used to determine proof of the feasibility of an observation protocol instrument facilitating instructional leadership in secondary science classrooms.

Theoretical Influences on Study Design

I drew from three primary methodological sources to design this study: Creamer (2018), who described mixed methods research; Wilson (2005), who described instrument development; and McKenney and Reeves (2013) who described the field of Educational Design Research (EDR), which focused more on practical outcomes than methodological procedures.

Development Designs

The design of this study is a type of mixed methods research generally known as a development design, which Creamer (2018) described as a multiphase sequential study in which an initial phase informs the direction of another. This design is primarily used for the development of an instrument, where “one phase informs or is directly linked to the other”

(Creamer, 2018, p. 29). This multiphase design is sequential, where one phase is completed before continuing to another. According to Creamer (2018), designs of this nature would typically begin with a phase that uses a qualitative focus group followed by a phase that involves the pilot testing of an instrument prototype. This allows the focus group to provide feedback on the format and content of the instrument before pilot testing is done in a subsequent phase.

Instrument Development

Wilson (2005) described a detailed methodology for the development of instruments in the book *Constructing Measures*. Wilson outlines his approach as having four building blocks: the construct map, the item design, the outcome space, and a measurement model (Wilson, 2005, pp.18-19). The construct map is a step that defines and examines the nuances of the construct being measured. The item design step is where items are written. The outcome space is where items are scored and analyzed. The measurement model step is a way to relate the scores on the items back to the construct that is being measured. The measurement model is often, but not always, achieved by use of a statistical model. This study will generally follow this methodology, with some deviation to accommodate for the unique nature of observation protocols when compared to other instruments, such as surveys or cognitive tests.

Educational Design Research

EDR is a field that is focused on contributing to educational practice (McKenney & Reeves, 2013). It is a tradition that is less defined by its methodology than its goals. EDR is rooted in finding pragmatic solutions to complex educational problems. EDR embraces innovation as it seeks to simultaneously accomplish two goals: 1) the generation of research-based solutions that translate well to real-world settings; and 2) the generation of new knowledge and theory that can be used by others in similar settings (McKenney & Reeves, 2013).

EDR is used primarily to generate educational "products, processes, programs, and policies" (McKenney & Reeves, 2013, p.138). Teacher professional development is a prominent application of EDR. The value of EDR is judged by its ability to improve educational practice. While EDR is not confined to any specific methodology, it might be found integrated within qualitative genres such as participatory action research, case studies, and other phenomenological work. It includes practices from quantitative research. Though EDR tends to be an iterative process as it looks for design-based solutions, there are generally three phases of the work: 1) analysis of and orientation to the problem; 2) design and development of potential solutions; and 3) evaluation and retrospection of the interventions (McKenney & Reeves, 2013, p. 143). While this may sound similar to other research paradigms, EDR practitioners McKenney and Reeves (2013) described one difference by saying, "Instead of tossing innovations over the metaphorical walls of classrooms and online learning environments, educational design researchers are working hand in hand with practitioners to conduct design and research in ways that make substantive change possible" (p. 153). The goals of EDR make it a potentially useful guiding principle for encouraging the implementation of 3DT in secondary science classrooms. This study will involve both instructional leaders and teachers in the development of an instrument because they are the potential end-users of the instrument.

In one study that utilized EDR principles, Sekano et al. (2023), investigated the needs of teachers regarding online professional development. The paradigm of EDR helped them determine that teachers needed PD that is "affordable, flexible, intensive, and ongoing" (p. 1) and went on to make recommendations for best practices in PD based on their findings.

Proposal

An essential first step in the development of an instrument is to define the purposes for which it is intended, which is vital for the validation of a tool (American Educational Research Association [AERA], et al., 2014). Most, though not all, of the instruments, tools, and procedures reviewed in chapter 2 are for the purpose of science education experts conducting research. The purpose of this proposed research was to determine the feasibility of using PCI-OP for instructional leaders and science teachers to work collaboratively in implementing inquiry-based instruction in secondary classrooms with greater fidelity and frequency than what may have been previously done. This conflict in purposes means that the existing instruments are not likely to be valid for instructional leaders, mentors, or other colleagues to use within a school. In addition, the current push in the Commonwealth of Virginia for observation tools that reflect VA SEP makes the development of new tools even more relevant.

Purpose

The purpose of this study was to determine the potential for developing an observation protocol tool based on observable features of the SEP. Because this observation process of SEP has never been approached before, and as established earlier, teaching is a complex phenomenon, this study will focus on an observation protocol specifically for one of the SEP: Planning and Carrying Out Investigations. Upon completion of this study, the proof of concept will show whether the protocol is feasible, and whether further development can proceed on the other SEP. The purpose of protocol itself is to provide secondary science teachers with feedback on how SEPs are used in an observed lesson to encourage them to implement 3DT teaching in their classrooms. This purpose aligns well with the principles of EDR because the instrument developed would be used in the professional lives of both teachers and instructional leaders. The abbreviated observation protocol will be referred to as the Planning and Carrying Out Investigations Observation Protocol (PCI-OP). PCI is a practice both in the VA SOLs and in the NGSS. However, since Virginia is the context in which this research will take place, I will focus on it primarily as being a VA SEP.

Research Questions

Two research questions guided this investigation, aimed at developing and testing an observational protocol instrument based on observable features of the Science and Engineering Practices (SEP):

1. Is it feasible to utilize the Science and Engineering Practices as a framework for the development of an observation protocol intended for use by instructional leaders in secondary science classrooms?
2. How do instructional leaders and science teachers assess a science lesson differently, and what are the implications of these potential similarities or differences on the feasibility of the observation protocol outlined in Research Question 1?

Setting

Previously, I mentioned that NGSS has not been adopted by all states in the United States. Although Virginia has not adopted the NGSS, the state has standards that are similar in some ways. For a convenience sample, Virginia served as the setting for this investigation.

Virginia Science Standards. Instead of using the NGSS, Virginia has standards that are referred to as the Standards of Learning (SOLs). The science SOLs for each grade level do include six "Scientific and Engineering Practices" that are very similar to and were influenced by the eight NGSS Science and Engineering Practices (A. Petersen, personal communication, May 19, 2023). These skills are described in specific terms in the Virginia Science SOLs. To start, "investigate" refers to scientific methodology and implies the systematic use of the following inquiry and engineering skills:

- Asking questions and defining problems.
- Planning and carrying out investigations;
- Interpreting, analyzing, and evaluating data;
- Constructing and critiquing conclusions and explanations;
- Developing and using models;
- Obtaining, evaluating, and communicating information. (VDOE, 2018, p. *v*).

The SOLs include those six skills within the standards for every grade level from K-12. Four of the Virginia skills match the NGSS SEP verbatim; the other two Virginia skills are a combination of other SEP. The only NGSS SEP not represented in the VA SEP is practice seven: engaging in argument from evidence (NRC, 2013). However, the Virginia SOLs elaborate on each practice differently, becoming more sophisticated with each successive grade. Some grade levels do mention argumentation from evidence within the other practices. For example, Table 1 compares the differences in elaboration of the VA SEP PCI at the Kindergarten level and the eighth-grade level. Notice that while PCI is a practice to be included at every grade, the bulleted portion elaborates each practice differently. These elaborations were used in drafting preliminary items for PCI-OP.

Table 1.

Virginia Standards of Learning: Descriptions of Planning and Carrying Out Investigations

Scientific and Engineering Practice	Elaboration at the Kindergarten Level	Elaboration at the Eighth-Grade Level
Planning and carrying out investigations	• make observations to collect data • identify characteristics and properties of objects through observations • measure the relative length and weight of common objects • record information from investigations	• independently and collaboratively plan and conduct observational and experimental investigations; identify variables, constants, and controls where appropriate and include the safe use of chemicals and equipment • evaluate the accuracy of various methods for collecting data • take metric measurements using appropriate tools and technologies • apply scientific ideas or principles to design, construct, and/or test a design of an object, tool, process or system

Note. VDOE, 2018, pp. 9; 33

Researcher's Entry, Role, and Ethics

I currently teach chemistry part time at a high school within the Commonwealth of Virginia. I have 18 years of experience teaching at two different Virginia public high schools, where I taught chemistry, earth science, and geology. I have four years of experience as a university supervisor, where I mentored, observed, and evaluated secondary science student teachers. I have led PD for pre-service and in-service teachers on campus at my university and at state and national conferences. While I do not believe the realities of classrooms allow teachers to use inquiry-based methods 100% of every moment, every class, every day, every year, I believe that inquiry is the most important pedagogical paradigm for science teachers to adopt. I am hopeful that the shift in emphasis to 3DT will facilitate science teaching in as many contexts as possible, leading to more inquiry-based teaching and learning in more schools. I am excited about the ways that organizations like NSTA have highlighted 3DT and the robust PD they have been offering that aligns with 3DT. I am currently serving a three-year term on NSTA's standing committee for the Coordination and Supervision of Science Teaching.

My previous experiences with receiving feedback from instructional leaders have generally been positive, and those that were not positive were simply more lackluster than they were negative. As a first-year teacher, I recall that my supervisor described a lesson I delivered as boring. Some years later, I included "increasing my use of inquiry-based teaching" as one of my professional goals. During the required meeting with my principal to discuss my professional goals, the principal expressed concern that some of my students were not ready to be taught using inquiry-based methods and that they might need more hand-holding (her words, not mine) than what she believed inquiry-based teaching allowed. As far as I remember, that was the first negative experience I had with an instructional leader not supporting inquiry-based teaching, and it made an impression on me. In more recent years, my evaluations have all been positive experiences, and my instructional leaders have expressed being very happy with what they saw. However, very little feedback that I have received has ever been specific to science teaching, as opposed to teaching in general.

I will ethically manage my biases by working with a committee, by using expert reviewers for PCI-OP, and by checking in with participants to review their responses and involvement during the focus group phase of this study, to be sure I have represented their views with fidelity.

Data Collection

Data collection for this study involved identifying and recruiting three different groups of participants for each of three different phases.

Participant Selection

This study required the input of both expert reviewers, who helped refine the items, as well as teacher and instructional leaders, who acted as focus group participants and raters who tested the instrument by watching the same pre-recorded clips of science lessons. The criteria and process for identifying, inviting, and recruiting participants were different for different phases and is outlined below.

Expert Reviewers. Phase 1 of this study involved the use of expert reviewers. I initially invited three experts from the field of science education to serve as reviewers (sometimes referred to in measurement literature as informants), who reviewed drafts of the items. Had any of the three declined to be a reviewer, I was ready to contact alternate potential expert reviewers. I recruited reviewers on a purposeful basis, considering 1) a level of familiarity with and previous research with 3DT and SEP, and 2) experience in observing and rating teachers and students in science classrooms. In keeping with the goals of EDR, it was important that this tool be evaluated for its ability to work for professionals who need it, and I emphasized that to the reviewers. I emailed potential reviewers and explained the project briefly, followed by a specific list of the tasks I asked them to do (for example, comment on the type of scale they thought would be best for these items). Reviewers read and commented on the items and returned their comments to me within a designated timeframe. See Appendix C for the invitation email and informed consent, and Appendix B for the survey that were used in Phase 1.

Focus Group Participants. I invited two science teachers and two administrators to participate in a focus group during Phase 2 of this study. I selected secondary science teachers and administrators from the Commonwealth of Virginia by convenience.

Raters. I purposely recruited participants for Phase 3 from the Commonwealth of Virginia by convenience. I recruited any licensed secondary science teachers who currently taught at any public middle or high school in Virginia, regardless of the district in which they were employed. I also recruited any instructional leaders who worked at the time at any public high school in Virginia, any Virginia public school district offices, or the Virginia Department of Education. I contacted potential participants by sending emails and advertisements through social media groups such as the Virginia Association of Science Teachers (VAST), and by personal communication with teachers and administrators at events such as the NSTA Annual Conference, which took place in Kansas City, Missouri, in October 2023; and the VAST Annual Professional Development Institute in Roanoke, Virginia, in November 2023. Participants were also encouraged to pass on the information to others who they thought might also be interested. Recruitment materials can be seen in Appendix J. I obtained as many publicly available email addresses as possible for Virginia secondary science teachers and instructional leaders from school district websites and contacted them by e-mail, inviting them to take part in the project. Instructional leaders and secondary science teachers are the potential end-users of this type of observation instrument; using them as participants is in keeping with the tenets of EDR (McKenney & Reeves, 2013), which directly involves the users of procedures and products in their development, to use participants who would actually use the PCI-OP.

Rater participants actively participated in the study by progressing through a survey I designed, which familiarized them with the prototype instrument and directed them to view and evaluate video clips. A detailed overview of the survey structure is provided in Appendix J. Participants independently completed the survey on their personal computers, answering questions via Question Pro. They had the option to save their progress and revisit the survey later. This module was completed asynchronously before March 4, 2024.

Instrumentation

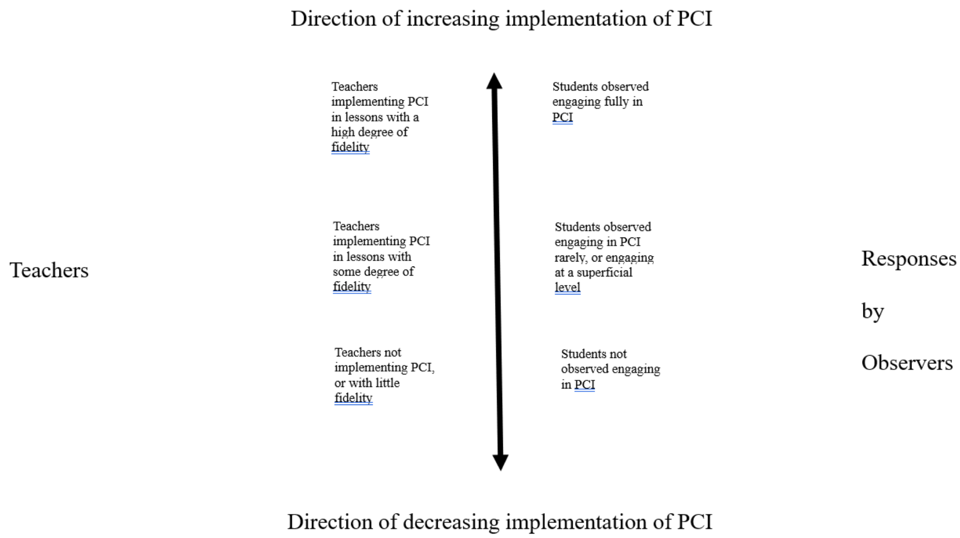
In order to construct a preliminary version of PCI-OP, I followed the procedure for developing instruments as outlined in Wilson (2005), which I previously described.

Construct Map

Construct maps help to define the construct more clearly for the researcher. The construct this observation protocol measured concerns the implementation of the scientific and engineering practice (VA SEP) "Planning and Carrying Out Investigations," as described in the Virginia SOLs (VDOE, 2018), and depicted in Figure 6.

Figure 6.

Construct Map of PCI Implementation



Instrument Blueprint and Item Design

I constructed a test blueprint (see Appendix M) consisting of three categories with the following number of items:

Category 1: Proposing, Planning, or Comparing Different Ideas for Investigations: 4 items;

Category 2: Identifying or Discussing Different Components or an Investigation: 8 items;

Category 3: Students Taking Investigative Actions: 10 items

I began with more items than I wanted on my final observation protocol to test each item's effectiveness. Wilson (2005) said the following regarding developing items: "In many, if not most cases, the construct is not clearly defined until a large set of items has been developed and tried out with respondents" (p. 42). Additionally, some items were redundant due to the need to test the ability of multiple items to measure a dimension of the construct. The preliminary version of PCI-OP had 22 items.

Items were in the form of statements that described the behaviors of students and teachers that might be observed in a secondary science classroom. I sought feedback from experts regarding the number of points the items should be rated on. For example, some experts indicate that a simple "observed" versus "not observed" choice would be sufficient. Other experts, in contrast, preferred the availability of more than two choices, such as "fully or frequently observed," "observed incompletely or briefly," or "not observed."

Preliminary Items. See Appendix H.

Methodological Plan

There were three sequential phases of the study, detailed as follows.

Phase 1: Expert Reviewers

This phase focused on soliciting feedback from three expert reviewers on the structure and content of PCI-OP. Expert reviewers were contacted as described above, using the recruitment materials in Appendix C. I sent a copy of the preliminary items for PCI-OP in December 2023 so they could study it. I then directed them to a Question Pro (QP) survey. I identified the expert reviewers as Expert Reviewer #1, Expert Reviewer #2, and Expert Reviewer #3. In the survey, the first task was for the expert reviewers to specify how they preferred their qualifications to be described in the research report; these descriptions are detailed in Chapter 4. The second task required them to indicate which items they would retain on the protocol, which ones needed editing, and which ones should be removed entirely. For the third task, they provided comments and edits on individual items. They also commented on the structure of the scale used in the fourth task. Finally, any additional feedback not previously provided was collected in the fifth task. The survey itself is included in Appendix B.

These tasks were completed by December 15, 2023. Subsequently, on January 2, 2024, I shared an edited version of PCI-OP with the expert reviewers and requested any final feedback by January 8, 2024. The outcomes of their feedback are discussed in Chapter 4 and can be accessed in Appendix D.

Phase 2: Focus Group

The aim of this phase was to conduct a focus group with participants to explore their cognitive processes while using the PCI-OP instrument with video recordings of science lessons. Participants were recruited as previously described, during January 2024. Upon volunteering for the focus group, participants were directed to a Question Pro (QP) survey where they chose a pseudonym for reporting purposes and answered questions regarding their professional background. This survey is available in Appendix I.

The focus group convened on January 28, 2024, with lunch provided for the participants. The focus group protocol is detailed in Appendix I. Participants introduced themselves, and I provided an overview of the project and the objectives of Phase 2. Paper copies of PCI-OP were distributed, and we viewed video clips of science classes, pausing frequently to discuss how an observer might complete PCI-OP if observing the lesson live. We also tested the web-based version of PCI-OP, identifying any potential issues raters might encounter. The focus group session was audio recorded, and I used Microsoft Word to transcribe the recording. We discussed possible modifications, revisions, and any concerns participants had regarding PCI-OP. Focus group participants received lunch and $20. In reporting the focus group findings, I employed member checking to verify that I accurately represented the participants' perspectives and words. This process involved following up with participants to ensure the accuracy of their views as presented in the report.

Phase 3: Raters

Following Phase 2, I edited PCI-OP again. I began recruiting participants for Phase 3 in January 2024, but made sure they knew data collection for this phase would not begin until February 9, 2024. I recruited via the methods previously discussed, and informed potential participants that as an incentive, they could enter a drawing for one of eight Amazon gift cards (two worth $50 each and six worth $20 each). The drawing took place after data collection was completed. The purpose of this phase was to conduct a pilot test using the most recent version of PCI-OP with as many Virginia science teachers and administrators as possible. Volunteers for Phase Three were asked to complete three tasks. They were directed to a QP survey where they completed some questions about their professional background. The second task was to view three separate 15-minute video clips of science lessons while completing the most recent version of PCI-OP. The raters started a new copy of the protocol with each new video clip. The third task was to reflect on their experience using the PCI-OP and comment on its future feasibility. Raters also had the opportunity to comment on each individual item in the version of PCI-OP that they used. At the end of the third task, participants entered their email address in the QP survey if they wished to enter the drawing for incentives. All three tasks were in the same QP survey, so that demographic and background questions could be used for disaggregation. The QP survey used in this phase can be found in Appendix J.

Data Analysis Strategies

Data collected during the first two phases of the study were primarily used to edit the PCI-OP used during Phase 3. From the data in Phase 3, I calculated the means for each item by group—science teachers and instructional leaders—after which I conducted t-tests to answer RQ 2. Data collected during Phase 3 was also analyzed by calculating separate values of Cronbach's alpha for each video, as well as by preparing a classical item analysis. These analyses provided a robust answer to RQ 1. I also used the data from the second and third phases to propose future expansion of this work, such as a longer protocol with items that align with other VA SEP, and feasibility of using PCI-OP in other contexts, such as at the elementary level.

Chapter Summary

I wrote a prototype instrument, the Planning and Carrying Out Investigations Observation Protocol (PCI-OP), which preliminarily consisted of twenty-two items in the form of observable statements and using a Likert scale. I used expert reviewers to provide feedback on PCI-OP and subsequently edited. I conducted a focus group with science teachers and administrators while they used PCI-OP to rate a science lesson. I edited PCI-OP a second time based on their experiences. I recruited science teachers and instructional leaders to pilot test the PCI-OP by rating videos of science lessons. I solicited reflective feedback from the raters to make a determination regarding the feasibility of using PCI-OP in the future.

Chapter 4. Findings

The purpose of this study was to determine the potential for developing an observation protocol tool based on observable features of the SEP. Only one SEP was chosen for this study to determine whether such a protocol would be feasible, and whether further developments could proceed on the other SEP. Two research questions guided this investigation:

1. Is it feasible to utilize the Science and Engineering Practices as a framework for the development of an observation protocol intended for use by instructional leaders in secondary science classrooms?
2. How do instructional leaders and science teachers assess a science lesson differently, and what are the implications of these potential similarities or differences on the feasibility of the observation protocol outlined in Research Question 1?

Findings from this study are organized by sequential phases.

Phase 1: Expert Reviewers

Expert reviewers were contacted during the first week of December 2023 via the means described in Chapter 3 and detailed in Appendix C. Five reviewers were initially contacted and three ultimately agreed to participate in this study, each of whom completed Phase 1 of data collection. Reviewers were asked to describe their qualifications in their own words; Reviewer #3 requested that their personal details not be included in this study. Descriptions can be found in Table 2.

Table 2.

Qualifications of Expert Reviewers

Reviewer	Self-Identified Qualifications of Reviewer
Expert Reviewer #1	K-12 science education for 30 years; 23 years of secondary classroom teacher; 8 years of division level leadership; 8 years of state leadership experience; led state science standards development and implementation; published on a variety of topics to include science observation tools and STEM education.
Expert Reviewer #2	Science teacher for 10 years; 2016 Presidential Award for Excellence in Mathematics and Science Teaching; 2015 [STATE] Science Teacher of the Year; nationally recognized Science Education speaker, High School Life Science Director for Association of Presidential Awardees in Science Teaching.
Expert Reviewer #3	Thirty years experience as secondary science teacher, secondary administrator, county administrator; state and national professional organization leader; NGSS writer.

Expert reviewers were given a link to the first Question Pro (QP) survey (see Appendix B), in which they described their reactions to instructions for the instrument, the type of scale for each item, and other stylistic and format concerns. Two reviewers indicated that a Likert scale of one to three was optimal, while the third reviewer did not have a strong preference except that they believed that there should be a space to make note of evidence, regardless of the type of scale used. All three reviewers felt strongly that information regarding the frequency of which item might be observed should be discussed. Additionally, Reviewer #2 felt strongly about being able to remind users of PCI-OP that when observing science classes, one should look for SEPs to be rooted in phenomena, stating, "Focus on phenomena-driven instruction. This is what everyone gets wrong. They crosswalk old learning standards with new DCIs [disciplinary core ideas] and forget there are these other dimensions [i.e., SEP and CCC] present that are necessary to explaining phenomena." This quote from an expert reviewer underscores the importance of integrating SEP in 3DT and validates the importance of establishing SEP as a viable basis for administrators to be looking for in science classes.

As a result of the feedback received from the first QP survey from expert reviewers, I wrote an introduction for PCI-OP that discussed the importance of considering phenomena-based instruction, as well as encouraged observers to consider the limited context that one single observation using PCI-OP would allow an observer to make. I constructed individual items with a three-point Likert scale and left space available for comments or evidence that would add valuable context to the rating given. An excerpt of the format can be seen in Figure 7 and the full first draft of PCI-OP can be found in Appendix D.

Figure 7.

Excerpt of Formatting for PCI-OP Draft v.1

1. Students planned investigations designed to answer a question or questions.

☐ **Not Observed**	☐ **Observed Partially**	☐ **Observed Fully**
Evidence and/or Comments:		

2. Students had the opportunity to propose investigations they are interested in.

☐ **Not Observed**	☐ **Observed Partially**	☐ **Observed Fully**
Evidence and/or Comments:		

The second QP survey for expert reviewers (refer to Appendix B) was circulated after the completion of PCI-OP draft version one. In this subsequent survey, expert reviewers provided feedback on individual survey items, indicating their preference to either keep, delete, or revise each item. The outcomes are detailed in Table 3. As per the expert reviewers' assessments, seven out of 22 items received unanimous votes to retain, while no items received unanimous votes for deletion.

Table 3.

Expert Reviewer Ratings of Individual Survey Items, PCI-OP v. 1

Item	Keep, no revisions	Keep, revise	Remove
1. Students planned investigations designed to answer a question or questions.	2	1	0
2. Students had the opportunity to propose investigations they are interested in.	1	2	0
3. Students identified or discussed materials necessary for a certain investigation.	2	1	0
4. Students identified or discussed the variables in a certain investigation.	2	1	0
5. Students conducted investigations.	2	0	0
6. Students were asked to construct hypotheses or predictions.	0	2	0
7. Students identified or discussed measurements to be collected in a certain investigation.	2	0	0
8. Students observed and described phenomena (e.g. weather events).	2	1	1
9. Students sorted or categorized phenomena (e.g. how animals move).	1	1	1
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	0	2	1
11. Students practiced using lab equipment (e.g. beakers, Bunsen burners).	1	0	2
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	1	1	1
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	1	1	1
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	0	2	1
15. Students recorded data or information from observations.	3	0	0

Item	Keep, no revisions	Keep, revise	Remove
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	3	0	0
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	3	0	0
18. Students identified or discussed the data that should be collected to answer a testable question.	3	0	0
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	3	0	0
20. Students compared the accuracy of different methods of collecting the same data.	2	1	0
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	3	0	0
22. Students analyzed or evaluated data or information they collected.	3	0	0

Following completion of this second QP survey by the expert reviewers during Phase 1, I made minor revisions and changes to the items in PCI-OP. Version 2 can be found in Appendix E. Many of the expert reviewers' comments regarding items they preferred to see deleted centered around redundancy of various items. For example, Expert Reviewer #1 said, "A lot of items that are very similar and these do not build. In other words, they recorded data in 15, discussed tools in 16, the determined data in 18?" I chose to delay any deletion of individual items until after Phase 3 to pilot test all of them and compare their performance. This approach enabled me to select the best items to keep among the ones that were found to be redundant. This issue will be discussed in Chapter 5.

Phase 2: Focus Group

Potential focus group participants were contacted by e-mail during the first week of January 2024, as described in Chapter 3. Four focus group participants were ultimately recruited.

Demographic information regarding the four participants can be found in Table 4. Note that all questions about age, race, and gender identity were open-ended in format, giving participants flexibility to choose their own terms for their identity. The prior professional experience of the focus group participants is provided in Table 5. The focus group participants were similar in that they were all female. Two of the participants were high school science teachers and two were middle school administrators.

Table 4.

Demographic Characteristics of Focus Group Participants

Pseudonym	Age	Gender	Race or Ethnicity	Current Profession
Neicy	54	Female	African American	School Administrator
Katerina	39	Female	Black	School Administrator
Claire	44	Female	White	Science Teacher
Suzanne	33	Female	White	Science Teacher

Table 5.

Previous Professional Experience of Focus Group Participants

Participant	Summary of Professional Experience
Neicy	Eight years of experience as a secondary school administrator in Virginia (5 in middle school and 3 in elementary school); 15 years of experience as a middle school teacher (8 years in history and 7 years in math).
Katerina	Six months of experience as a secondary school administrator in Virginia; one year of experience as an elementary school administrator; one year of experience as an instructional coach; four years of experience as a student teacher supervisor; ten years of experience as an elementary school teacher.

Participant	Summary of Professional Experience
Claire	Ten years of experience as a secondary school science teacher in Virginia and eight years of experience as a secondary school science teacher outside of Virginia.
Suzanne	Eight years of experience as a secondary school science teacher in Virginia.

The focus group met on January 28, 2024. The group spent approximately 15 minutes chatting while waiting for all participants, distributing lunch, and distributing materials. They then reviewed the purpose of the project and the role of the focus group within the study. They were then shown two clips of one video from *Ambitious Science Teaching,* who granted permission for their videos to be used in this research (personal communication, J. Thompson, October 8, 2023; see Appendix M). Focus group participants first viewed five minutes of the video under the heading "Building a consensus model," followed by a pause and discussion about what was observed, and how PCI-OP could be used as an observation tool by administrators. They were then shown an additional four minutes of the same video, followed by more discussion and suggestions for PCI-OP. The focus group lasted one hour and two minutes in total.

Focus Group Themes

Following the focus group, I began the process of inductively analyzing the data. I began with listening to the recording of the focus group to familiarize myself with the flow of the event and the types of comments that participants made. I then transcribed the recording using Microsoft Word and corrected any errors by listening to the recording while reading the transcripts. I categorized comments into emerging themes, using an iterative process to refine and recategorize excerpts of discussion.

One prominent recurring theme emerged from the focus group, which centered around concerns regarding the context of the lesson within the sequence of events in a classroom. This theme informed the bulk of the revisions, leading to the creation of PCI-OP Version 3 (see Appendix F). Other minor themes that did not result in any PCI-OP revisions were also noted. All themes are discussed below.

Focus group participants found it especially important to discuss the context of the lesson within the larger picture of what might be taking place over time in a classroom. For example, the participants discussed whether a lesson might be an introduction to a concept or a follow-up activity after multiple days of other learning activities about the same content. Both teachers and administrators felt that it would be helpful to know how much background the students had and what expectations the teachers had for any given day’s lesson. The topic originated when Suzanne expressed concerns with the lesson itself, saying,

> She [the teacher] did that thing where they want the kids to come up with the answer, but the kids don't know when to come up with the answer, and that always annoys the crap out of me when they want people to do it that way. It's like I get the whole point of science and you're supposed to learn, but also they have to have some background. You have to help them get there to a certain extent and I kind of think she was letting them flounder a little too long.

After a few moments of back-and-forth conversation, Katerina added,

> Yeah, I think that's one of the things. That when you're coming in, just seeing a snapshot, we don't know what she [the teacher] talked to them about before. So like she may have been teaching on this for a couple days. And so she expected them to be like they could

> give her an answer. But we don't know. That's like…(pause) as administrators could be at the time to want to see what their unit plans are.

Conversation continued after this point, centering around different ways the teacher in the video might have structured their unit, which of course would have been unknowable to the outside observer. Later, I asked Katerina to comment as an administrator on what types of questions she might want to ask the teacher either before and or after a lesson such as this. She responded that knowing some things about inquiry would be useful, adding that she knew some teachers did not follow methods that are as inquiry-based as others; thus, Katerina indicated that she would want to know more about what methods the teacher was following. The group discussed the importance of knowing the type of scaffolding a teacher had provided and the expectations of students at the end of any lesson.

I used these findings related to lesson context to add a short pre-observation and post-observation section to the third version of PCI-OP, which the teacher would fill out. The pre-observation section was worded as follows:

> To be answered by the teacher prior to the class being observed. Where does this lesson fall within the unit/concept that you're currently teaching? For example, is it an introduction to the concepts you'll be teaching over the next week? Is it a review after several days of developing conceptual understandings? Something in-between?

The post-observation section was worded thusly:

> To be answered by the teacher after the class that was observed. Where are you and your students going with these concepts from here? For example, do you plan to revisit some of the concepts they developed during this lesson and address possible misconceptions?

What learning events will you be using next? Or maybe you plan to assess this concept next and then potentially move on to new material?

Three additional themes emerged from focus group discussions. First, participants stressed the importance of being aware that specific terminology that might have a different meaning to scientists versus non-scientists (such as *accuracy* and *phenomenon*). Second, the administrators in the group appreciated that some of the items included specific examples given as a guide for observers, and that more explicit questions would be more helpful to an administrator. Third, the teachers in the group, who were both life science teachers, suggested that equal attention should be given to life science examples as to physical science examples.

Phase 3: Raters

Teachers and administrators were recruited beginning in February 2024 via social media, word of mouth, and by collecting email addresses where available from school and school district websites. Approximately 1900 Virginia teachers and administrators were contacted; all of whom were directed to a website conveying essential information about this study (see Appendix G). Question Pro (QP) data indicate that the survey link was accessed 348 times; 146 respondents began the survey; and 13 respondents completed every page of the survey and clicked "finish" at the end. A brief outline of the survey is as follows:

Task 1: Demographic Section

Task 2: Video Section (comprising Videos 1, 2, and 3)

Task 3: Reflection Section

Survey Completion and Respondent Demographics

Of the 146 respondents who began the survey, 94 completed Task 1 (demographics); a total of 32 respondents started and completed the first video of Task 2; a total of 21 respondents started and completed the second video of Task 2; and a total of 18 respondents started and completed the third video of Task 2. A further 14 respondents started and completed Task 3 (the reflection section), but one respondent did not click the final "finished" button at the end of the survey. This completion data indicate that the biggest drop-off point was between Task 1 (demographics) and Task 2 (video ratings). It should be noted that all data were captured for all respondents who completed an item, whether they completed the full survey or not. Tables 6 and 7 summarize demographic characteristics for respondents, separated by teachers and administrators.

Table 6.

Demographic Characteristics of Phase 3 Raters: Teachers

n=	Gender	Race or Ethnicity	Age	Years of Teaching Experience
86	76 female or woman (88%); 9 male (10%); 1 no response (1%)	1 Filipino (1%); 5 Black or African-American (6%); 6 Asian or Asian-American (7%); 6 Hispanic or Latino/a (7%); 68 Caucasian, White, or European (79%)	Mean: 43.4 Range: 23-64	Mean: 12.8 Range: 1-33

Table 7.

Demographic Characteristics of Phase 3 Raters: Administrators

n=	Gender	Race or Ethnicity	Age	Years of Administrator Experience
8	5 female (62%); 3 male (38%)	1 Black (12%); 7 Caucasian or White (88%)	Mean: 45.1 Range: 39-56	Mean: 8.5 Range: 8-18

Table 8 summarizes respondents' opinions on various aspects of inquiry and 3DT.

Table 8.

Familiarity with Science Education by Percent of Respondents

	Teachers (n=86)	Administrators (n=8)
Have not heard of inquiry	0%	0%
Have heard of inquiry	25.6%	37.5%
Well-versed in inquiry	48.8%	12.5%
Inquiry is not helpful to students	0%	0%
Inquiry is helpful to students	57%	75%
Have not read 2018 SOLs	4.65%	0%
Have read 2018 SOLs	39.53%	50%
Have been to training for 2018 SOLs	50%	50%
Have adjusted their teaching/instructional leadership for 2018 SOLs	65.11%	12.5%
Have heard of NGSS/3DT	31.4%	12.5%
Have not heard of NGSS/3DT	16.28%	50%
Have read about NGSS/3DT	45.35%	37.5%
Have been to training about NGSS/3DT	20.93%	0%
Use NGSS in their teaching/instructional leadership	23.26%	0%
Currently involved in science-specific professional organization	43.02%	25%

Note. Respondents were able to choose multiple responses.

Results from Task 2: Video Ratings

Raters were tasked with viewing three different videos from the website Ambitious Science Teaching. The first video that raters were asked to watch was entitled *Supporting Ongoing Changes in Student Thinking*, which was approximately 11 minutes in length. This video was of a portion of a middle school lesson from a unit on energy and phase changes. The second video was entitled *Day 2 Building a Consensus Model*, which was 16 minutes in length. The second video was an excerpt from a high school lesson from a unit on gas laws. The third video was entitled *Day 2 Eliciting Ideas Using Tuning Forks*, which was the longest at 19 minutes total. The third video was also from a high school class taken from a lesson from a unit on the physics of sound. I chose these three videos due to their varied classroom environments (such as grade level) and subject matter, which lends to the generalizability of PCI-OP in multiple classrooms and lessons. Respondents were asked to indicate the degree to which they had viewed the statement as part of the lesson in each video using a scale of 1-3, with 1 being "Not Observed," 2 being "Observed Partially," and 3 being "Observed Fully".

For the first video, 29 teachers and three administrators completed the 22 items being pilot tested in PCI-OP. Table 9 presents the means and standard deviations calculated for the teachers as a collective and for the administrators as a group. T-tests were not conducted for each item due to the very small sample size of administrators.

Table 9.

PCI-OP Ratings for Video 1

	Teachers (n=29)		Administrators (n=3)	
Item	Mean	Standard Deviation	Mean	Standard Deviation
1. Students planned investigations designed to answer a question or questions.	1.59	0.67	2.67	0.47
2. Students had the opportunity to propose investigations they are interested in.	1.38	0.61	2.00	0.82

	Teachers (n=29)		Administrators (n=3)	
Item	Mean	Standard Deviation	Mean	Standard Deviation
4. Students identified or discussed the variables in a certain investigation.	2.69	0.53	3.00	0.00
5. Students conducted investigations.	1.93	0.78	2.67	0.47
6. Students were asked to construct hypotheses or predictions.	2.59	0.67	3.00	0.00
7. Students determined or discussed measurements to be collected in a certain investigation.	2.48	0.68	3.00	0.00
8. Students observed and described events (e.g. weather events).	2.41	0.77	3.00	0.00
9. Students sorted or categorized objects or events (e.g. how animals move).	1.62	0.72	2.33	0.94
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	2.34	0.60	2.67	0.47
11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).	1.55	0.72	1.67	0.47
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	1.97	0.81	1.33	0.47
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	2.41	0.67	2.33	0.47

	Teachers (n=29)		Administrators (n=3)	
Item	Mean	Standard Deviation	Mean	Standard Deviation
15. Students recorded data or information from observations.	2.76	0.51	3.00	0.00
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	1.48	0.68	2.67	0.47
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	1.79	0.81	2.33	0.47
18. Students identified or discussed the data that should be collected to answer a testable question.	2.38	0.61	3.00	0.00
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	1.34	0.66	1.67	0.94
20. Students compared different methods of collecting the same data.	1.28	0.58	1.33	0.47
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	1.59	0.77	1.33	0.47
22. Students analyzed or evaluated data or information they collected.	2.69	0.53	2.67	0.47

For video 2, only two administrators completed the observation items; for video 3, only one administrator completed the observation items. Table 10 presents only teacher mean ratings and standard deviations for each item. Video 2 was rated by 19 teachers and video 3 was rated by 17 teachers. T-tests were not performed for ratings of videos 2 and 3 due to the small number of administrators.

Table 10.

PCI-OP Ratings for Videos 2 and 3, Teachers Only

	Video 2 (n=19)		Video 3 (n=17)	
Item	Mean	Standard Deviation	Mean	Standard Deviation
1	2	0.92	1.59	0.77
2	1.89	0.85	1.76	0.73
3	2.26	0.71	1.71	0.67
4	2.32	0.73	1.76	0.81
5	3.00	0.00	2.88	0.32
6	2.89	0.45	2.24	0.73
7	2.00	0.73	1.65	0.84
8	2.68	0.73	2.59	0.60
9	1.84	0.87	1.65	0.84
10	2.89	0.31	2.53	0.78
11	2.95	0.22	2.65	0.59
12	1.89	0.91	1.59	0.84
13	2.58	0.75	2.76	0.42
14	2.21	0.89	2.00	0.97
15	2.84	0.49	2.65	0.59
16	2.05	0.94	1.94	0.94
17	1.47	0.75	1.53	0.78
18	2.16	0.81	1.88	0.83
19	2.58	0.67	1.65	0.76
20	1.79	0.89	1.71	0.82
21	1.53	0.88	1.41	0.77
22	2.95	0.22	2.41	0.69

Item Analysis.

By combining the ratings for all three videos from both teachers and administrators, a total of 71 responses was obtained. For polytomously scored surveys such as PCI-OP, a classical item analysis includes item difficulty and item discrimination. Item difficulty for Likert-style questions is represented by the mean score for the item. In some cases, this value is normalized on a scale of 0 to 1 for easier comparison, achieved by dividing the mean by the maximum possible score. Rocconi (2023) referred to this as more of an item endorsability than a difficulty measure. Items with endorsabilities close to 0 or close to 1 may suggest the need for further scrutiny. In the context of PCI-OP, this could indicate that observers of a lesson consistently tend to overrate or underrate specific observed classroom behaviors for certain reasons. For example, items endorsabilities higher than 0.7 might imply that respondents find it too easy to rate the item highly. Table 11 displays item endorsability on a scale of 0 to 1 for all 22 items. None of the items had an endorsability low enough (less than 0.3) to raise concerns, but items 5, 6, 8, 10, 13, 15, and 22 exhibited high endorsabilities (highlighted below).

Item discrimination is assessed by calculating Pearson's product-moment correlation coefficient, which indicates the extent to which an item correlates with the construct measured by the instrument as a whole. The possible values for discrimination range from -1 (indicating perfect negative correlation) to +1 (indicating perfect positive correlation). Item discriminations close to zero suggest that the item might not be related to measuring the same construct as all of the other items. Negative item discriminations suggest that the item might measure something opposite; however, in some cases, this might imply that the item should be reverse-scored. Item discrimination exceeding 0.7 may indicate redundancy compared to other items. Items 3, 9, and

17 showed discriminations that were slightly high (highlighted below), while none of the items had discriminations that were negative or that were considered low or close to zero (< 0.3).

Table 11.

Item Analysis (n=71)

Item	Item Endorsability (scale 0 to 1)	Item Discrimination (scale -1 to +1)	Variance
1. Students planned investigations designed to answer a question or questions.	0.58	0.70	0.68
2. Students had the opportunity to propose investigations they are interested in.	0.55	0.63	0.60
3. Students identified or discussed materials necessary for a certain investigation.	0.64	0.70	0.58
4. Students identified or discussed the variables in a certain investigation.	0.78	0.35	0.63
5. Students conducted investigations.	0.84	0.35	0.54
6. Students were asked to construct hypotheses or predictions.	0.87	0.42	0.44
7. Students determined or discussed measurements to be collected in a certain investigation.	0.71	0.54	0.69
8. Students observed and described events (e.g. weather events).	0.86	0.39	0.50
9. Students sorted or categorized objects or events (e.g. how animals move).	0.57	0.70	0.66
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	0.86	0.56	0.39
11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).	0.75	0.43	0.73
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	0.61	0.63	0.75
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	0.85	0.40	0.42

Item	Item Endorsability (scale 0 to 1)	Item Discrimination (scale -1 to +1)	Variance
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	0.61	0.50	0.80
15. Students recorded data or information from observations.	0.92	0.36	0.26
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	0.61	0.59	0.80
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	0.55	0.72	0.63
18. Students identified or discussed the data that should be collected to answer a testable question.	0.73	0.61	0.62
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	0.58	0.64	0.76
20. Students compared different methods of collecting the same data.	0.50	0.58	0.60
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	0.50	0.56	0.62
22. Students analyzed or evaluated data or information they collected.	0.90	0.38	0.30

Task 3: Rater Reflections on Using PCI-OP

Task 3 for raters involved assessing statements about the feasibility of future use of PCI-OP. While teacher participants generally expressed that they found PCI-OP easy to use and believed it could be beneficial to teachers, their perception regarding whether administrators have time or capability to use PCI-OP was somewhat lower. Only one administrator completed this section, and rated all questions with a score of three out of four, as can be seen in Table 12.

The rating options for the reflection questions were 1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree.

Table 12.

Assessment of Future Use of PCI-OP

Item	Teachers (n=13)		Administrators (n=1)	
	Mean (scale 1-4)	Standard Deviation	Mean (scale 1-4)	Standard Deviation
I thought the survey items on the Observation form were easy to understand.	3.54	0.52	3.00	N/A
I understood what I was looking for when using the Observation form to watch the videos.	3.54	0.52	3.00	N/A
I think feedback from administrators on an Observation form like this would be helpful to science teachers.	3.15	0.80	3.00	N/A
I think an Observation form like this could help administrators be a valuable resource for science teachers implementing three-dimensional teaching or other inquiry-based strategies.	2.92	0.64	3.00	N/A
I think administrators could easily learn to use this Observation form.	2.77	1.01	3.00	N/A
I think administrators would have time to use this Observation form, in comparison to observation forms they currently use.	2.46	0.66	3.00	N/A

Note. 1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree.

Reliability

Once more, pooling the ratings for all three videos from both teachers and administrators yielded a total of 71 responses. Video One had 32 responses with an overall Cronbach's alpha of

0.87. Video Two had 19 responses with an overall Cronbach's alpha of 0.85. Video Three had 17 responses with an overall Cronbach's alpha of 0.93. These figures (0.87, 0.85, and 0.93) indicate good to excellent reliability for the 22 items in the third draft version of PCI-OP.

Chapter Summary

Data collection for this study occurred in three phases. Phase 1 involved gathering feedback from the three expert reviewers who suggested using a three-point Likert scale and including open comment space for each item. Additionally, they recommended adding instructions highlighting that observers should not expect to observe all behaviors in one lesson, as well as emphasized the importance of identifying science and engineering practices (SEP) within science content.

Phase 2 comprised a focus group with four participants, including two science teachers and two administrators. The primary insight from this focus group was the need to facilitate communication between teachers and administrators regarding the lesson context, such as the type of preceding lesson and the subsequent lesson trajectory.

Phase 3 involved a pilot test of the latest revision after incorporating feedback from the focus group. Participants provided background information, observed three video lessons, and answered reflection questions. Data analysis revealed good to excellent reliability of PCI-OP items (indicated by Cronbach's alphas ranging from 0.85 to 0.93), along with substantial agreement between administrators and teachers in rating the video lessons. However, some items may require deletion or revision, a topic to be explored further in Chapter 5.

Chapter 5. Conclusions and Discussion

The goal of this study was to assess the feasibility of developing an observation protocol tool focusing on observable features of a specific Science and Engineering Practice (SEP). A single SEP was selected to gauge the viability of such a protocol and to ascertain if further development can proceed on the other SEP. Two research questions guided this investigation:

1. Is it feasible to utilize the Science and Engineering Practices as a framework for the development of an observation protocol intended for use by instructional leaders in secondary science classrooms?
2. How do instructional leaders and science teachers assess a science lesson differently, and what are the implications of these potential similarities or differences on the feasibility of the observation protocol outlined in Research Question 1?

Data collection occurred in three consecutive phases, outlined in Chapters 3 and 4. This section provides a summary of the data collection procedures from Chapter 3 and the findings from Chapter 4, followed by conclusions organized by research question. A discussion of changes to PCI-OP that were informed by Phase 3 of this study is included in this chapter. The limitations of this study are then discussed, concluding with an exploration of future research avenues and implications for educators.

Summary of Data Collection and Findings

Phase 1 involved gathering feedback from three expert reviewers. The reviewers' primary feedback on PCI-OP was (a) to adopt a three-point Likert scale, (b) incorporate space for observers to provide evidence and comments for each item, and (c) to include a section regarding guidance for observers. The resulting revisions to guidance emphasized that observers should not expect to observe all behaviors on the instrument in any one lesson and highlighted the importance of identifying science and engineering practices within science content.

Phase 2 included a focus group with four participants, including two science teachers and two administrators. The key takeaway from this group regarding PCI-OP revisions was the need to facilitate communication between teachers and administrators regarding the context of the lesson, such as previous lesson types and future lesson plans.

Phase 3 comprised a pilot test of the most recent revision following the focus group. The survey link was distributed to approximately 1900 teachers and administrators within the Commonwealth of Virginia. Participants provided background information, observed three video lessons, and answered reflection questions. Data analysis revealed good to excellent reliability of PCI-OP items (indicated by Cronbach's alphas for the three different videos ranging from 0.85-0.93), substantial agreement between administrators and teachers in rating the video lessons, and potential for future use of PCI-OP or a similarly designed observation protocol.

Conclusions to Research Questions

RQ 1) Is it feasible to utilize the Science and Engineering Practices as a framework for the development of an observation protocol intended for use by instructional leaders in secondary science classrooms?

Considering the high alpha and agreement between the two groups, as discussed later in relation to RQ 2, along with feedback from participants in each phase, PCI-OP (or a similarly designed observation protocol) indeed demonstrates potential. However, feasibility could be further enhanced with additional refinement. It would be beneficial to expand its scope to encompass other science and engineering practices (SEPs), and accompanying guidelines and materials for both teachers and administrators would need to be developed. Of particular interest to the researcher was a subtle underlying theme observed in both Phase 2 and Phase 3, suggesting a greater interest among administrators in subject area-specific instruments like this compared to teachers, which is an outcome that will be explored later in this chapter.

RQ 2) How do instructional leaders and science teachers assess a science lesson differently, and what are the implications of these potential similarities or differences on the feasibility of the observation protocol outlined in RQ 1?

The number of administrators who responded to the recruitment request for participants was low (eight initially with only three completing the first video and just one completing all three videos). Therefore, any potential conclusions drawn should be approached with caution due to this limited sample size. Surprisingly, the administrators and the teachers did not differ significantly in how they rated lessons on PCI-OP. In contrast, I had expected more variation, especially in a science-specific observation protocol like PCI-OP. If administrators and science teachers had diverged in their ratings, it might have suggested that administrators struggle to evaluate science lessons from the same perspective as teachers. This suggests that administrators are capable of identifying scientific and engineering practices in ways that are similar to teachers. Nonetheless, further research is necessary to validate this conclusion, which will be discussed later in this chapter.

Final Version of PCI-OP

Usually, after completing the three phases involved in developing a research or assessment instrument, a researcher would construct a final draft of the instrument, which can be found in Appendix O. However, since this instrument study was conducted as a proof-of-concept, in practice, I would incorporate additional items related to science and engineering practices and then conduct another pilot test in future versions.

Meyer (2014) recommended that items with difficulties and discriminations ranging from 0.3 to 0.7 and with larger variances should be prioritized for retention in future versions of a survey instrument. The data collected in this study can be utilized to remove certain items—particularly those with high item discrimination—as they might indicate redundancy. Table 13 summarizes the changes and deletions to items and formatting of PCI-OP that would follow.

Table 13.

Final Revision to Items and Sections for PCI-OP Version 4

Item Wording (from version 3)	Item Endorsability and Discrimination	Action Needed and Rationale (new wording italicized, where appropriate)
1. Students planned investigations designed to answer a question or questions.	0.58 0.70	Keep for v.4
2. Students had the opportunity to propose investigations they are interested in.	0.55 0.63	Keep for v.4
3. Students identified or discussed materials necessary for a certain investigation.	0.64 0.70	Keep for v.4
4. Students identified or discussed the variables in a certain investigation.	0.78 0.35	Keep for v.4
5. Students conducted investigations.	0.84 0.35	Delete for redundancy. High endorsability.
6. Students were asked to construct hypotheses or predictions.	0.87 0.42	Keep for v.4
7. Students determined or discussed measurements to be collected in a certain investigation.	0.71 0.54	Keep for v.4
8. Students observed and described events (e.g. weather events).	0.86 0.39	Combine with item 10, see below.

Item Wording (from version 3)	Item Endorsability and Discrimination	Action Needed and Rationale (new wording italicized, where appropriate)
9. Students sorted or categorized objects or events (e.g. how animals move).	0.57 0.70	Combine with item 10, see below.
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	0.86 0.56	Combine with items 8, 9, and 13. All similar and with high endorsabilities. Should re-pilot test to see if endorsability changes. *Students observed, described, or categorized characteristics of physical objects or events (e.g. phase of mitosis, type of volcanic eruption, type of cloud, shape of leaf).*
11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).	0.75 0.43	Keep for v.4
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	0.61 0.63	Keep for v.4
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	0.85 0.40	Combine with item 10, see above.
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	0.61 0.50	Keep for v.4
15. Students recorded data or information from observations.	0.92 0.36	Combine with 18; see below.
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	0.61 0.59	Keep for v.4

Item Wording (from version 3)	Item Endorsability and Discrimination	Action Needed and Rationale (new wording italicized, where appropriate)
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	0.55 0.72	Combine with item 21; see below.
18. Students identified or discussed the data that should be collected to answer a testable question.	0.73 0.61	Combine with item 15. Similar. *Students identified, discussed, and/or recorded data and observations to answer a testable question.*
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	0.58 0.64	Keep for v.4
20. Students compared different methods of collecting the same data.	0.50 0.58	Keep for v.4
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	0.50 0.56	Combine with item 17. Similar. *Students identified or discussed the appropriate tool for a certain observation or measurement (e.g. a microscope for observing a cell, a thermometer for temperature, a ruler for length, a scale for mass).*
22. Students analyzed or evaluated data or information they collected.	0.90 0.38	Delete for redundancy. High endorsability.

In addition to the suggested edits described above, I also constructed a function where teachers would pre-select which items on PCI-OP that they would expect observers to see in a given lesson (see Appendix O). This could be done in the pre-observation phase and would allow communication between teacher and instructional leader regarding alignment between lesson planning and lesson implementation. Additionally, it shortened and therefore simplifies use of PCI-OP in classrooms.

Limitations of this Study

While this study yielded valuable data, it also features certain limitations that must be addressed. One limitation to the study pertains to the use of videos originally created for different purposes, which may not align with the context of this research. This artificial environment could have caused some confusion among participants. For instance, during the focus group, Suzanne's comments on aspects she disliked about the lesson highlighted potential confusion about the research focus, specifically regarding the development of the instrument rather than issues related to teaching in the video.

Another limitation is the small number of participants who completed the survey during Phase 3 of data collection, especially among administrators, which makes generalizability challenging. Additionally, the fact that 348 people clicked on the link indicates that the recruitment methods did generate interest, but something deterred potential participants from completing the survey. For future research employing a similar survey, potential strategies could include shortening the survey, using just one video instead of three, and restructuring the incentive structure; the latter could entail visiting schools to talk to science teachers during department meetings and providing a small incentive beforehand to encourage participation. Furthermore, engaging with professional groups focused on instructional leadership may facilitate connections with administrators for future survey participation. The limited number of administrators may have also influenced the results regarding differences between teachers and administrators as related to the research question.

While this study was limited to a local context (both geographically and the use of abbreviated VA SEP, which differ slightly from the NGSS SEP), the fact that the participant cohort for this study was restricted to teachers and administrators who were working within the Commonwealth of Virginia should be considered another limitation. It is likely that expanding the participant pool, and particularly to including those working in states with a more rigid adoption of NGSS, would provide more robust data.

For a future study, I would recommend establishing the specific area of science potential participants might be licensed to teach, currently teach, or have taught in the past. Such data would assist researchers in ensuring that an instrument such as PCI-OP could be optimally deployed for a specific purpose—for example in life science classes as opposed to physical science classes. It must also be noted that there was a potential response bias, particularly in Phase 3, related to the time required to watch the videos and complete the surveys. Other teachers or administrators who did not have the time to complete the process did not have the opportunity to share their thoughts.

Lastly, a potential miswording of a question during Phase 3 data collection could have limited the usefulness of data related to participants' familiarity with aspects of science education. Specifically, I asked participants to check all of the items that applied among statements such as, "I am familiar with inquiry" and "I consider myself well-versed in inquiry." However, resulting data indicated that some participants may not have checked all items that applied to them in that several checked that they considered themselves to be well-versed, but not that they were familiar with inquiry. This misperception skewed at least some of the data from this section of Phase 3, causing the percentages listed in Table 8 to be inaccurate.

Discussion of Future Research

On a fundamental level, additional testing of PCI-OP is needed to streamline and optimize its efficacy. The use of videos during this study's pilot test was convenient and enabled the researcher to ensure that participants were all observing the same lessons. Moreover, using three different videos allowed for the inclusion of a variety of stimuli being measured by PCI-OP. That said, the videos were not designed for the specific purpose of pilot testing PCI-OP, which could have led to a disconnect in what observers were watching and what was written in the observation form. This issue could be mitigated by tightly controlling the creation of a specific lesson video for observers to watch for further testing of PCI-OP—one that intentionally includes some items on the instrument and not others. The end result is that the survey would be considered as a criterion-referenced instrument as opposed to a norm-referenced instrument. In so doing, a research team could identify true scores for every item in response to the planned lesson, after which research participants would rate the lesson, followed by analysis to determine how close their ratings matched the previously defined true scores. For example, a lesson could purposely be designed to include some SEPs but not others. The true scores for items that were not included in the lesson would be 1 for "not observed;" whereas the true scores for items that were intentionally included would be 3 for "observed fully." While a good observation protocol needs to apply to as many different classroom settings as possible, a redesigned study such as this might confirm with greater confidence whether participants can use it accurately.

Some participant responses indicate the possibility of confusion over the purpose of the research. For example, Expert Reviewer #1 commented on the order of items not making sense regarding the chronological order of events that would likely unfold in a classroom. This issue indicates they believed that the study would involve PCI-OP being used in actual classrooms as opposed to pilot testing with videos. Additionally, Suzanne (in the focus group) commented on how she felt about the lesson itself, as opposed to the use of PCI-OP, indicating that she may have misunderstood the purpose of this study, but also suggests that she may have some negative attitudes and beliefs towards 3DT. Additional research that correlates teacher beliefs and attitudes about inquiry-based teaching and learning to the support they do or could receive from administrators might lead to a richer understanding of how to increase the use of 3DT and inquiry in classrooms.

The results from this investigation support the need for further research to explore the degree to which administrators would benefit from having subject-specific observation forms. The administrator feedback during the focus group, coupled with the fact that the one administrator who completed reflection questions in Phase 3 (see Table 8) rated questions higher than the group of teachers, suggests that administrators might *want* to use subject-specific observation forms and other tools for instructional leadership—but simply lacked the resources to have done so previously. Similarly, comparison studies using a current type of observation form and one like PCI-OP to see how both administrators and teachers respond might be beneficial. In short, additional studies are recommended to determine if administrators would prefer to utilize tool like PCI. If related data were available to support this idea, it could be extremely helpful for administrators in offering support to science teachers in the field.

Another future study could involve a researcher and administrator observing a teacher's lesson together, followed by debriefing with the teacher. Such an assessment approach would allow both research and practice angles to be addressed in the validation of PCI-OP or a similar instrument. Frequently, when pre-service administrators take licensure courses to earn their educational leadership credentials, they study the nature of educational research; in contrast, teachers often do not. According to the Virginia Tech School of Education (2024), their program for educational leadership requires a course called Research, Assessment, and Evaluation for PK-12 Leaders. Thus, reinforcing a robust three-way collaboration between teacher, administrator, and researcher could be fruitful, as the processes of research and practice could become more fully integrated.

Implications for Educators

The findings from this study on PCI-OP have implications for teacher educators of all subjects, not just science. Teachers who receive training in subject-specific instructional methodologies may rarely (or never) interact with other teachers who can support them with subject-specific instructional support once they begin their careers. Teacher educators could work with pre-service teachers about ways to become advocates for more subject-specific leadership over the course of their careers. Additionally, teacher educators should review current common processes of teacher evaluation with their pre-service teachers, with the goal of ensuring that they know what to expect and can communicate with school leadership effectively about how they learned to teach.

The study's findings underscore the potential value of integrating concise modules on subject-specific pedagogy into educational leadership courses. Doing so could significantly benefit aspiring administrators by equipping them with the necessary tools to foster high-quality instructional leadership among their future faculty members. As the landscape of resources aimed at supporting administrators in leading science teachers continues to expand, educators stand to gain from leveraging these advancements. For instance, platforms such as Ambitious Science Teaching have begun offering resources tailored to the needs of educational leaders. A recent observation tool for administrators, similar to that in Phase 3 of the study, has offered and exemplifies the increasing availability of practical support crucial for those seeking to enhance their leadership capabilities in science education.

Moreover, recent developments within the National Science Teaching Association (NSTA) further underscore the growing emphasis on school leadership within the field. At NSTA's national conference in March 2024, specialized sessions dedicated to addressing the needs of educational leaders were featured as a platform for educators to explore innovative strategies and best practices in educational leadership. Consider, for example, the following excerpt from two sessions:

> These two sessions will focus on leadership. These sessions will include a discussion of the structures that need to happen to support these shifts in science instruction. How do you get teachers on board to engage in this work? For example, what are examples of the stories you need to tell to support teacher interest and buy in. What are examples of sustainable structures at a system level to support this work? What are examples of observation and instructional tools to support this work? In session 1, we will draw on case studies to engage how storytelling can be used to support onboarding teachers and to strengthen a science infrastructure. In session 2, we will share observation and instructional tools for leaders to support this work. (NSTA, 2024)

Lastly, in districts and states where district science supervisors are present, collaborative professional development sessions could be organized involving both science teachers and instructional leaders. Such sessions offer the dual benefit of refreshing science pedagogy for teachers and introducing administrators to the nuances of the field. As relationships strengthen, tools like PCI-OP could be introduced, facilitating structured observation and discussion between administrators and science teachers. This would empower administrators to identify key actions in science classrooms while providing teachers the chance to elucidate concepts like 3DT and other inquiry-based learning approaches.

Chapter Summary

Based on three phases of data collection and subsequent analysis, PCI-OP holds promise for administrators to facilitate effective instructional leadership with secondary school science teachers. The potential facilitation lies in its capacity to support the implementation of 3DT by promoting the integration of science and engineering practices within the science curriculum. Furthermore, the data collected in this study indicates that both teachers and administrators provided similar ratings for science lessons, which underscores the potential value of a science-specific observation protocol, suggesting its merit in providing valuable insights into classroom practices.

The research trajectory and analysis in this study suggests numerous avenues for ongoing exploration. These potential studies could delve deeper into how both teachers and administrators perceive 3DT and the broader landscape of science teaching and learning. Moreover, future revisions and testing of instruments like PCI-OP, especially those incorporating elements beyond a single science and engineering practice, would contribute significantly to advancing this area of study and practice.

References

Acheson, K. A., & Gall, M. D. (2003). *Clinical supervision and teacher development: Preservice and inservice applications* (5th ed.). John Wiley & Sons. https://eric.ed.gov/?id=ED476151

Ambitious Science Teaching. (2023). *Principal AST look-for tool*. https://ambitiousscienceteaching.org/principal-ast-look-for-tool/

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association. https://www.aera.net/Publications/Books/Standards-for-Educational-Psychological-Testing-2014-Edition

Aoki, J. M., Foster, A. S., & Ramsey, J. M. (2005). Inquiry perceptions held by responding greater Houston-area science supervisors. *The Texas Science Teacher, October 2005*, 17-22. https://ui.adsabs.harvard.edu/abs/2003PhDT........68A/abstract

Barrow, L. H. (2006). A brief history of inquiry: From Dewey to standards. *Journal of Science Teacher Education*, *17*(3), 265–278. https://doi.org/10.1007/s10972-006-9008-5

Binns, I. C., & Popp, S. (2013). Learning to teach science through inquiry: Experiences of preservice teachers. *Electronic Journal of Science Education*, *17*(1), 1–24. https://doi.org/10.1007/s10972-012-9314-z

Biological Sciences Curriculum Study. (n.d.). *BSCS Story.* https://bscs.org/about/our-story/

Bodzin, A. M., & Beerer, K. M. (2003). Promoting inquiry-based science instruction: The validation of the Science Teacher Inquiry Rubric (STIR). *Journal of Elementary Science Education*, *15*(2), 39.l https://doi.org/10.1007/bf03173842

Bradford, C., & Braaten, M. (2018). Teacher evaluation and the demoralization of teachers. *Teaching and Teacher Education*, *75*, 49–59. https://doi.org/10.1016/j.tate.2018.05.017

Capps, D. K., & Crawford, B. A. (2013). Inquiry-based instruction and teaching about nature of science: Are they happening? *Journal of Science Teacher Education, 24*(3), 497-526. https://doi.org/10.1007/s10972-012-9314-z

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(8), 947-967. https://doi.org/10.1016/S0742-051X(02)00053-7

Crawford, B. A. (2014). *From inquiry to scientific practices in the science classroom.* In Handbook of research on science education, volume II (pp. 529-556). Routledge. https://doi.org/10.4324/9780203097267-36

Creamer, E. G. (2018). *An introduction to fully integrated mixed methods research.* SAGE Publications Inc. https://doi.org/10.4135/9781071802823

DeBoer, G. E. (1991). *A History of Ideas in Science Education: Implications for Practice*. Teachers College Press. https://eric.ed.gov/?id=ED329428

Elliott, S. (2021). Proof of concept research. *Philosophy of Science, 88*(2), 258-280. https//www.doi.org/10.1086/711503

Enochs, L. G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *Science Education, 74*(6), 625–637. https://doi.org/10.1002/sce.3730740605

Forbes, C. T., Biggers, M., & Zangori, L. (2013). Investigating essential characteristics of scientific practices in elementary science learning environments: The practices of science observation protocol (P-SOP). *School Science and Mathematics, 113*(4), 180-190. https://doi.org/10.1111/ssm.12014

Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2014). *SuperVision and instructional leadership: A developmental approach* (9th ed.). Pearson. https://www.pearson.com/en-us/subject-catalog/p/supervision-and-instructional-leadership-a-developmental-approach/P200000009874/9780137871827

Gray, R. O. N. (2014). The distinction between experimental and historical sciences as a framework for improving classroom inquiry. *Science Education, 98*(2), 327-341. https://doi.org/10.1002/sce.21098

Hayes, K. N., Lee, C. S., DiStefano, R., O'Connor, D., & Seitz, J. C. (2016). Measuring science instructional practice: A survey tool for the age of NGSS. *Journal of Science Teacher Education*, *27*, 137-164. https://doi.org/10.1007/s10972-016-9448-5

Heredia, S. C. & Yu, J. H. (2017). A matter of choice: Opportunities for informal science institutions to support science teacher induction. *Journal of Science Teacher Education, 28*(6), 549-565. doi: 10.1080/1046560X.2017.1388085

Hill, H. C., & Grossman, P. (2013). *Learning from Teacher Observations: Challenges and Opportunities Posed by New Teacher Evaluation Systems.* Harvard Educational Review. https://doi.org/10.17763/haer.83.2.d11511403715u376

Lochmiller, C. R. (2016). Examining administrators' instructional feedback to high school math and science teachers. *Educational Administration Quarterly*, *52*(1), 75–109. https://doi.org/10.1177/0013161X15616660

Lochmiller, C. R., Huggins, K. S., & Acker-Hocevar, M. A. (2012). Preparing leaders for math and science: Three alternatives to traditional preparation. *Planning and Changing*, *43*(1/2), 198–220. https://eric.ed.gov/?id=EJ977555

Luft, J. A. (1999). Assessing science teachers as they implement inquiry lessons: The extended inquiry observational rubric. *Science Educator*, *8*(1), 9-18. https://www.researchgate.net/publication/234737999_Assessing_Science_Teachers_as_T

Luft, J. A. (2001). Changing inquiry practices and beliefs: The impact of an inquiry-based professional development programme on beginning and experienced secondary science teachers. *International Journal of Science Education, 23*(5), 517-534. https://doi.org/10.1080/09500690121307

Lustick, D. (2009). The failure of inquiry: Preparing science teachers with an authentic investigation. *Journal of Science Teacher Education, 20*(6), 583-604. https://doi.org/10.1007/s10972-009-9149-4

Marshall, J. C., Smart, J., & Horton, R. M. (2010). The design and validation of EQUIP: An instrument to assess inquiry-based instruction. *International Journal of Science and Mathematics Education*, *8*, 299-321. https://doi.org/10.1007/s10763-009-9174-y

McKenney, S. & Reeves, T. C. (2013). Educational design research. In J.M. Spector, M.D., Merrill, J. Elen, M. J. Bishop (Eds.), *Handbook of research on educational communications technology* (pp. 131-140). Springer. https://www.doi.org/10.1007/978-1-4614-3185-5_11

Melville, W., Campbell, T., & Jones, D. (2016). Leading learning: Science departments and the chair. *School Science and Mathematics*, *116*(4), 189-198. https://doi.org/10.1111/ssm.12169

Meyer, J. P. (2014). *Applied measurement with jMetrik*. Routledge.

Montgomery County Public Schools. (n.d.) *Science-Montgomery County Public Schools.* https://www.mcps.org/about_us/departments/curriculum_pages/science

National Research Council. (2012). *A framework for K-12 science education: Practices, cross-cutting concepts, and core ideas*. The National Academies Press. https://doi.org/10.17226/13165

National Research Council. (1996). *National science education standards*. The National Academies Press. https://nap.nationalacademies.org/catalog/4962/national-science-education-standards

National Research Council. (2000). Supporting inquiry-based teaching and learning. In S. Olson & S. Loucks-Horsley (Eds.), *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning* (pp. 143–152). The National Academies Press. https://doi.org/10.17226/9596

National Research Council. (2005). *How students learn: History, mathematics, and science in the classroom*. The National Academies Press. https://doi.org/10.4135/9781452219011.n2

National Research Council. (2013). *Next generation science standards: For states, by states.* The National Academies Press. https://my.nsta.org/ngss/AccessStandardsByTopic

National Science Teaching Association (NSTA) (2024). *Sessions and Workshops NSTA National Conference Denver 2024.* https://www.nsta.org/national-conference-science-education-denver-2024-sessions-and-workshops.

National Science Teaching Association Board of Directors. (2004). *NSTA position statement: Scientific inquiry.* National Science Teaching Association. https://static.nsta.org/pdfs /positionstatement_scientificinquiry.pdf

National Science Teaching Association Board of Directors. (2018). *NSTA position statement: Transitioning from scientific inquiry to three-dimensional teaching and learning.* National Science Teaching Association. https://www.nsta.org/nstas-official-positions/transitioning-scientific-inquiry-three-dimensional-teaching-and-learning

National Science Teaching Association. (n.d.). *Leaders Institute 2023.* https://www.nsta.org/leaders-institute-2023

National Science Teaching Association. (n.d.). *NGSS hub.* https://ngss.nsta.org/About.aspx

Navy, S. L., Maeng, J. L., & Bell, R. L. (2019). Learning from a state professional development conference for science teachers: Beginning secondary science teachers' experiences. *Journal of Science Teacher Education*, *30*(4), 409-428. https://doi.org/10.1080/1046560X.2019.1584512

Peters, G. & Woolley, J.T. (n.d.). *Jimmy Carter, Department of Education Nomination of F. James Rutherford To Be an Assistant Secretary*. The American Presidency Project. https://www.presidency.ucsb.edu/node/249760

Ritter, J. M., Boone, W. J., & Rubba, P. A. (2001). Development of an instrument to assess prospective elementary teacher self-efficacy beliefs about equitable science teaching and learning (SEBEST). *Journal of Science Teacher Education*, *12*(3), 175-198. https://doi.org/10.1023/A:1016747713585

Rocconi, L. (2023). *Making the most of your survey items: Item analysis.* University of Tennessee Knoxville Educational Leadership and Policy Studies. https://elps.utk.edu/2024/01/15/making-the-most-of-your-survey-items-item-analysis/

Roehrig, G. H., & Kruse, R. A. (2005). The role of teachers' beliefs and knowledge in the adoption of a Reform-Based curriculum. *School Science and Mathematics, 105*(8), 412-

422. https://doi.org/10.1111/j.1949-8594.2005.tb18061.x

Roehrig, G. H., & Luft, J. A. (2004). Constraints experienced by beginning secondary science teachers in implementing scientific inquiry lessons. *International Journal of Science Education, 26*(1), 3-24. https://doi.org/10.1080/0950069022000070261

Rutherford, F. J. (1964). The role of inquiry in science teaching. *Journal of Research in Science Teaching*, *2*(2), 80–84. https://doi.org/10.1002/tea.3660020204

Rutherford, F. J., & Ahlgren, A. (1990). *Science for all Americans*. Oxford University Press, Inc.

Sampson, V. (2004). The science management observation protocol. *The Science Teacher*, *71*(10), 30. https://cursa.ihmc.us/rid=1Q3XDD98G-3Z7579-F2BK/Science%20Management%20Protocol.pdf

Sandoval, W.A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*, 634-656. https://doi.org/10.1002/sce.20065

Searles, W. E., & Kudeki, N. (1987). A comparison of teacher and principal perception of an outstanding science teacher. *Journal of Research in Science Teaching*, *24*(1), 1–13. https://doi.org/10.1002/tea.3660240102

Sekano, G., Laubscher, D., & Bailey, R. (2023). Challenging traditional teacher professional development by implementing technology-supported cooperative learning. *International Journal of Learning, Teaching and Educational Research, 22.* 542-543. https://doi.org/10.26803/ijlter.22.3.31

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/http://www.jstor.org/stable/1175860

Silm, G., Tiitsaar, K., Pedaste, M., Zacharia, Z. C., & Papaevripidou, M. (2017). Teachers' readiness to use inquiry-based learning: An investigation of teachers' sense of efficacy and attitudes toward inquiry-based learning. *Science Education International, 28*(4), 315-325. https://files.eric.ed.gov/fulltext/EJ1161535.pdf

Smolleck, L. D., Zembal-Saul, C., & Yoder, E. P. (2006). The development and validation of an instrument to measure preservice teachers' self-efficacy in regard to the teaching of science as inquiry. *Journal of Science Teacher Education*, *17*, 137-163. https://doi.org/10.1007/s10972-006-9015-6

Stein, M. K., & Nelson, B. S. (2003). Leadership content knowledge. *Educational Evaluation and Policy Analysis*, *25*(4), 423–448. https://doi.org/10.3102/01623737025004423

Virginia Department of Education. (2018). *Science standards of learning for Virginia public schools.* https://www.doe.virginia.gov/teaching-learning-assessment/k-12-standards-instruction/science/standards-of-learning

Virginia Department of Education. (2021). *Guidelines for uniform performance standards and evaluation criteria for teachers.* https://www.doe.virginia.gov/teaching-learning-assessment/teaching-in-virginia/performance-evaluation/teachers

Virginia Tech School of Education (2024). *About the Program for the Preparation of School Principals and Supervisors*. https://liberalarts.vt.edu/departments-and-schools/school-of-education/academic-programs/educational-leadership-program/about-maed-eds.html

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Looking inside the classroom: A study of K-12 mathematics and science education in the United States*. Horizon Research, Inc. https://horizonresearch.com/insidetheclassroom /reports/looking/frontmatter.pdf

Wheeler, L. B., Navy, S. L., Maeng, J. L., & Whitworth, B. A. (2019). Development and validation of the classroom observation protocol for engineering design (COPED). *Journal of Research in Science Teaching*, *56*(9), 1285-1305. https://doi.org/10.1002/tea.21557

Wilson, M. (2005). *Constructing Measures.* Psychology Press. https://doi.org/10.4324/9781410611697

Windschitl, M. (2003). Inquiry projects in science teacher education: What can investigative experiences reveal about teacher thinking and eventual classroom practice? *Science Education, 87*(1), 112-143. https://doi.org/10.1002/sce.10044

Appendix A. Acronyms and Definitions

3DT: Three-Dimensional Teaching (NRC, 2012; NRC, 2013)

CCC: Cross-Cutting Concepts (NRC, 2012; NRC, 2013)

CFA: Confirmatory Factor Analysis

DCI: Disciplinary Core Ideas (NRC, 2012; NRC, 2013)

EDR: Educational Design Research (McKenney & Reeves, 2013)

NGSS: Next Generation Science Standards (NRC, 2013)

NRC: National Research Council

NSES: National Science Education Standards (NRC, 1996)

NSF: National Science Foundation

NSTA: National Science Teaching Association (formerly the National Science Teachers Association)

PD: Professional Development

QP: Question Pro (survey construction and distribution platform)

SEP: Science and Engineering Practices (NRC, 2012; NRC, 2013)

SOL: Standards of Learning

VA SEP: Virginia Scientific and Engineering Practices

Appendix B. Surveys for Phase 1

Survey #1

In the reporting of the findings from this research, there will be a table of Expert Reviewer Qualifications. You will only be identified as "Expert Reviewer # _____". The researchers in this study would prefer to have you describe your own qualifications related to science education. What do you want the table to say about your qualifications as an expert reviewer of this observation protocol? For example, you might choose for your qualifications to read, "Science Education professor for 12 years; published three observation protocols; researcher of three-dimensional teaching in urban schools" or "Mid-Atlantic state Department of Education Science Curriculum Administrator for five years; co-wrote textbook for pre-service science methods course; former NSTA president".

After you and other expert reviewers make suggestions and comments on the Planning and Carrying Out Investigations Observation Protocol (PCI-OP), it will be edited and reviewed during a focus group of science teachers and administrators. Following the focus group, PCI-OP may be edited a second time. Following that, it will be piloted with science teachers and administrators who will watch video segments of science classes and rate the classes using PCI-OP. The raters will also have the opportunity to reflect on and provide feedback of their experiences using PCI-OP. One major objective of the development of PCI-OP is the potential for usefulness by administrators (who may not have robust science backgrounds) to give meaningful feedback to secondary science teachers. Please keep this objective in mind when providing comments and suggestions. Thank you.

What should be included in the way of instructions for observers who use PCI-OP? For example, it would be wise to remind observers that they should not expect to see all of the actions described in PCI-OP during one single science lesson. How would you recommend wording this, and are there other points to include in the instructions?

Some existing observation protocols separate items that focus on teacher actions or behaviors from items that focus on student actions or behaviors. (For example, the Science Instructional Practices protocol (Hayes, et al., 2016) has a section where teachers self-report what they do in a lesson, and then self-report what their students are asked to do during a lesson). Do you have experience with or thoughts about this idea? What would you recommend for PCI-OP? What are advantages and disadvantages of having two separate sections for observation of students and observation of teachers?

How should the response format of the items be constructed?

1. A Likert Scale 1-3 (1: not observed, 2: observed sometimes, 3: observed frequently)
2. Two options: Yes, observed; No, not observed
3. Other (please select this option and then type your suggestion in the blank) __________

Please provide any other feedback for the construction of PCI-OP and any suggestions for subsequent research using PCI-OP. Please leave any comments or feedback below. After all expert reviewers complete Survey #1, I will edit PCI-OP and send you an updated draft. After you review it, I will send you the link for a second and final survey.

Survey #2

After you and other expert reviewers make suggestions and comments on the Planning and Carrying Out Investigations Observation Protocol (PCI-OP), it will be edited and reviewed during a focus group of science teachers and administrators. Following the focus group, PCI-OP may be edited a second time. Following that, it will be piloted with science teachers and administrators who will watch video segments of science classes and rate the classes using PCI-OP. The raters will also have the opportunity to reflect on and provide feedback of their experiences using PCI-OP. One major objective of the development of PCI-OP is the potential for usefulness by administrators (who may not have robust science backgrounds) to give meaningful feedback to secondary science teachers. Please keep this objective in mind when providing comments and suggestions. Thank you.

What is your name?

Please select one option for each of the preliminary survey items on PCI-OP.

	Keep this item with no revisions	Keep this item, but consider revisions (to be given below)	Remove this item
1. Students planned investigations designed to answer a question or questions.	❑	❑	❑
2. Students had the opportunity to propose investigations they are interested in.	❑	❑	❑
3. Students identified or discussed materials necessary for a certain investigation.	❑	❑	❑
4. Students identified or discussed the variables in a certain investigation.	❑	❑	❑
5. Students conducted investigations.	❑	❑	❑
6. Students were asked to construct hypotheses or predictions.	❑	❑	❑
7. Students identified or discussed measurements to be collected in a certain investigation.	❑	❑	❑
8. Students observed and described phenomena (e.g. weather events).	❑	❑	❑
9. Students sorted or categorized phenomena (e.g. how animals move).	❑	❑	❑
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	❑	❑	❑
11. Students practiced using lab equipment (e.g. beakers, Bunsen burners).	❑	❑	❑
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	❑	❑	❑
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	❑	❑	❑
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	❑	❑	❑
15. Students recorded data or information from observations.	❑	❑	❑
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	❑	❑	❑

17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	❑	❑	❑
18. Students identified or discussed the data that should be collected to answer a testable question.	❑	❑	❑
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	❑	❑	❑
20. Students compared the accuracy of different methods of collecting the same data.	❑	❑	❑
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	❑	❑	❑
22. Students analyzed or evaluated data or information they collected.	❑	❑	❑

If you would revise any of these items, please type your suggested wording in the blank next to the item. If you recommended the item for removal or you don't have suggestions for revisions, please leave the item blank. You can click the up arrow if you need to go back to the previous question.

	Suggested Wording
1. Students planned investigations designed to answer a question or questions.	
2. Students had the opportunity to propose investigations they are interested in.	
3. Students identified or discussed materials necessary for a certain investigation.	
4. Students identified or discussed the variables in a certain investigation.	
5. Students conducted investigations.	
6. Students were asked to construct hypotheses or predictions.	
7. Students identified or discussed measurements to be collected in a certain investigation.	
8. Students observed and described phenomena (e.g. weather events).	
9. Students sorted or categorized phenomena (e.g. how animals move).	
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	
11. Students practiced using lab equipment (e.g. beakers, Bunsen burners).	
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	
15. Students recorded data or information from observations.	

16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	
18. Students identified or discussed the data that should be collected to answer a testable question.	
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	
20. Students compared the accuracy of different methods of collecting the same data.	
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	
22. Students analyzed or evaluated data or information they collected.	

Please provide any other feedback for the construction of PCI-OP and any suggestions for subsequent research using PCI-OP. For example, you may have thoughts about the order in which items are presented, or the instructions presented to observers using the protocol. Please leave any comments or feedback below. After all expert reviewers complete this process, I will edit PCI-OP and send you an updated draft to make any final comments.

Appendix C. Materials for Phase 1

INITIAL INVITATION

Dear __________________,

I hope this email finds you well! I am contacting you based on your expertise in science education. ***insert specific information about the reviewer here*** I have 18 years of experience as a Virginia chemistry and earth science teacher. Currently, I am a doctoral candidate in science education at Virginia Tech. I am in the process of my dissertation research, which involves writing and evaluating the feasibility of a prototype observation protocol based on the Science and Engineering Practice "Planning and Constructing Investigations." To strengthen the overall quality of this research, I kindly invite you to assist in the review process of this instrument.

Should you agree, I will send you a draft of the Planning and Carrying Out Investigations Observation Protocol (PCI-OP). I will also send a survey with questions about the wording and format of the items on the instrument. Your constructive and candid comments would be appreciated. After I collect feedback from expert reviewers such as yourself, future phases of validation will involve a focus group and pilot test with administrators and science teachers, so your thoroughness and attention to detail will assist me in improving and revising the instrument prior to that time.

I kindly request you please respond to me no later than November 15 to let me know if you are willing to participate in this project. I estimate that your involvement would take no more than 2-3 hours between now and December 15. Please feel free to contact me if you have any questions. I look forward to hearing from you soon!

Sincerely,

Jennifer Maguire

INFORMED CONSENT

Title of Research Study: A Proof of Concept Study in the Development of an Observation Protocol Based on Scientific and Engineering Practices

Principal Investigators:
Jennifer Maguire, maguirej@vt.edu
Dr. Natalie Ferand, nferand@vt.edu
Dr. Nancy Bodenhorn, nanboden@vt.edu
Virginia Tech, School of Education

Study Information: You have been identified as an expert in science education. If you agree to participate in the study, you will be asked to review a prototype observation protocol and give feedback on its content and structure.

Time required: Participating as an expert reviewer in this study is expected to require 2-3 hours of your time.

Risks and benefits: There are no anticipated risks or benefits to participating in the study. Only the researchers will have access to the information we collect. There is a minimal risk that security of any online data may be breached, but the online hosts (QuestionPro and GoogleDrive) use several forms of encryption and other protections, so it is unlikely that a security breach of the online data will result in any adverse consequences for you.

Compensation: You will not receive any compensation for participating in the study.

Confidentiality: Your identity will be kept confidential to the extent possible. You will be identified as "Expert Reviewer # ______" in the data collection and the findings that result from this project. Data will be kept on password protected computers and GoogleDrive file folders only accessible to the research team. Your name will not be used in any report of findings.

Voluntary participation: Your participation in this study is completely voluntary. You will not be penalized in any way for not participating.

Right to withdraw from the study: You have the right to withdraw from the study at any point. You will not be penalized in any way for withdrawing.

Whom to contact if you have questions about the study:
If you have questions, concerns, or complaints, or think the research has hurt you, talk to the principal investigators.
Jennifer Maguire maguirej@vt.edu
Dr. Natalie Ferand nferand@vt.edu
Dr. Nancy Bodenhorn nanboden@vt.edu

This research has been reviewed and approved by the Virginia Tech Human Research Protection Program (HRPP). You may communicate with them at 540-231-3732 or irb@vt.edu if:

· You have questions about your rights as a research subject.
· Your questions, concerns, or complaints are not being answered by the research team.
· You cannot reach the research team.
· You want to talk to someone besides the research team to provide feedback about this research.

Thank you very much for your time and support. Please click Next below to consent to participating in this research.

Appendix D. PCI-OP Draft v.1 (12.13.23)

Introduction and Instructions: This tool is intended to be used by instructional leaders and secondary school science teachers. It is meant to help provide support to teachers incorporating Science and Engineering Practices (NGSS, 2013; VDOE, 2018) in their lessons. Not all Science and Engineering Practices (SEPs) are appropriate to incorporate in every lesson. Priority and emphasis should be placed on examples in which the SEP is used in the context of trying to figure out a solution to a problem or explain a phenomenon. Explaining a phenomenon or designing a solution to a problem provides context that makes the incorporation of SEPs meaningful. This prototype tool does not include all SEPs, but instead focuses on the practice "Planning and Carrying Out Investigations." Future versions of the tool may include more of the SEPs.

The PCI-OP tool is not intended to be applied to a single classroom observation or in order to form a judgement about the quality of a teacher's lessons. PCI-OP is intended to help encourage dialogue between instructional leaders and teachers and to help provide teachers with guidance for the incorporation of SEPs within their content.

1. Students planned investigations designed to answer a question or questions.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

2. Students had the opportunity to propose investigations they are interested in.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

3. Students identified or discussed materials necessary for a certain investigation.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

4. Students identified or discussed the variables in a certain investigation.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

5. Students conducted investigations.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

6. Students were asked to construct hypotheses or predictions.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

7. Students identified or discussed measurements to be collected in a certain investigation.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

8. Students observed and described phenomena (e.g. weather events).

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

9. Students sorted or categorized phenomena (e.g. how animals move).

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

11. Students practiced using lab equipment (e.g. beakers, Bunsen burners).

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).

□ Not Observed	□ Observed Partially	□ Observed Fully

Evidence and/or Comments:		

14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

15. Students recorded data or information from observations.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

18. Students identified or discussed the data that should be collected to answer a testable question.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

20. Students compared the accuracy of different methods of collecting the same data.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

22. Students analyzed or evaluated data or information they collected.

□ Not Observed	□ Observed Partially	□ Observed Fully
Evidence and/or Comments:		

Appendix E. PCI-OP Draft v.2 (1.28.24)

Introduction and Instructions: This tool is intended to be used by instructional leaders and secondary school science teachers. It is meant to help provide support to teachers incorporating Science and Engineering Practices (NGSS, 2013; VDOE, 2018) in their lessons. Not all Science and Engineering Practices (SEPs) are appropriate to incorporate in every lesson. Priority and emphasis should be placed on examples in which the SEP is used in the context of trying to figure out a solution to a problem or explain a phenomenon. Explaining a phenomenon or designing a solution to a problem provides context that makes the incorporation of SEPs meaningful. This prototype tool does not include all SEPs, but instead focuses on the practice "Planning and Carrying Out Investigations." Future versions of the tool may include more of the SEPs.

The PCI-OP tool is not intended to be applied to a single classroom observation or in order to form a judgement about the quality of a teacher's lessons. PCI-OP is intended to help encourage dialogue between instructional leaders and teachers and to help provide teachers with guidance for the incorporation of SEPs within their content.

1. Students planned investigations designed to answer a question or questions.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

2. Students had the opportunity to propose investigations they are interested in.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

3. Students identified or discussed materials necessary for a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

4. Students identified or discussed the variables in a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

5. Students conducted investigations.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

6. Students were asked to construct hypotheses or predictions.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

7. Students determined or discussed measurements to be collected in a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

8. Students observed and described phenomena (e.g. weather events).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

9. Students sorted or categorized phenomena (e.g. how animals move).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**

Evidence and/or Comments:

14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

15. Students recorded data or information from observations.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

17. Students determined or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

18. Students determined or discussed the data that should be collected to answer a testable question.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

20. Students compared the accuracy of different methods of collecting the same data.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

22. Students analyzed or evaluated data or information they collected.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

Appendix F. PCI-OP Draft v.3 (2.1.24)

Introduction and Instructions: This tool is intended to be used by instructional leaders and secondary school science teachers. It is meant to help provide support to teachers incorporating Science and Engineering Practices (NGSS, 2013; VDOE, 2018) in their lessons. Not all Science and Engineering Practices (SEPs) are appropriate to incorporate in every lesson. Priority and emphasis should be placed on examples in which the SEP is used in the context of trying to figure out a solution to a problem or explain a phenomenon. Explaining a phenomenon or designing a solution to a problem provides context that makes the incorporation of SEPs meaningful. This prototype tool does not include all SEPs, but instead focuses on elements of the practice "Planning and Carrying Out Investigations." Even on a day where the teacher plans for the students to be investigating for the duration of class time, it is unreasonable to expect that all questions on this observation protocol will be observed fully.

The PCI-OP tool is not intended to be applied to a single classroom observation or in order to form a judgement about the quality of a teacher's lessons. PCI-OP is intended to help encourage dialogue between instructional leaders and teachers and to help provide teachers with guidance for the incorporation of SEPs within their content.

Future versions of this tool may include more of the SEPs. Again, this prototype version is being tested with questions that fall under just one SEP, "Planning and Carrying Out Investigations."

Pre-Observation: To be answered by the teacher prior to the class being observed.

Where does this lesson fall within the unit/concept that you're currently teaching? For example, is it an introduction to the concepts you'll be teaching over the next week? Is it a review after several days of developing conceptual understandings? Something in-between? _______________

____

____

____

Observation: To be completed by the instructional leader while watching a class. "Observed partially" might mean that you observed some students engaged in this behavior but not others, or it might mean that you saw it taking place for the whole class, but only briefly or at a surface level. Comments and items of evidence are optional.

1. Students planned investigations designed to answer a question or questions.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

2. Students had the opportunity to propose investigations they are interested in.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

3. Students identified or discussed materials necessary for a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**

Evidence and/or Comments:

4. Students identified or discussed the variables in a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

5. Students conducted investigations.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

6. Students were asked to construct hypotheses or predictions.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

7. Students determined or discussed measurements to be collected in a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

8. Students observed and described phenomena (e.g. weather events).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

9. Students sorted or categorized phenomena (e.g. how animals move).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**

Evidence and/or Comments:

12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

15. Students recorded data or information from observations.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

17. Students determined or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

18. Students determined or discussed the data that should be collected to answer a testable question.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

20. Students compared the accuracy of different methods of collecting the same data.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

22. Students analyzed or evaluated data or information they collected.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

Post-Observation: To be answered by the teacher after the class that was observed.

Where are you and your students going with these concepts from here? For example, do you plan to revisit some of the concepts they developed during this lesson and address possible misconceptions? What learning events will you be using next? Or maybe you plan to assess this concept next and then potentially move on to new material? ______________________________

____

____

____

Appendix G. PCI-OP Recruitment Website

Web Address: http://sites.google.com/vt.edu/pci-op/home
Screenshots retrieved March 28, 2024

Phase 3 has ended! Thank you SO much to everyone who participated in this project! Stay tuned for information about findings and where we go next!

PCI-OP Home Phase 1: Expert Reviewers Phase 2: Focus Group Participants Phase 3: Raters

Planning and Carrying Out Investigations:

An Observation Protocol

Welcome!

Thank you for your willingness to participate in this research! Your efforts will help researchers make recommendations about support services for science teachers and their administrators. If you'd like to volunteer to be part of the study, please click Phase 3 above and go through the steps indicated. If you'd like to learn more about the steps this research has been through already, click above to read more about Phases 1 and 2.

FAQ

Q: What is this PCI-OP research project about?

The PCI-OP is being developed as the dissertation research for one of the researchers, Jennifer Maguire. To contact Jennifer, see below. PCI-OP will explore the potential for building an observation protocol based on the science and engineering practice "Planning and Carrying Out Investigations". The study will involve both science teachers and instructional leaders as participants, keeping in mind that both teachers and administrators are the end users of such a tool. The research questions for this study are:

1) Does using the Science and Engineering Practices as a framework for the development of an observation protocol for use by instructional leaders in secondary science classrooms show feasibility?

2) Do instructional leaders and science teachers rate a science lesson differently; what are the implications of those potential similarities or differences on the feasibility of the observation protocol from RQ 1?

Q: What is an observation protocol?

An observation protocol is a tool that is used to help guide an observer of classroom instruction. There are many different published observation protocols. They differ in the context used (such as grade level or subject), the focus of the observation (such as classroom management or formative assessment), or by user (such as an administrator or a student teacher).

Q: What does the state of Virginia say about teacher evaluation and using observation protocols?

The VDOE does not mandate the use of any one observation tool for administrators to use when evaluating teachers. According to the Guidelines for uniform performance standards and evaluation criteria for teachers from VDOE, the optional use of validated observation protocols is suggested as just one part of a holistic evaluation system. The evaluation system in Virginia lists eight "unified performance standards" for all teachers, listed and described below.

Performance Standard 1: Professional Knowledge

Performance Standard 2: Instructional Planning

Performance Standard 3: Instructional Delivery

Performance Standard 4: Assessment of/for Student Learning

Performance Standard 5: Learning Environment

Performance Standard 6: Culturally Responsive Teaching and Equitable Practices

Performance Standard 7: Professionalism

Performance Standard 8: Student Academic Progress

Additionally, during the spring of 2023, the Office of STEM Instruction at VDOE developed a tool for administrators in the state of Virginia to use specifically with science teachers. For more information about this tool, the Science Classroom Observation Tool (SCOT), visit here.

Q: What are the NGSS?

The Next Generation Science Standards are a set of comprehensive K-12 science standards that have been fully adopted by 20 states and the District of Columbia. They are laid out in three dimensions that describe the science that students should know and be able to do by certain grade levels. Visit here and here for more information.

Q: What is Three-Dimensional Teaching?

Three-dimensional teaching and learning (3DT) involves the incorporation of three dimensions: 1) science and engineering practices; 2) cross-cutting concepts; and 3) disciplinary core ideas.

Q: What are Science and Engineering Practices?

Science and Engineering Practices are one of the dimensions of the Next Generation Science Standards (NGSS):

1. Asking questions (for science) and defining problems (for engineering)

2. Developing and using models

3. Planning and carrying out investigations

4. Analyzing and interpreting data

5. Using mathematics and computational thinking

6. Constructing explanations (for science) and designing solutions (for engineering)

7. Engaging in argument from evidence

8. Obtaining, evaluating, and communicating information

In Virginia, the Science Standards of Learning modify these slightly and call them the Scientific and Engineering Practices. They are included as part of the Science SOLs for every grade level and class:

1. asking questions and defining problems

2. planning and carrying out investigations

3. interpreting, analyzing, and evaluating data

4. constructing and critiquing conclusions and explanation

5. developing and using models

6. obtaining, evaluating, and communicating information

Q: Does Virginia use the NGSS?

Virginia has not adopted the NGSS. However, the most recent revisions to the Virginia Science SOLs were informed in multiple ways by the NGSS. In particular, the NGSS's Science and Engineering Practices and Virginia's Scientific and Engineering Practices are very similar.

For more information, contact:

Jennifer Maguire

Doctoral Candidate, Virginia Tech

maguirej@vt.edu

This research has been approved as Exempt

VT IRB #23-1222 and VT IRB #23-1290

Phase 3 has ended! Thank you SO much to everyone who participated in this project! Stay tuned for information about findings and where we go next!

PCI-OP

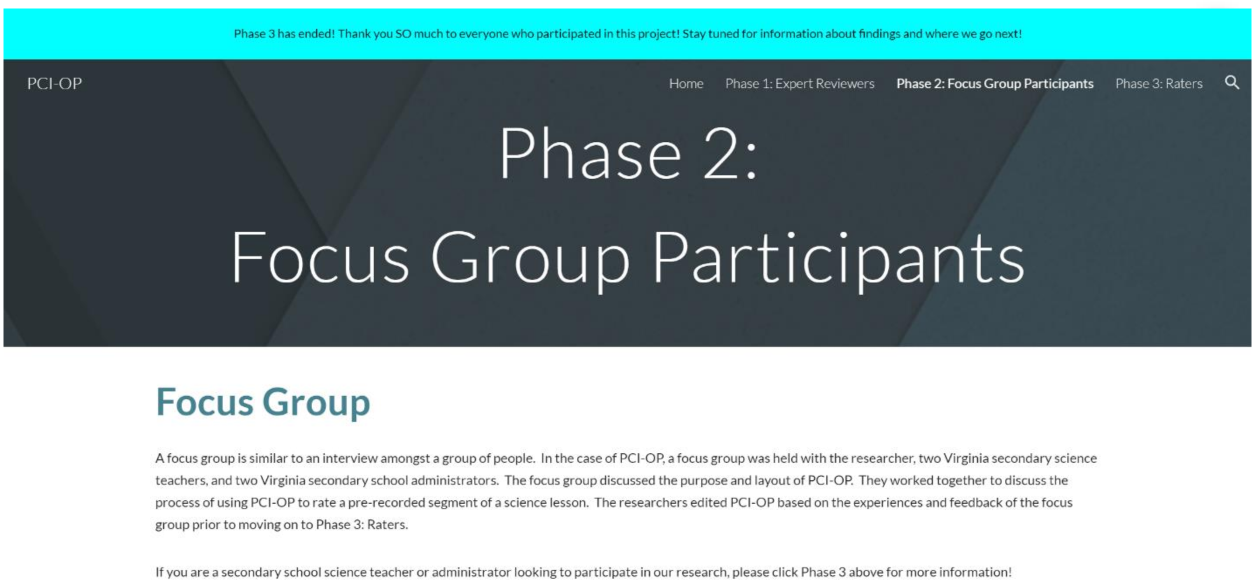
Home | **Phase 1: Expert Reviewers** | Phase 2: Focus Group Participants | Phase 3: Raters

Phase 1:
Expert Reviewers

Expert Reviewers

Eliciting the feedback of experts is a common early phase when developing and testing a research instrument. In the case of PCI-OP, feedback was gathered from three experts within the field of science education. They commented on the structure of the instrument, formatting for instructions, and wording of the individual items on PCI-OP. The researchers edited PCI-OP based on the comments from the expert reviewers prior to moving on to Phase 2: Focus Group.

If you are a secondary school science teacher or administrator looking to participate in our research, please click Phase 3 above for more information!



Phase 3 has ended! Thank you SO much to everyone who participated in this project! Stay tuned for information about findings and where we go next!

PCI-OP

Home | Phase 1: Expert Reviewers | **Phase 2: Focus Group Participants** | Phase 3: Raters

Phase 2:
Focus Group Participants

Focus Group

A focus group is similar to an interview amongst a group of people. In the case of PCI-OP, a focus group was held with the researcher, two Virginia secondary science teachers, and two Virginia secondary school administrators. The focus group discussed the purpose and layout of PCI-OP. They worked together to discuss the process of using PCI-OP to rate a pre-recorded segment of a science lesson. The researchers edited PCI-OP based on the experiences and feedback of the focus group prior to moving on to Phase 3: Raters.

If you are a secondary school science teacher or administrator looking to participate in our research, please click Phase 3 above for more information!

Phase 3 has ended! Thank you SO much to everyone who participated in this project! Stay tuned for information about findings and where we go next!

PCI-OP

Home | Phase 1: Expert Reviewers | Phase 2: Focus Group Participants | **Phase 3: Raters**

Phase 3: Raters

We're thrilled you're here.

It's amazing that you're considering this. To help us as researchers dig into new ideas that might work for teachers and administrators- and that can ultimately affect the education that children are getting.... that's just incredible, thank you for even thinking about it!

Who can participate?

If you work in Virginia:

-as a public secondary school science teacher

-a public secondary school administrator

-OR a district, county or state level science curriculum leader,

then yes, we need you!

What will participants need to do?

If you agree to participate in the study, you will be asked to: 1) complete a survey about your professional background (e.g. number of years teaching, subjects taught, etc.); 2) watch several video clips of science classes and complete an observation survey while you watch; and 3) complete a survey about your experience using the observation survey.

The three tasks can be done all at once, or you can break them up over time, as long as they are all completed by Thursday, February 29, 2024. In total, we estimate that the three tasks would take you 1-2 hours to complete.

Ready to get started??

Click here to visit our survey. All three tasks are on this survey, but you should be able to save your progress after Task 1 or Task 2 and return to it later if you don't want to complete it all at once. If you choose to save your progress, the survey will ask for your email so it can send you your saved survey. Your email will not be used for any other purpose other than redirecting you to finish your survey.

At the conclusion of Task 3, you'll be asked if you want to enter a drawing for one of **eight** digital Amazon gift cards. **Two** gift cards are worth **$50** each and **six** are worth **$20** each. To enter the drawing, you'll submit your email address on a link to a Google form given at the end of Task 3. This ensures your identity is kept separate from your answers, thus preserving your anonymity. The drawing will be held after all data has been collected and we'll contact you by email if you're selected (we estimate that the drawing will take place some time in mid-late March 2024).

Current Draft of PCI-OP

Below you'll find the current draft of PCI-OP, which is being used during Phase 3 of this research.

Planning and Carrying Out Investigations-Observation Protocol (PCI-OP)

Introduction and Instructions: This tool is intended to be used by instructional leaders and secondary school science teachers. It is meant to help provide support to teachers incorporating Science and Engineering Practices (NGSS, 2013; VDOE, 2018) in their lessons. Not all Science and Engineering Practices (SEPs) are appropriate to incorporate in every lesson. Priority and emphasis should be placed on examples in which the SEP is used in the context of trying to figure out a solution to a problem or explain a phenomenon. Explaining a phenomenon or designing a solution to a problem provides context that makes the incorporation of SEPs meaningful. This prototype tool does not include all SEPs, but instead focuses on elements of the practice "Planning and Carrying Out Investigations." Even on a day where the teacher plans for the students to be investigating for the duration of class time, it is unreasonable to expect that all actions referred to in items on this observation protocol would be observed fully.

The PCI-OP tool is not intended to be applied to a single classroom observation or in order to form a judgment about the quality of a teacher's lessons. PCI-OP is intended to help encourage dialogue between instructional leaders and teachers and to help provide

Appendix H. Preliminary Items for PCI-OP

1. Students planned investigations designed to answer a question or questions.
2. Students had the opportunity to propose investigations they are interested in.
3. Students identified or discussed materials necessary for a certain investigation.
4. Students identified or discussed the variables in a certain investigation.
5. Students conducted investigations.
6. Students were asked to construct hypotheses or predictions.
7. Students identified or discussed measurements to be collected in a certain investigation.
8. Students observed and described phenomena (e.g. weather events).
9. Students sorted or categorized phenomena (e.g. how animals move).
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).
11. Students practiced using lab equipment (e.g. beakers, Bunsen burners).
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).
15. Students recorded data or information from observations.
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).
18. Students identified or discussed the data that should be collected to answer a testable question.
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.
20. Students compared the accuracy of different methods of collecting the same data.
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).
22. Students analyzed or evaluated data or information they collected.

Appendix I. Materials for Phase 2

INITIAL INVITATION

Dear __________________,

I hope this email finds you well! I am contacting you because you are an administrator/science teacher in Virginia. I have 18 years of experience as a Virginia chemistry and earth science teacher. Currently, I am a doctoral candidate in science education at Virginia Tech. I am in the process of my dissertation research, which involves testing the feasibility of a method for science teachers and administrators to work together. To strengthen the overall quality of this research, I would like to kindly invite you to participate in a focus group that will review an observation form specifically for science classrooms. The focus group will involve a conversation amongst 5-6 science teachers and administrators. It will last for approximately 90 minutes and refreshments will be provided. It will be held at a place convenient for all parties.

Should you agree to be one of the focus group participants, I will send you a survey regarding your availability over the next few weeks. I will also send a survey about your professional background.

I kindly request you please respond to me no later than December 1 to let me know if you are willing to participate in this project. I estimate that your involvement would take approximately two hours (focus group and survey completion) between January 2 and 10. Please feel free to contact me if you have any questions. I look forward to hearing from you soon!

Sincerely,

Jennifer Maguire

INFORMED CONSENT

Title of Research Study: A Proof of Concept Study in the Development of an Observation Protocol Based on Scientific and Engineering Practices

Principal Investigators:

Jennifer Maguire, maguirej@vt.edu

Dr. Natalie Ferand, nferand@vt.edu

Dr. Nancy Bodenhorn, nanboden@vt.edu

Virginia Tech, School of Education

Study Information: You have been identified as a secondary science teacher or secondary school administrator in Virginia. If you agree to participate in the study, you will be asked to complete a survey about your professional background and then meet with 4-5 other teachers and administrators to participate in a focus group where you will review how a prototype observation protocol might be used in a secondary science classroom.

Time required: Participating as a focus group participant in this study is expected to require approximately 2 hours of your time.

Risks and benefits: There are no anticipated risks or benefits to participating in the study. Only the researchers will have access to the information we collect. There is a minimal risk that security of any online data may be breached, but the online hosts (QuestionPro and GoogleDrive) use several forms of encryption and other protections, so it is unlikely that a security breach of the online data will result in any adverse consequences for you.

Compensation: You will not receive any monetary compensation for participating in the study, but complimentary refreshments will be offered during the focus group. If the observation form proves to be useful, this should benefit future science teachers and administrators.

Confidentiality: Your identity will be kept confidential to the extent possible. You will be identified using a pseudonym of your choosing in the data collection and the findings that result from this project. You will have the opportunity to review quotes and other feedback you gave during the focus group meeting. Data will be kept on password protected computers and GoogleDrive file folders only accessible to the research team. Your name will not be used in any report of findings.

Voluntary participation: Your participation in this study is completely voluntary. You will not be penalized in any way for not participating.

Right to withdraw from the study: You are given the right to withdraw from the study at any point. You will not be penalized in any way for withdrawing.

Whom to contact if you have questions about the study:

If you have questions, concerns, or complaints, or think the research has caused any harm, talk to the principal investigators.

Jennifer Maguire maguirej@vt.edu

Dr. Natalie Ferand nferand@vt.edu

Dr. Nancy Bodenhorn nanboden@vt.edu

This research has been reviewed and approved by the Virginia Tech Human Research Protection Program (HRPP). You may communicate with them at 540-231-3732 or irb@vt.edu if:

· You have questions about your rights as a research subject.

· Your questions, concerns, or complaints are not being answered by the research team.

· You cannot reach the research team.

· You want to talk to someone besides the research team to provide feedback about this research.

Thank you very much for your time and support. Please click Next below to consent to participating in this research.

FOCUS GROUP PROFESSIONAL BACKGROUND SURVEY

Which of the following do you have experience with?

	Enter number of years' experience in Virginia. Enter 0 if you have no experience with this in Virginia.	Enter number of years' experience outside of Virginia. Enter 0 if you have no experience with this outside of Virginia.
Worked as an administrator at a middle or high school		
Worked as a teacher for any grade of science or type of science at a middle or high school		
Worked as a school-based science specialist/instructional coach (even if you worked in more than one school at a time)		
Worked as a county- or district-level administrator for science curriculum and instruction		
Worked as a state-level administrator for science curriculum and instruction		

At the time of taking this survey, what is your primary job?

1. Science teacher in a Virginia middle or high school
2. Administrator in Virginia (middle or high school, district, state, or other level)

What is your age?

With what gender do you identify?

How would you describe your own race or ethnicity?

How familiar are you with inquiry-based instruction in science education?

1. I haven't heard of it
2. I have heard of it
3. I don't think it is useful/helpful for student learning at the secondary level
4. I do think it is useful/helpful for student learning at the secondary level
5. I consider myself well-versed in inquiry-based instruction

6. Other; please describe

(Teachers only) How familiar are you with the newest Virginia Science SOLs (2018 version)? Check all that apply.

1. I haven't read them
2. I've read them
3. I've adjusted my teaching to reflect changes to the new SOLs
4. I've been to training or professional development related to the new SOLs
5. Other; please describe __________

(Teachers only) How familiar are you with three-dimensional teaching or the Next Generation Science Standards? Check all that apply.

1. I have heard of it
2. I have never heard of it
3. I have read about it
4. I've been to training or other professional development where it was discussed
5. I use it in my teaching
6. Other; please describe

Are you currently involved in any professional organizations specifically for science teaching? (for example, the National Science Teaching Association or the Virginia Association of Science Teachers)

1. Yes
2. No

(Administrators only) How familiar are you with the newest Virginia Science SOLs (2018 version)? Check all that apply.

1. I haven't read them
2. I've read them
3. I've discussed with teachers that I supervise ways to adjust their teaching in response to changes in the new SOLs
4. I've been to training or professional development related to the new SOLs
5. Other; please describe __________

(Administrators only) How familiar are you with three-dimensional teaching or the Next Generation Science Standards? Check all that apply.

1. I have heard of it
2. I have never heard of it
3. I have read about it
4. I've been to training or other professional development where it was discussed
5. I've discussed with teachers that I supervise ways to incorporate components of it in their teaching

6. Other; please describe __________

(Administrators only) Briefly describe (4-6 sentences) your experiences with and/or philosophy of supervising instruction and evaluating teachers.

(Administrators only) On a scale of 1-5 (1 being not at all and 5 being very much), to what extent do you think supervising and evaluating teachers should differ with regard to grade level or subject area?

	1: Not at all	2	3	4	5: Very much
Rate on a scale of 1-5	❏	❏	❏	❏	❏

When reports on the results of this focus group and survey are reported, you will be identified using a pseudonym. It's customary for participants to choose their own pseudonym. What name would you like to be known as in the reporting of this research?

FOCUS GROUP PROTOCOL

Welcome participants. Explain dissertation process, purpose and layout of this study, and PCI-OP, to include answering any questions about inquiry or 3DT.

Participants introduce themselves to one another.

Distribute paper copies of PCI-OP. Allow group to read through it first. Ask for comments or questions.

Watch one video clip together, pausing frequently to discuss out loud the thought process participants use as they decide how to complete PCI-OP in response to what they observe.

Discuss potential benefits and concerns of using PCI-OP in the future. Ask for suggestions on individual items. Ask for suggestions with instructions for using PCI-OP.

Ask administrators to consider what they would discuss with a teacher following the use of PCI-OP to observe a class. Ask teachers to consider what they would want the administrator to know. Discuss together.

Any final thoughts or comments? Discuss the process of transcription and subsequent member checking.

Appendix J. Materials for Phase 3

EARLY RECRUITMENT FLYER

Are you an administrator or science teacher at a secondary school in the state of Virginia??

We need your help!

Researchers at Virginia Tech are seeking to learn more about how science teachers and administrators work together. Volunteering for our study will help us make valuable contributions to how teachers teach and how administrators lead!

As a thank you for your time, you will have the opportunity to be entered into a drawing for one of two $50 Amazon gift cards!

Visit here for more information.(LINK)

Data collection starts January 15, 2024!

INITIAL INVITATION

Dear Colleague,

I hope this email finds you well! I am contacting you because you are a secondary administrator or secondary science teacher in Virginia. I have 18 years of experience as a Virginia chemistry and earth science teacher. Currently, I am also a doctoral candidate in science education at Virginia Tech. I am in the process of my dissertation research, which involves testing the feasibility of a method for science teachers and administrators to work together. To strengthen the overall quality of this research, I kindly invite you to participate in testing this method. As a thank you for volunteering to participate, you will have the chance to win a $50 Amazon gift card!

Participation in this project will involve three components: 1) completing a survey about your professional background (e.g. number of years teaching/administering, subjects taught, etc.); 2) watching several video clips of science classes and completing an observation survey while you watch; and 3) providing feedback about your experience using the observation survey. Participating in this project will involve approximately 2-3 hours of time between now and January 31, and you can complete the components at your convenience. Once you complete the final component, you will have the option of providing your contact information to be entered into a drawing for one of two $50 Amazon gift cards!

Should you agree to be one of the participants in this project, please click the link below and then choose “Phase 3: Raters” in the menu across the top. Please feel free to contact me if you have any questions. I hope you’ll decide to participate in this project!

LINK TO GOOGLE SITE

Sincerely,

Jennifer Maguire

INFORMED CONSENT

Title of Research Study: A Proof of Concept Study in the Development of an Observation Protocol Based on Scientific and Engineering Practices

Principal Investigators:
Jennifer Maguire, maguirej@vt.edu
Dr. Natalie Ferand, nferand@vt.edu
Dr. Nancy Bodenhorn, nanboden@vt.edu
Virginia Tech, School of Education

Study Information: You have been identified as a secondary school administrator in Virginia. If you agree to participate in the study, you will be asked to: 1) complete a survey about your professional background (e.g. number of years teaching, subjects taught, etc.); 2) watch several video clips of science classes and complete an observation survey while you watch; and 3) complete a survey about your experience using the observation survey.

Time required: Participating as a rater in this study is expected to require 2-3 hours of your time; it does not need to be completed all at once.

Risks and benefits: There are no anticipated risks or benefits to participating in the study. Only the researchers will have access to the information we collect. There is a minimal risk that security of any online data may be breached, but the online hosts (QuestionPro and GoogleDrive) use several forms of encryption and other protections, so it is unlikely that a security breach of the online data will result in any adverse consequences for you.

Compensation: You will not receive any compensation for participating in the study. However, at the conclusion of the study, a drawing will take place for one of two $50 Amazon gift cards. You and all other raters may choose to participate in the drawing by providing your email address so we may contact you if your email address is drawn as a winner.

Confidentiality: Your identity will be kept confidential to the extent possible. You will be identified using a pseudonym of your choosing in the data collection and the findings that result from this project. You will have the opportunity to review quotes and other thoughts you offered during the surveys. Data will be kept on password protected computers and GoogleDrive file folders only accessible to the research team. Your name will not be used in any report of findings.

Voluntary participation: Your participation in this study is completely voluntary. You will not be penalized in any way for not participating.

Right to withdraw from the study: You are given the right to withdraw from the study at any point. You will not be penalized in any way for withdrawing.

Whom to contact if you have questions about the study:
If you have questions, concerns, or complaints, or think the research has hurt you, talk to the principal investigators.
Jennifer Maguire maguirej@vt.edu
Dr. Natalie Ferand nferand@vt.edu
Dr. Nancy Bodenhorn nanboden@vt.edu

This research has been reviewed and approved by the Virginia Tech Human Research Protection Program (HRPP). You may communicate with them at 540-231-3732 or irb@vt.edu if:

- You have questions about your rights as a research subject.
- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team to provide feedback about this research.

Thank you very much for your time and support. Please click Next below to consent to participating in this research.

OBSERVATION SURVEY
INFORMED CONSENT
Title of Research Study: A Proof of Concept Study in the Development of an Observation Protocol Based on Scientific and Engineering Practices
Principal Investigators:
Dr. Natalie Ferand, nferand@vt.edu
Jennifer Maguire, maguirej@vt.edu
Virginia Tech, School of Education

Study Information: You have been identified as a secondary school administrator or teacher in Virginia. If you agree to participate in the study, you will be asked to: 1) complete a survey about your professional background (e.g. number of years teaching, subjects taught, etc.); 2) watch several video clips of science classes and complete an observation survey while you watch; and 3) complete a survey about your experience using the observation survey.

Time required: Participating as a rater in this study is expected to require about two hours of your time. Risks and benefits: There are no anticipated risks or benefits to participating in the study. Only the researchers will have access to the information we collect. There is a minimal risk that security of any online data may be breached, but the online hosts (QuestionPro and GoogleDrive) use several forms of encryption and other protections, so it is unlikely that a security breach of the online data will result in any adverse consequences for you.

Compensation: You will not receive any guaranteed compensation for participating in the study. However, you will have the option of entering a random drawing for one of eight Amazon gift cards. Six of the gift cards will be worth $20 each and two will be worth $50 each. The drawing will be conducted after all data has been collected.

Confidentiality: Your identity will be kept confidential to the maximum extent possible. The survey you complete will not be attached to your name or other characteristics about you. Data will be kept on password protected computers and GoogleDrive file folders only accessible to the research team. Your name will not be used in any report of findings.

Voluntary participation: Your participation in this study is completely voluntary. You will not be penalized in any way for not participating. Right to withdraw from the study: You are given the right to withdraw from the study at any point. You will not be penalized in any way for withdrawing. Whom to contact if you have questions about the study: If you have questions, concerns, or complaints, or think the research has hurt you, talk to the principal investigators.

Dr. Natalie Ferand nferand@vt.edu
Jennifer Maguire maguirej@vt.edu

This research has been reviewed and approved by the Virginia Tech Human Research Protection Program (HRPP). You may communicate with them at 540-231-3732 or irb@vt.edu if:
· You have questions about your rights as a research subject.
· Your questions, concerns, or complaints are not being answered by the research team.
· You cannot reach the research team.

· You want to talk to someone besides the research team to provide feedback about this research.

Thank you very much for your time and support. Please click Next below to consent to participating in this research.

Task #1: Please answer the following questions about your background. This helps us look for patterns in our data analysis.

Which of the following do you have experience with?

	Enter number of years' experience in Virginia. Enter 0 if you have no experience with this in Virginia.	Enter number of years' experience outside of Virginia. Enter 0 if you have no experience with this outside of Virginia.
Worked as an administrator at a middle or high school		
Worked as a teacher for any grade of science or type of science at a middle or high school		
Worked as a school-based science specialist/instructional coach (even if you worked in more than one school at a time)		
Worked as a county- or district-level administrator for science curriculum and instruction		
Worked as a state-level administrator for science curriculum and instruction		

At the time of taking this survey, what is your primary job?

1. Science teacher in a Virginia middle or high school
2. Administrator in Virginia (middle or high school, district, state, or other level)

What is your age?

With what gender do you identify?

How would you describe your own race or ethnicity?

How familiar are you with inquiry-based instruction in science education?

1. I haven't heard of it
2. I have heard of it
3. I don't think it is useful/helpful for student learning at the secondary level
4. I do think it is useful/helpful for student learning at the secondary level
5. I consider myself well-versed in inquiry-based instruction
6. Other; please describe

How familiar are you with the newest Virginia Science SOLs (2018 version)? Check all that apply.

1. I haven't read them
2. I've read them
3. I've adjusted my teaching or instructional leadership to reflect changes to the new SOLs
4. I've been to training or professional development related to the new SOLs
5. Other; please describe __________

How familiar are you with three-dimensional teaching or the Next Generation Science Standards? Check all that apply.

1. I have heard of it
2. I have never heard of it
3. I have read about it
4. I've been to training or other professional development where it was discussed
5. I use it in my teaching or instructional leadership
6. Other; please describe __________

Are you currently involved in any professional organizations specifically for science teaching? (for example, the National Science Teaching Association (NSTA), the Virginia Association of Science Teachers (VAST), the Virginia Science Education Leadership Association (VSELA), etc.)

1. Yes
2. No

Task #2 involves watching three different video clips of secondary science classes from the website Ambitious Science Teaching, which is a robust resource for enriching science teaching in all grades. While the researchers for PCI-OP are not affiliated with Ambitious Science Teaching, the researchers for PCI-OP requested and were granted permission to use the videos on the Ambitious Science Teaching website to test PCI-OP.

Right click here and select "open link in new tab." View the video under the heading "Supporting ongoing changes in student thinking." Start the video at about 12:00 and view the remainder of it (approx. 11 minutes total). Answer the questions below about what you observe students doing. If you have comments about any of the questions, make note of them to share later. For example, if you found a question confusing or difficult to answer, we'd like you to share that when you complete Task #3.

	Not observed	Observed Partially	Observed Fully
1. Students planned investigations designed to answer a question or questions.	❑	❑	❑
2. Students had the opportunity to propose investigations they are interested in.	❑	❑	❑
3. Students identified or discussed materials necessary for a certain investigation.	❑	❑	❑
4. Students identified or discussed the variables in a certain investigation.	❑	❑	❑
5. Students conducted investigations.	❑	❑	❑
6. Students were asked to construct hypotheses or predictions.	❑	❑	❑
7. Students determined or discussed measurements to be collected in a certain investigation.	❑	❑	❑
8. Students observed and described events (e.g. weather events).	❑	❑	❑
9. Students sorted or categorized objects or events (e.g. how animals move).	❑	❑	❑
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	❑	❑	❑
11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).	❑	❑	❑
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	❑	❑	❑
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	❑	❑	❑
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	❑	❑	❑
15. Students recorded data or information from observations.	❑	❑	❑

16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	❑	❑	❑
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	❑	❑	❑
18. Students identified or discussed the data that should be collected to answer a testable question.	❑	❑	❑
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	❑	❑	❑
20. Students compared different methods of collecting the same data.	❑	❑	❑
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	❑	❑	❑
22. Students analyzed or evaluated data or information they collected.	❑	❑	❑

Right click here and select "open link in new tab." View the video under the heading "Day 2 Building a consensus model." Start the video at about 9:00 and view it until about 25:00 (approx. 16 minutes total). Answer the questions below about what you observe students doing. If you have comments about any of the questions, make note of them to share later. For example, if you found a question confusing or difficult to answer, we'd like you to share that when you complete Task #3.

	Not observed	Observed Partially	Observed Fully
1. Students planned investigations designed to answer a question or questions.	❑	❑	❑
2. Students had the opportunity to propose investigations they are interested in.	❑	❑	❑
3. Students identified or discussed materials necessary for a certain investigation.	❑	❑	❑
4. Students identified or discussed the variables in a certain investigation.	❑	❑	❑
5. Students conducted investigations.	❑	❑	❑
6. Students were asked to construct hypotheses or predictions.	❑	❑	❑
7. Students determined or discussed measurements to be collected in a certain investigation.	❑	❑	❑
8. Students observed and described events (e.g. weather events).	❑	❑	❑
9. Students sorted or categorized objects or events (e.g. how animals move).	❑	❑	❑
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	❑	❑	❑

11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).	❑	❑	❑
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	❑	❑	❑
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	❑	❑	❑
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	❑	❑	❑
15. Students recorded data or information from observations.	❑	❑	❑
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	❑	❑	❑
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	❑	❑	❑
18. Students identified or discussed the data that should be collected to answer a testable question.	❑	❑	❑
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	❑	❑	❑
20. Students compared different methods of collecting the same data.	❑	❑	❑
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	❑	❑	❑
22. Students analyzed or evaluated data or information they collected.	❑	❑	❑

Right click here and select "open link in new tab." View the video under the heading "Day 2 eliciting ideas using tuning forks." Start the video at about 6:00 and view it until about 25:00 (approx. 19 minutes total). Answer the questions below about what you observe students doing. If you have comments about any of the questions, make note of them to share later. For example, if you found a question confusing or difficult to answer, we'd like you to share that when you complete Task #3.

	Not observed	Observed Partially	Observed Fully
1. Students planned investigations designed to answer a question or questions.	❑	❑	❑
2. Students had the opportunity to propose investigations they are interested in.	❑	❑	❑
3. Students identified or discussed materials necessary for a certain investigation.	❑	❑	❑
4. Students identified or discussed the variables in a certain investigation.	❑	❑	❑
5. Students conducted investigations.	❑	❑	❑

6. Students were asked to construct hypotheses or predictions.	❑	❑	❑
7. Students determined or discussed measurements to be collected in a certain investigation.	❑	❑	❑
8. Students observed and described events (e.g. weather events).	❑	❑	❑
9. Students sorted or categorized objects or events (e.g. how animals move).	❑	❑	❑
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	❑	❑	❑
11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).	❑	❑	❑
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	❑	❑	❑
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	❑	❑	❑
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	❑	❑	❑
15. Students recorded data or information from observations.	❑	❑	❑
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	❑	❑	❑
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	❑	❑	❑
18. Students identified or discussed the data that should be collected to answer a testable question.	❑	❑	❑
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	❑	❑	❑
20. Students compared different methods of collecting the same data.	❑	❑	❑
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	❑	❑	❑
22. Students analyzed or evaluated data or information they collected.	❑	❑	❑

During this final task, please reflect on your experience participating in this research. Right click on this document and select "open link in new tab." This is the current draft of the PCI-OP observation tool, which was edited following Phase 1 (where researchers collaborated with science education experts) and again following Phase 2 (where researchers held a focus group with science teachers and administrators) of this research project. Please read the document before completing this final task. Pay special attention to the instructions at the beginning, and the two pre- and post-observation questions. You might want to keep the document open after you read it to refer to as you answer the questions in this final task.

Following Phase 1 of this research project, the experts that researchers consulted with felt strongly that instructions for PCI-OP should include 1) information about using Science and Engineering Practices in context of student sense-making; and 2) a statement about how a science teacher should not be expected to include all behaviors on the PCI-OP in every lesson. Do you have any comments or other suggestions for the instructions of PCI-OP? Keep in mind that PCI-OP is being developed primarily for administrators (who may not have science backgrounds) to use with secondary science teachers.

If no comment on the Introduction and Instructions section, write "none" and continue.

Following Phase 2 of this research project, the focus group felt strongly that teachers should have an opportunity to share with administrators the circumstances and setting of the lesson being observed. For example, participants felt they would want to know how far into a unit a particular lesson took place. Teachers would be answering the pre- and post-observation questions. Do you have any comments or other suggestions for the pre- and post-observation questions on the current draft of PCI-OP?

If no comment on the Pre- and Post-observation sections, write "none" and continue.

Rate the following regarding your experience watching the videos during Task #2.

	Strongly disagree	Disagree	Agree	Strongly agree
I thought the survey items on the Observation form were easy to understand.	❏	❏	❏	❏
I understood what I was looking for when using the Observation form to watch the videos.	❏	❏	❏	❏
I think feedback from administrators on an Observation form like this would be helpful to science teachers.	❏	❏	❏	❏
I think an Observation form like this could help administrators be a valuable resource for science teachers implementing three-dimensional teaching or other inquiry-based strategies.	❏	❏	❏	❏

I think administrators could easily learn to use this Observation form.	❑	❑	❑	❑
I think administrators would have time to use this Observation form, in comparison to observation forms they currently use.	❑	❑	❑	❑

Do you have any suggestions for improving the wording or presentation of any of the questions on the PCI-OP observation tool? For any question that you don't have suggestions, enter a 0 in the blank.

	Suggestions
1. Students planned investigations designed to answer a question or questions.	
2. Students had the opportunity to propose investigations they are interested in.	
3. Students identified or discussed materials necessary for a certain investigation.	
4. Students identified or discussed the variables in a certain investigation.	
5. Students conducted investigations.	
6. Students were asked to construct hypotheses or predictions.	
7. Students determined or discussed measurements to be collected in a certain investigation.	
8. Students observed and described events (e.g. weather events).	
9. Students sorted or categorized objects or events (e.g. how animals move).	
10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).	
11. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).	
12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).	
13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).	
14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).	
15. Students recorded data or information from observations.	
16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.	
17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).	
18. Students identified or discussed the data that should be collected to answer a testable question.	
19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.	
20. Students compared different methods of collecting the same data.	
21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).	
22. Students analyzed or evaluated data or information they collected.	

Are you interested in entering a drawing for an Amazon gift card (two cards worth $50 each and six worth $20 each)? The drawing will take place after all data has been collected.

1. Yes
2. No

(If yes) Please right click on this Google form and select "open link in new tab" to enter your information for the drawing. You'll be contacted if you win.

Appendix K. VDOE Office of STEM Instruction Protocol

Science Classroom Observation Tool for School Administrators
June 2023

The Science Classroom Observation Tool for School Administrators was developed to support school administrators when observing students and teachers in the classroom, constructing feedback to teachers, and evaluating science instruction. This tool is to be used to construct feedback loops to ensure instruction and pedological methodologies are aligned to a science standards-based classroom and to inform the teacher on how to identify areas of focus for planning, acting, and continuously improving.

The observation tool is meant to be used in parts and not as a whole; teachers and administrators should identify areas of focus upon which to direct their attention.

"Look Fors" in the Science classroom include:

1) Teachers start each science unit with a unit **phenomenon**[1] or problem or question.
2) Students **ask questions** that drive the learning.
3) When prompted, students should be able to tell what they are doing for that day, make connections to the phenomena, and relate what they are learning to real-world problems.
4) Students use their own language when conducting an activity and before science specific vocabulary is introduced (conceptual understanding).
5) Students "DOING" science using the **Scientific and Engineering Practices (SEPs)** to make sense of a phenomenon.
6) Students engage in discourse as they **make sense of the science** versus the teacher telling students the answers.
7) The teacher uses questioning techniques that involve students in higher level questions and engages many students.

Standard Alignment: The lesson meets the expectations of the standards.			
Standard or lesson objective:			
Lesson Alignment: The instruction of the lesson meets the expectations of the standards.	Lesson fully meets	Lesson partially meets	Lesson does not meet

A phenomenon[1], problem or question that is aligned to the standards drive the instruction.	Is fully aligned	Is partially aligned	Is not aligned
The activities, materials, and/or tasks align to the expectations of the standards.	Is fully aligned to expectations	Is partially aligned to expectations	Is not aligned to expectations
Observation Notes:			

Student Actions	**Evidence: Select all that apply**	**Other observations**
Students are engaged in the phenomena[1] or problem.	☐ Students express interest in the phenomena or problem. ☐ Students ask questions about phenomena or problem. ☐ When asked, students can articulate what they are doing and why they are doing it as it relates phenomena or problem.	
Students connect the lesson(s) to their experiences, culture, and community.	☐ Students connect the lesson to their experience. ☐ Students connect lesson to something in their community.	
Students can use their understanding of content to extend their thinking and make sense of the phenomena or problem.	☐ Students extend their thinking through making connections to their experiences and the real-world. ☐ Students explain their understanding as they make connections to the phenomena. ☐ Depending on where students are in the learning process, students use their own language to	

[1] Phenomena- Natural phenomena are observable events that occur in our communities, the world, and the universe. The use of natural phenomena in science is intended to shift science learning from learning about a topic to figuring out why or how something happens. The phenomena should be selected that are relevant and engaging to students.

Student Actions	Evidence: Select all that apply	Other observations
	explain understanding of the phenomena or problem. ☐ Students support claims and explanations using evidence. ☐ Students extend their thinking through making connections to other concepts in science such as *stability and change, patterns, systems, cause and effect, energy and matter, and structure and function*	
Students use the scientific and engineering practices (SEPs) to gather or critique evidence to explain phenomena or solve a problem.	Students are: ☐ asking scientific questions ☐ planning and conducting investigations, including collecting data ☐ interpreting, analyzing or evaluating data ☐ constructing and critiquing conclusions, explanations, and arguments ☐ developing or using models ☐ evaluating and communicating information	

Instructional practices of the teacher	Evidence: Select all that apply	Observations
Teacher provides structures for students to express understanding and make explicit connections to prior and/or upcoming lessons.	☐ Teacher elicits students understanding of the phenomena. ☐ Teacher uses multiple strategies throughout instruction to allow students to make connections. ☐ Teacher provides scaffolds to support all students as they express understanding of phenomena or problem.	
Teacher lesson(s) design and implementation provides opportunities for students to explain the phenomena or problem.	☐ Opportunities for students to explore core content as they engage with a phenomenon or conduct an investigation. ☐ Opportunities for students to use data and resources to provide evidence of conceptual understanding. ☐ Teacher integrates questions to prompt student reflection and to have students revise misconceptions. ☐ Direct instruction, as needed, is used to address gaps in student understanding.	
Teacher supports students as they engage in the SEPS to gather, make sense of, and/or critique evidence to figure out phenomena.	☐ Safe classroom practices are consistently reinforced in the classroom. ☐ Instruction allows time for students to develop and implement SEPs. ☐ Teacher provides scaffolds and support for development of SEPs. ☐ Lesson provides opportunities for students to engage in SEPs to include investigations, modeling, and using evidence to construct arguments.	

Instructional practices of the teacher	Evidence: Select all that apply	Observations
Teacher promotes student sensemaking through opportunities for discourse.	☐ Discourse allows students' opportunities to use their own language to explain their understanding of the phenomena or problem. ☐ Questioning techniques are used to facilitate the sensemaking process.	
Teacher utilizes feedback to determine students' changes in thinking and builds student understanding through whole class discussion.	☐ Teacher asks questions to elevate/prompt student thinking. ☐ Teacher supports students in the development and revision of models throughout the learning process. ☐ Teacher uses a variety of strategies to determine student understanding.	
Observation Notes:		

Appendix L. Virginia Tech IRB Approval Letters

Division of Scholarly Integrity and Research Compliance
Institutional Review Board
North End Center, Suite 4120 (MC 0497)
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-3732
irb@vt.edu
http://www.research.vt.edu/sirc/hrpp

MEMORANDUM

DATE: November 29, 2023

TO: Natalie Louise Kincy Ferand, Jennifer Lynn Maguire

FROM: Virginia Tech Institutional Review Board (FWA00000572)

PROTOCOL TITLE: A Proof of Concept Study in the Development of an Observation Protocol Based on Scientific and Engineering Practices, Phase 1

IRB NUMBER: **23-1222**

Effective November 28, 2023, the Virginia Tech Human Research Protection Program (HRPP) determined that this protocol meets the criteria for exemption from IRB review under 45 CFR 46.104 (d) category(ies) 2(ii).

Ongoing IRB review and approval by this organization is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these activities impact the exempt determination, please submit an amendment to the HRPP for a determination.

This exempt determination does not apply to any collaborating institution(s). The Virginia Tech HRPP and IRB cannot provide an exemption that overrides the jurisdiction of a local IRB or other institutional mechanism for determining exemptions.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

https://secure.research.vt.edu/external/irb/responsibilities.htm

(Please review responsibilities before beginning your research.)

PROTOCOL INFORMATION:

Determined As: **Exempt, under 45 CFR 46.104(d) category(ies) 2(ii)**
Protocol Determination Date: **November 28, 2023**

ASSOCIATED FUNDING:

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

Division of Scholarly Integrity and Research Compliance
Institutional Review Board
North End Center, Suite 4120 (MC 0497)
300 Turner Street NW
Blacksburg, Virginia 24061
540/231-3732
irb@vt.edu
http://www.research.vt.edu/sirc/hrpp

MEMORANDUM

DATE:	January 11, 2024
TO:	Natalie Louise Kincy Ferand, Jennifer Lynn Maguire
FROM:	Virginia Tech Institutional Review Board (FWA00000572)
PROTOCOL TITLE:	A Proof of Concept Study in the Development of an Observation Protocol Based on Scientific and Engineering Practices, Phases 2 and 3
IRB NUMBER:	**23-1290**

Effective January 11, 2024, the Virginia Tech Human Research Protection Program (HRPP) determined that this protocol meets the criteria for exemption from IRB review under 45 CFR 46.104 (d) category(ies) 2(ii).

Ongoing IRB review and approval by this organization is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these activities impact the exempt determination, please submit an amendment to the HRPP for a determination.

This exempt determination does not apply to any collaborating institution(s). The Virginia Tech HRPP and IRB cannot provide an exemption that overrides the jurisdiction of a local IRB or other institutional mechanism for determining exemptions.

All investigators (listed above) are required to comply with the researcher requirements outlined at:

https://secure.research.vt.edu/external/irb/responsibilities.htm

(Please review responsibilities before beginning your research.)

PROTOCOL INFORMATION:

Determined As:	**Exempt, under 45 CFR 46.104(d) category(ies) 2(ii)**
Protocol Determination Date:	**January 11, 2024**

ASSOCIATED FUNDING:

The table on the following page indicates whether grant proposals are related to this protocol, and which of the listed proposals, if any, have been compared to this protocol, if required.

Invent the Future

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
An equal opportunity, affirmative action institution

Appendix M. Email Requesting Permission to Use Ambitious Science Teaching

From: Jessica Thompson <jjthomps@uw.edu>

Sent: Sunday, October 8, 2023 3:36 PM

To: Maguire, Jennifer

Subject: Re: Ambitious Science Teaching videos

Hi Jennifer,

You are more than welcome to use the videos for your dissertation. We will launch a new website version in a month or so. The links to the current videos should stay the same, and we are adding more videos, particularly at the elementary level. =)

All the best,

Jessica

On Sun, Oct 8, 2023 at 12:09PM Maguire, Jennifer <maguirej@exchange.vt.edu> wrote:

Dear Jessica,

Hi, I'm Jennifer. I'm a doctoral candidate in science education at Virginia Tech. I'm working on my dissertation research, and I have been discussing with my committee the need for some videos of science classrooms for my participants to watch as part of my research. My mentor, Takumi Sato, suggested I reach out to you and ask about potentially using some of the videos from the Ambitious Science Teaching site. Let me know if you think this is a possibility and I'd love to talk more about it.

Thank you!

Jennifer

Jennifer Maguire, M.A.Ed.

Pronouns: she/her/hers

Doctoral Candidate, Science Education

Graduate Assistant for Assessment & Program Evaluation

Office of Undergraduate Education, Virginia Tech

--

Jessica Thompson, Ph.D.

Professor

Teaching, Learning and Curriculum, Science Education

University of Washington, College of Education

http://ambitiousscienceteaching.org/

faculty page

office hours

(206) 221-4630 w

She/Her/Hers

The UW is on the Coast Salish homelands. I am grateful to live and work on these lands. In acknowledging the people and the land, we attend to the longer history and our place in that history. We recognize these lands and waters and their significance for the peoples who lived and continue to live in reciprocity with the land, whose practices were and are tied to the land and the water, and whose lives continue to enrich and develop in relationship to the land, waters and other inhabitants today.

--

Jessica Thompson, Ph.D.

Professor

Teaching, Learning and Curriculum, Science Education

University of Washington, College of Education

http://ambitiousscienceteaching.org/

faculty page

office hours

(206) 221-4630 w

She/Her/Hers

The UW is on the Coast Salish homelands. I am grateful to live and work on these lands. In acknowledging the people and the land, we attend to the longer history and our place in that history. We recognize these lands and waters and their significance for the peoples who lived and continue to live in reciprocity with the land, whose practices were and are tied to the land and the water, and whose lives continue to enrich and develop in relationship to the land, waters and other inhabitants today.

Appendix N. Preliminary Instrument Blueprint

Category 1: Proposing, Planning, or Comparing Different Ideas for Investigations: 4 items

1. Students planned investigations designed to answer a question or questions.

2. Students had the opportunity to propose investigations they are interested in.

16. Students designed, constructed, or tested a tool or system intended to solve a certain problem.

20. Students compared the accuracy of different methods of collecting the same data.

Category 2: Identifying or Discussing Different Components or an Investigation: 8 items

3. Students identified or discussed materials necessary for a certain investigation.

4. Students identified or discussed the variables in a certain investigation.

6. Students were asked to construct hypotheses or predictions.

7. Students identified or discussed measurements to be collected in a certain investigation.

17. Students identified or discussed the appropriate tool needed for a certain measurement (e.g. a ruler for length, a scale for mass).

18. Students identified or discussed the data that should be collected to answer a testable question.

19. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.

21. Students identified or discussed the appropriate tool for a certain observation (e.g. a microscope for a cell).

Category 3: Students Taking Investigative Actions: 10 items

5. Students conducted investigations.

8. Students observed and described phenomena (e.g. weather events).

9. Students sorted or categorized phenomena (e.g. how animals move).

10. Students identified or discussed characteristics of physical objects or events to look for (e.g. color changes, sounds).

11. Students practiced using lab equipment (e.g. beakers, Bunsen burners).

12. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).

13. Students observed and described physical objects or events in front of them (e.g. rocks, leaves).

14. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).

15. Students recorded data or information from observations.

22. Students analyzed or evaluated data or information they collected.

Appendix O. Planning and Carrying Out Investigations- Observation Protocol (PCI-OP)

Final Draft

Introduction and Instructions: This tool is intended to be used by instructional leaders and secondary school science teachers. It is meant to help provide support to teachers incorporating Science and Engineering Practices (NGSS, 2013; VDOE, 2018) in their lessons. Not all Science and Engineering Practices (SEPs) are appropriate to incorporate in every lesson. Priority and emphasis should be placed on examples in which the SEP is used in the context of trying to figure out a solution to a problem or explain a phenomenon. Explaining a phenomenon or designing a solution to a problem provides context that makes the incorporation of SEPs meaningful. This prototype tool does not include all SEPs, but instead focuses on elements of the practice "Planning and Carrying Out Investigations." Even on a day where the teacher plans for the students to be investigating for the duration of class time, it is unreasonable to expect that all questions on this observation protocol will be observed fully.

The PCI-OP tool is not intended to be applied to a single classroom observation or in order to form a judgement about the quality of a teacher's lessons. PCI-OP is intended to help encourage dialogue between instructional leaders and teachers and to help provide teachers with guidance for the incorporation of SEPs within their content.

Future versions of this tool may include more of the SEPs. Again, this prototype version is being tested with questions that fall under just one SEP, "Planning and Carrying Out Investigations."

Pre-Observation: To be answered by the teacher prior to the class being observed. After answering the pre-observation question below, teachers should check which observation items an observer should expect to see for this particular observation session.

Where does this lesson fall within the unit/concept that you're currently teaching? For example, is it an introduction to the concepts you'll be teaching over the next week? Is it a review after several days of developing conceptual understandings? Something in-between? _______________

____

____

____

____

Observation: To be completed by the instructional leader while watching a class. Observe only items that the teacher has checked off below; other items are not applicable to this lesson on this day. "Observed partially" might mean that you observed some students engaged in this behavior but not others, or it might mean that you saw it taking place for the whole class, but only briefly or at a surface level. Comments and items of evidence are optional.

☐ 1. Students planned investigations designed to answer a question or questions.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 2. Students had the opportunity to propose investigations they are interested in.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 3. Students identified or discussed materials necessary for a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 4. Students identified or discussed the variables in a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 5. Students were asked to construct hypotheses or predictions.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 6. Students determined or discussed measurements to be collected in a certain investigation.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 7. Students observed, described, or categorized characteristics of physical objects or events (e.g. phase of mitosis, type of volcanic eruption, type of cloud, shape of leaf).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 8. Students used appropriate, available lab equipment (e.g. beakers, Bunsen burners).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 9. Students practiced making measurements (e.g. length, mass, volume, temperature, elapsed time).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 10. Students observed and described representations of physical objects or events, such as photos or videos (e.g. clouds, volcanic eruptions).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 11. Students designed, constructed, or tested a tool or system intended to solve a certain problem.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 12. Students identified, discussed, and/or recorded data and observations to answer a testable question.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 13. Teacher and students reviewed safety or ethical concerns related to use of chemicals, equipment, or lab procedures.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

☐ 14. Students compared the accuracy of different methods of collecting the same data.

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**

Evidence and/or Comments:

☐ 15. Students identified or discussed the appropriate tool for a certain observation or measurement (e.g. a microscope for observing a cell, a thermometer for temperature, a ruler for length, a scale for mass).

□ **Not Observed**	□ **Observed Partially**	□ **Observed Fully**
Evidence and/or Comments:		

Post-Observation: To be answered by the teacher after the class that was observed.

Where are you and your students going with these concepts from here? For example, do you plan to revisit some of the concepts they developed during this lesson and address possible misconceptions? What learning events will you be using next? Or maybe you plan to assess this concept next and then potentially move on to new material? ____________________________

____

____

____