Improving Science Education in International Schools Through Professional Development Targeting Next Generation Science Standards Assessment Design

by

Wyatt Wilcox

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

Approved March 2020 by the
Graduate Supervisory Committee:

Gustavo Fischman, Chair
Leigh Graves Wolf
Shirley Droese

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

This study explores the impact of a professional development (PD) activity conducted for teachers of the *Next Generation Science Standards* (NGSS) at 15 American-curriculum international schools. The intervention involved teachers utilizing the 3D-PAST screening tool to systematically evaluate the alignment of teacher-designed assessments with the constructs of the NGSS and best practices in science instruction. Data about the way the intervention enhanced or challenged teachers' understanding of the NGSS were collected via a multiple methods approach. The *New Framework of Science Education Survey of Teacher Understanding* (NFSE-STU) was used in a retrospective pretest-posttest fashion to assess changes in teachers' understanding of NGSS constructs. Subsequently, interviews were conducted with participants which provided data that expanded upon the NFSE-STU findings. The Refined Consensus Model of Pedagogical Content Knowledge (RCM-PCK) was used to interpret the findings and situate the study within the extant literature on teacher PCK. The intervention was found to have a statistically significant effect on teachers' understanding of the NGSS in all areas measured by the NFSE-STU. Additionally, data suggest that the intervention elicited changes in teachers' classroom practices and improved collaborative professional practices. Also highlighted in the analysis was the significance of the relationship between the intervention moderator and the participants as a strong predictor of the way the intervention was perceived by teachers. The findings strongly support the suggestion that international school administrators seeking to maximize the impact of science teacher professional development should consider PD activities that train teachers

in the use of aids to align NGSS assessments, because doing so simultaneously enhances teacher understanding of the NGSS while encouraging meaningful changes to professional practice. The study contributes to the nascent body of literature utilizing the RCM-PCK to situate understanding of science-teacher PCK, and fills a void in literature examining PD in American curriculum international schools, and highlights issues with potential to serve as foci for additional cycles of action research in the areas of international schools, science teacher and NGSS-related professional development, and the use of tools similar to 3D-PAST within other teaching disciplines.

DEDICATION

To my wife, Susannah,

my daughter, “Ella” Susannah,

my son, Gideon,

and to my parents, Wayne and Karen.

You each, in your own ways, have given a part of your life,

so that this might be a part of mine.

ACKNOWLEDGMENTS

This dissertation would not have been possible without support and guidance from a number of individuals. Chief among the individuals who have guided me along this journey is Dr. Gustavo Fischman, my chair and LSC advisor. His encouragement to stop worrying about being a good student and to rather focus on being a great researcher, allowed me to find my own voice for this project. His extensive knowledge and his wisdom in knowing when to push, and also when to reassure me, were instrumental at every critical junction. I am grateful and thankful to have been under his tutelage.

It has also been a great privilege to have Dr. Leigh Graves Wolf serve as a committee member. One indicator of a great teacher is the number of students who, long after the final paper is submitted, continue to seek out their wisdom and assistance; I was just one among many of my cohort members who solicited her to sit on a committee. She is adept at providing meaningful and challenging feedback, yet make her students feel thankful for the learning opportunity rather than defeated by it. I can't recall how many times an iteration of the phrase, "I spoke with Leigh about…it was a *really* helpful conversation." was passed between my LSC members. It was many. I am grateful for her support and humbled by her example.

Dr. Shirley Droese has been a mentor from the very beginning of this journey. She holds a wealth of knowledge about the international school context, curriculum, and teacher professional development that have been indispensable resources. Her commitment to this work, unwavering support, and periodic check-ins have helped me to see this study through to the end. It has been an honor to have had Dr. Droese's support.

My LSC members, Allison Ellsworth, Matthew Romanoski, and Chantel Estes were my teammates through it all. We shared many laughs, many more struggles, and we now will also share one great accomplishment.

Paul Andersen must also be recognized. I am thankful that he remained helpful and accommodating as I put his product "under the microscope." Modeling the scientific mindset that he seeks for his students, he viewed the study with an open mind and eagerness to learn from its findings.

Finally, undertaking a research project that encompasses 15 schools in 13 countries and on four continents is no small task. Without the willingness and efforts of administrators at each school who engaged their faculty with my requests, this study simply would not have been possible. Each of these individuals played a small but important role in contributing to the success of the research:

Robin Appleby	Sheldon Guenther	Dr. Jennifer Ricks
David Brooker	Joseph Hart	Dr. Jeffrey Rosen
Dr. Tamela Canale	Glen Karlsson	Joseph Stanzione
Dr. Ginger Carlson	Shaysann Kaun	Stacy Stephens
Dr. Johanna Cena	Debra Lechner	Dr. David Straffon
Robyn Chapel	Steve Meade	Dr. Dana Specker Watts
Andrea Fossum	Dr. Leigh Miller	

TABLE OF CONTENTS

APPENDIX Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

3D-PAST	Three-Dimensional Performance Assessment Screening Tool
BP	Best Practices in Science Instruction
CCC	Crosscutting Concepts
CK	Content Knowledge
cPCK	Collective Pedagogical Content Knowledge
DCI	Disciplinary Core Ideas
ePCK	Enacted Pedagogical Content Knowledge
EARCOS	East Asia Regional Council of Overseas Schools
I3D	Integration of the Three Dimensions (of the NGSS)
MSA-CESS	Middle States Association of Commissions on Elementary and Secondary Schools
NESA	Near East South Asia
NRC	National Research Council
NFSE-STU	New Framework of Science Education Survey of Teacher Understanding
NGSS	Next Generation Science Standards
PCK	Pedagogical Content Knowledge
PD	Professional Development
PE	Performance Expectation
PK	Pedagogical Knowledge
pPCK	Personal Pedagogical Content Knowledge
RCM	Refined Consensus Model of Pedagogical Content Knowledge for Science Instruction
RPP	Retrospective Pretest-Posttest
SAS	Singapore American School
SEP	Science and Engineering Practices
SIPS	Science Instructional Practices Survey
STEBI	Science Teaching Efficacy Beliefs Instrument
WASC	Western Association of Schools and Colleges

CHAPTER 1

INTRODUCTION

The Challenges of Science Education Reform

Science instruction, as recognized today, entered the U.S. public school curriculum during the 19th century. This was in part because scientists themselves argued for its value amongst the studies of humanities, which were the primary focus of the time. Thomas Huxley, Herber Spencer, Charles Lyell, Michael Faraday, John Tyndall, and Charles Eliot were notable scientists who were outspoken about bringing science instruction into mainstream classrooms (DeBoer, 1991). In the midst of the transformations brought on by the industrial revolution, scientists argued the discipline's practical applications and inductive reasoning processes provided superior intellectual training over the deductive reasoning processes prevalent in education in the late 1800s (DeBoer, 1991).

As time went on, it was science's practical application and societal relevance that eventually justified its place in formal education, rather than the inductive reasoning and logic skills that are inherently linked to the field (National Education Association, 1918; National Society for the Study of Education, 1932; 1947). In the mid and latter parts of the 20th century, as nuclear proliferation continued and the United States and the Soviet Union battled over space, the centrality of science and technology to American geopolitical strength became increasingly apparent (DeBoer, 1991; Johanningmeier, 2010). It was significant then, when in 1983, the publication of *A Nation at Risk*: *The Imperative for Educational Reform* (National Commission on Excellence in Education,

1983) cast doubt on American's ability to compete in science and math. Considered to be one of the most noteworthy events in the history of the U.S. public education system, the 36-page report highlighted science and mathematics education as one of the key avenues through which the United States might continue its competitiveness on the global scene, while simultaneously suggesting a steady deterioration of American academic achievements in science (Klieger & Yakobovitch, 2011). *A Nation At Risk* created a notable public response and has been seen as one of, if not the primary, catalyst to creating the political will that led to subsequent decades of education reform movements, including an increased emphasis on improved standards for science education (Neumann, Fischer & Kauertz, 2010; Stevenson & Stigler, 1994).

While *A Nation at Risk* created the political will for large scale reforms, the reforms it ushered in very often materialized as predominantly content-focused standards, structured to disseminate discrete scientific knowledge to prepare students for international measures of science achievement (DeBoer, 1991; Wells, 2019). The resulting instruction was heavily dependent upon vocabulary and diagram memorization. Laboratory activities, if they existed, typically were of a so-called cookbook variety where students followed precise directions to arrive at predetermined outcomes (Bentley, Ebert & Ebert, 2007; Pruitt, 2014).

The Next Generation Science Standards

The Next Generation Science Standards (NGSS) are a more recent effort to improve K-12 science instruction and they are a significant departure from the prevailing content-focused standards because they place emphasis on developing conceptual and cross-disciplinary understandings as well as on the development of scientific skills and processes (Brunsell, Kneser & Niemi, 2014; Pruitt, 2014). First released in 2013, as of February 2020, twenty U.S. states had adopted the NGSS as the basis for their public school curriculum, and another 24 had developed their own standards based on the conceptual framework from which the NGSS is derived—the National Research Council's (NRC) *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core ideas* (National Research Council, 2012; NSTA, n.d.).

In developing *A Framework* (National Research Council, 2012), the NRC built upon major ideas from *Science for All Americans* (American Association for the Advancement of Science, 1990), *Benchmarks for Science Literacy* (1994), the *National Science Education Standards* (National Research Council, 1996), and other research conducted by the American Association for the Advancement of Science (National Research Council, 2012). Similar to prior reform movements the impetus for the NRC's development of *A Framework* came from indications that American students lagged behind their international counterparts in science achievement and were not being well-enough prepared for 21st-century economies. The NRC referenced a 2009 Carnegie Commission on Mathematics and Science Education report which stated:

> The nation's capacity to innovate for economic growth and the ability of American workers to thrive in the modern workforce depends on a broad foundation of math and science learning, as do our hopes for preserving a vibrant democracy and the promise of social mobility that lie at the heart of the American dream (Commission on Mathematics and Science Education, 2009, p. vii).

In creating *A Framework*, the NRC has interpreted the commission's call for a "broad foundation of math and science learning" (Commission on Mathematics and Science Education, 2009, p. vii) to be more than memorization of large quantities of scientific information. Rather, the NRC's framework seeks to build a base of scientific knowledge coupled with inquiry skills, understanding of science and engineering processes, and an ability to apply scientific concepts across disciplines. Accordingly, the Next Generation Science Standards, which are immediately derived from the NRC's conceptual framework, incorporate scientific inquiry and other scientific processes and skills in a way that previous content-heavy standards have not (Brunsell, et al., 2014.; Nollmeyer & Bangert, 2015; Pruitt, 2014). Each standard within the NGSS, referred to as student performance expectations (PE), has a three-dimensional nature whereby a discipline-specific core idea, a science or engineering-related practice, and a broader cross-disciplinary concept are all integrated into a single cohesive standard (Pruitt, 2014).

The three-dimensional nature entailed in each NGSS PE challenges science teachers to not only have a strong grasp of science content and practices, but also the pedagogical implications for the way the three dimensions are tied together (Bybee, 2014; Krajcik, 2015; Krajcik, Codere, Dahsah, Bayer &Mun, 2014; Windschiti & Stroupe, 2017). For example, using models effectively, using evidence as a basis for argumentation, incorporating engineering design, and constructing explanations of scientific phenomena are new and unique instructional techniques for many teachers (Bybee, 2014; Reiser, 2013). Consequently, science educators have expressed feeling unprepared to fully implement the NGSS. Haag and Megawon (2015) conducted a study to describe U.S. middle and high school teachers' preparedness to teach the NGSS. The mixed-methods study collected data from 710 middle and high school science teachers from 38 states and focused on three aspects of teacher quality related to the NGSS: teacher motivation to teach the NGSS, teacher preparedness to teach the NGSS, and how experience with modeling instruction affected motivation and preparedness to teach the NGSS. The study concluded that typically only teachers with significant amounts of professional development (PD) with modeling instruction felt prepared and motivated to teach the standards (Haag & Megowan, 2015). Harris, Sithole, and Kibirige (2017) conducted a similar study sampling teachers from across 16 states and found only about 50% of teachers considered themselves familiar with the NGSS. Wilde (2018) conducted a smaller study of high school science teachers in California—a state that had adopted the NGSS nearly five years earlier—which revealed that only 41% of teachers considered themselves highly familiar with how the NGSS PE related to student assessment (California Department of Education, 2015).

It should be self-evident that having many teachers unprepared, unmotivated, or even unwilling to use the new methods associated with the NGSS has significant repercussions for their successful implementation, because teachers are the primary mechanism by which any set of standards are ultimately enacted (Bybee, 2014; Haag & Megowan, 2015; Loveland, 2004). A key contributor to the development of the NGSS, Bybee (2014), conceded shortly after their release that teachers themselves might be the Achilles heel of the standards when he lamented, "The responsible individuals [teachers] have their ideas about teaching and learning and those ideas do not necessarily align with the NGSS." (p. 218). Furthermore, even while some educators welcome pedagogical changes others may quickly return to traditional teaching methods if not engaged with sustained support and accountability (Lam, Cheng & Choy, 2010). Effectively engaging, preparing, and sustaining teachers in their work to understand the new standards will be essential to the success of the NGSS (Bianchini & Kelly, 2003; Bybee, 2014).

Science Education Reform in American-Curriculum International Schools

The effects of U.S. reform movements in science education extend beyond the U.S. national boundaries because there are a number of schools around the world that employ U.S. trained educators to simulate home-country educational experiences for the children of American expatriates. Singapore American School (SAS), where I was employed at the time of this action research study, may be the quintessential American international school. Like many such schools, SAS was founded in the mid-1900s by an expatriate parent population seeking an educational experience that would allow their children to easily reassimilate upon return to their home countries. When SAS opened in 1957, just 105 students received instruction in small colonial-style bungalows. However, over the next several decades, as Singapore gained independence and transformed into a major economic force in Southeast Asia, the number of expatriates in the country also increased and the burgeoning expatriate population in Singapore likewise allowed SAS to grow its student population, expand its facilities, its academic and extracurricular offerings, and recruit quality educators from around the world. Currently, SAS serves approximately 4,000 non-Singaporean students hailing from 56 nations, including more than 2,300 from the North American continent. SAS students enjoy purpose-built facilities on a 36-acre campus situated approximately 10 miles north of Singapore's central business district (Singapore American School, n.d.-b).

Meanwhile, SAS has also established a reputation for academic excellence. In 2016, of all schools registered worldwide to offer Advanced Placement (AP) courses, SAS ranked in the 96th percentile for the percentage of students earning a score of three or higher on the AP exams. In the same year, students at the school ranked on average in the 94th percentile, or higher, worldwide in all subjects evaluated by the Northwest Education Association's Measure of Academic Progress assessment, and the senior class of 2016 had an average SAT score of 1930 which is 29% higher than the global average of 1490 and, as of January 2019, about 40 college-level courses were offered in the high school (Singapore American School, n.d.-a). Despite such success, SAS continues to pursue initiatives aimed at transforming the educational experience offered at the school in order to better prepare students for a rapidly changing and increasingly technology-driven global economy. One such initiative is the adoption of the NGSS to guide the school's K-12 science instruction.

As an eighth-grade science teacher at SAS, I welcomed the school's adoption of the standards during the 2015-2016 school year, and I was eager to cooperate with colleagues to implement them; shortly thereafter, I enrolled in Arizona State University's (ASU) Doctor of Education in Leadership and Innovation program with the intention to focus my dissertation research on PD practices aimed at improving NGSS implementation at the school. Concurrently with the beginning of my own cycles of action research for the ASU program, however, SAS also initiated its own regimen of

NGSS-focused PD activities. Through the subsequent years of working with colleagues in action research cycles I found that, increasingly, faculty at SAS were feeling overwhelmed by participating in NGSS-focused PD in addition to fulfilling other professional responsibilities and colleague requests[1]. Teachers recognized their need for, and indeed wanted, NGSS-related PD but they wanted it to focus on activities that helped them do better the things they were *already* doing, rather than being asked to do *new* things as add-ons. This sentiment is articulated further in chapter 2's section describing how previous cycles of action research influenced the design of the intervention for the dissertation cycle.

In August of 2018, a series of PD activities were initiated at a number of domestic and international schools that focused on evaluating the alignment of internal assessments (i.e. assessments designed by practitioners for use with students in their own classrooms) with the constructs of the NGSS. These activities were led by science consultant Paul Andersen, and involved the use of an assessment screening tool he and Lisa Brosnick, then president of the Science Teachers Association of New York State, had developed. By June of 2019, Andersen had led similar PD in at least 19 international schools in 15 countries, and teachers' reception of the PD seemed positive but anecdotal (P. Andersen, personal communication, Dec 19, 2018). This study entails scholarly investigation of how these PD activities, which involved a systematic reflection of

[1] Teachers were also feeling overwhelmed by requests to participate in other colleagues' doctoral research studies. SAS had recently organized a cohort of faculty to complete the University of Southern California's Doctor of Education program. At the time of this study's first cycle of action research, there were no less than 10 other doctoral students actively pursuing research projects involving students and faculty within the school.

internal assessments, affected teachers' understanding of the NGSS and its associated pedagogy. It is my hope that the results of this study will inform future cycles of action research as well as NGSS-related PD activities at SAS and peer institutions around the world.

Problem of Practice

Science teachers need help implementing the NGSS because the standards integrate science and engineering skills, broader conceptual understandings, and content knowledge ways previous standards have not, and they also task teachers to use unfamiliar pedagogical approaches (Bybee, 2013, 2014; Brunsell, Kneser & Niemi, 2014.; Nollmeyer & Bangert, 2015; Pruitt, 2014). Even those teachers who have been trained in science often lack experience with the sorts of authentic investigations envisioned in the NGSS (Kang, Donovan & McCarthy, 2018). These teachers need professional development activities that help them to understand the NGSS and their implications on pedagogy and assessment practices (Bybee, 2013, 2014; Haag & Megowan, 2015; Harris, Sithole & Kibirige, 2017).

Purpose of the Study

This study seeks to improve science education in American-curriculum international schools by contributing to the understanding of PD focused on the NGSS. The study specifically explores how science teachers in international schools come to understand the NGSS through PD that targets the design of NGSS assessments. The intervention in this study is a PD activity that engages teachers in a systematic reflection of NGSS assessments that have been developed by teachers for use with students in their own classrooms (i.e. internally designed). Findings from the study may be used to inform future PD activities in international schools.

Research Question

How does professional development mediated by the use of a screening tool (3D-PAST) enhance and/or challenge science teachers' understanding of the Next Generation Science Standards (NGSS) in American international schools?

CHAPTER 2

THEORETICAL PERSPECTIVES AND RESEARCH GUIDING THE PROJECT

This chapter begins with a description of the two bodies of knowledge that compose the conceptual framework guiding this investigation and the interpretation of the data that was collected. The first body of knowledge is pedagogical content knowledge (PCK). PCK has been embraced by the science education community as an important theoretical framework for researching the professional knowledge of science teachers (Abell, 2007; Chan & Hume, 2019). The Refined Consensus Model (RCM) of PCK for science instruction is presented as a model for considering PCK. Second, given their centrality to this project, the key components of the NGSS framework are used to guide the research. Subsequently, a review of the literature situating the study within the unique context of international schools, findings from previous cycles of action research, and a review of literature on the retrospective pretest-post design of quantitative data collection are presented.

Pedagogical Content Knowledge

Pedagogical content knowledge has been articulated as the unique domain of understanding at the intersection of teachers' content knowledge (CK) and pedagogical knowledge (PK) (Shulman, 1987). In Shulman's (1986) original formulation of PCK, content knowledge and pedagogical knowledge were mostly distinct and independent domains influencing PCK (Lederman & Gess-Newman, 1992; Shulman, 1986, 1987). CK was described as the factual, subject-specific expertise held by a teacher. For example, teachers of science may need to know facts about cell structure and laws of physics, how to operate microscopes or prepare biological specimens. A strong grasp of CK is considered a fundamental trait of effective teachers because teachers need content knowledge to make decisions about instruction, pose challenging questions which elicit students' critical thinking, contextualize facts and topics, and select appropriate materials (Anderson & Freebody, 2012; Ball, Thames & Phelps, 2008; Baumert, Kunter, Blum, Brunner, Voss, Jordan, Klusmann, Krauss, Neubrand & Tsai, 2010; Findel, 2008; Rovegno, 1995). Consequently, the acquisition of CK is typically a primary focus of teacher preparation and licensure programs (Howell, Cook, Miller, Thompson, Faulkner & Rintamaa, 2018; Ward, Tsuda, Dervent & Devrilmez, 2018).

However effective science teachers must possess more than content knowledge, they must also have a developed understanding of pedagogy, i.e. suitable methods of teaching (Bybee, 2014; Covay Minor, Desimone, Caines Lee & Hochberg, 2016; Kind, 2009; Shulman, 1987). In contrast to CK, PK encompasses this understanding of the many ways to create effective learning opportunities for students in different contexts. Among other things, PK may include knowledge of classroom management practices, learning processes, student characteristics, and methods of questioning and planning (Lederman & Gess-Newsome, 1992; Voss, Kunter, & Baumert, 2011). While teacher preparation programs may attempt to develop pedagogical knowledge outside of actual practice, it is recognized that a teacher's PK is primarily developed throughout the duration of a career and within the specific working environments a teacher experiences (Lederman & Gess-Newsome, 1992).

Shulman (1987) argued that the thing differentiating "a content specialist and a pedagogue" (p.8) is the extent to which they have developed an "amalgamation of content and pedagogy" (p. 8) which is able to elicit meaningful learning experiences for students; this amalgamation is the realm of PCK. Shulman (1987) described PCK as knowledge of "ways of which to represent and communicate a subject which makes it most comprehensible for others" (p. 9), and as the distinctive bodies of knowledge for teaching which represent the "blending of content and pedagogy into an understanding of how particular topics, problems, or issues are organized, represented, and adapted to the diverse interests and abilities of learners, and presented for instruction" (p. 8). It is important to note that PCK is domain specific. For example, a biology teacher's PCK is different from an English or history teacher's PCK, and even within a subject a teacher's PCK may vary by topic. For example, PCK for the teaching of genetics may be different than PCK for the teaching of cellular structure. It is also contextual; as teachers' knowledge of students and environments change, so also does a teacher's PCK. This is to say that as teachers' knowledge of content, students, and teaching context change, the methods they employ to help students learn particular content also changes.

Since its inception, however, various aspects of PCK have been debated. These debates have included whether it is a distinct body of knowledge, the extent to which PCK is a knowledge base or a skill set or both, what components should be included in the knowledge base of PCK, the extent to which PCK is context-specific or individual or collective, and the appropriate boundaries within which PCK can be considered (Chan & Hume, 2019; Krepf, Ploger, Scholl & Seifert, 2018). Unsurprisingly, then, the complexity of PCK has been investigated using a variety of models. For example, a study by Kind (2009) found at least nine different models of PCK that have been utilized to study teachers generally or science teachers specifically. Kind (2009) does note, however, that much of the variation in models are accounted for in the ways in which subcategories of knowledge within the PCK domain were classified. Nonetheless, despite these differences the concept of PCK as a distinct knowledge domain of teachers has generally been affirmed (Covay, et. al, 2016; Kind, 2009).

Refined Consensus Model of PCK for Science Instruction.

The Refined Consensus Model of PCK for science instruction seeks to address some of the challenges presented by earlier models of PCK. Developed from the contributions of more than two dozen researchers in science teacher education, the RCM seeks to provide researchers with a means to situate studies of student science learning in relationship to PCK, as well as a means to situate theories about the development of teacher PCK (Carlson & Daehler, 2019). The RCM is conceived as a dynamic layering of three distinct realms of science PCK: collective PCK (cPCK), personal PCK (pPCK), and enacted PCK (ePCK). At the center of the model, ePCK is the most specific and context-dependent realm, which consists of the knowledge utilized when a teacher is engaged in the practice of teaching (i.e. planning instruction, carrying out a lesson, etc.). Thus, ePCK is drawn from pPCK which is the larger reservoir of pedagogical knowledge and skills possessed by a teacher. Accordingly, pPCK is considered to be the knowledge developed over the course of a career through formal education, teaching experiences, and professional sharing. Both ePCK and pPCK are influenced by learning contexts which may include classroom environments and student attributes, school or district conditions, and the broader educational climate. In contrast to ePCK and pPCK, cPCK is the knowledge that has been developed and shared by the larger science research and education community, and is therefore more generalized and public (Carlson & Daehler, 2019). In addition to the knowledge present in a field's literature, cPCK may also include "a continuum of knowledge held by a group that extends what is present in the literature and recognizes that the knowledge about science teaching is also developed within school districts, school sites, departments, grade-level teacher teams, and professional learning communities" (p. 89). Figure 1 presents a graphical representation of the RCM.

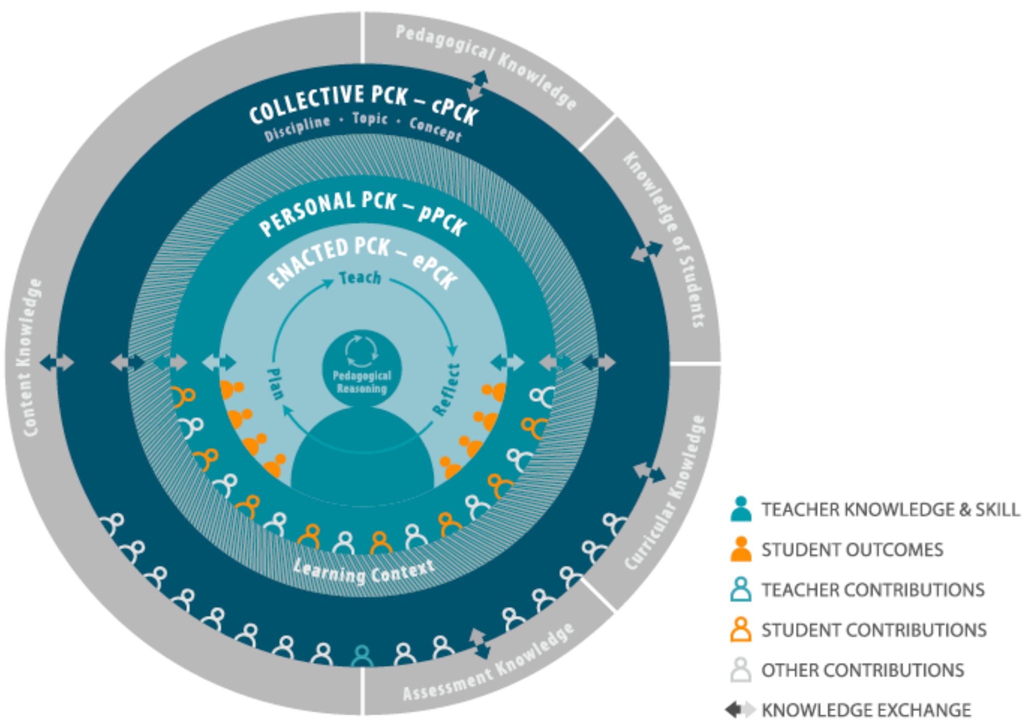

Figure 1. Representation of the Refined Consensus Model (RCM) of PCK. From *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science.* By A. Hume, R. Cooper, & A. Borowski, 2019. P. 83

Standards such as the NGSS occupy the realm of cPCK; they were developed through a collaboration of science and education experts, utilizing ideas drawn from a number of previous studies and science education reform efforts, and they are intended to be generally applicable for all science students (National Research Council, 2012). The RCM describes how, as cPCK, the NGSS may exist outside of and apart from an individual educator's knowledge and so do not impact classroom learning experiences until there is a knowledge transfer from the cPCK to the pPCK and then, ultimately, to the ePCK domain, where then a teacher's personal repertoire of PCK is enacted to have a direct impact on student learning. This translation of knowledge across PCK domains is a persistent problem in education and is frequently the goal of professional development programs for teachers (Hiebert, Gallimore & Sigler, 2002; Hume, Cooper & Borowski, 2019). In defining the boundaries of knowledge transfer, the RCM serves as a tool to investigate this problem (Hume, Cooper & Borowski, 2019).

Three-Dimensional Framework of the Next Generation Science Standards

Given their centrality to the project, the key concepts framing the NGSS also guide this research. The NGSS were derived from the conceptual framework articulated in *A Framework for K-12 Science Education* (2012), developed by the National Research Council Committee on A Conceptual Framework for New K-12 Science Education Standards. According to the framework, science literacy is described as composed of three distinct but integrated dimensions: disciplinary core ideas (DCI), science and engineering practices (SEP), and crosscutting concepts (CCC) (National Research Council, 2012; Pratt, 2013).

Disciplinary Core Ideas

Disciplinary core ideas constitute one dimension of the NGSS framework. DCI are the facts and conceptual understandings associated with the specific disciplines of physical sciences, life sciences, and earth and space sciences, as well as engineering and other applications of science. As put forth in the NGSS, DCI are not an exhaustive collection of the knowledge existing within a particular field, but are rather a limited set of knowledge and conceptual understandings selected because of their broad importance, ability to serve as key organizing principles of a discipline, or function as important tools for understanding more complex ideas. DCI may also be ideas of particular relevance to students because of their connection with societal or personal concerns. The NGSS also advance the notion that DCI should be accessible to younger students while also being broad enough to allow for progressively deeper investigation and understanding throughout students' K-12 experiences (National Research Council, 2012). DCI most closely correlate to what is typically considered content knowledge or subject matter knowledge in traditional understandings of K-12 science curriculum.

Science and Engineering Practices

A second dimension of the conceptual framework consists of common practices used by scientists and engineers. As they relate to science, these SEP are practices used to investigate the natural world, build models of concepts, and develop theories to explain phenomena. As related to engineering, they are a key set of activities engineers use to design and build systems (National Research Council, 2012). The SEP dimension consists specifically of eight practices considered essential for both scientists and engineers. Table 1 summarizes the science and engineering practices that are articulated in the NGSS.

Table 1.

Science and Engineering Practices Included in the NGSS Framework

Practice	Description
Asking questions and defining problems	Ability to create questions that can be answered empirically, and to define problems with identified constraints and criteria.
Developing and using models	Ability to use models to describe natural phenomena not observable with the naked eye, and to use models to analyze and test systems.
Planning and carrying out investigations	Ability to design and conduct systematic investigations with properly identified variables, and to use systematic investigations to test engineering designs.
Analyzing and interpreting data	Ability to use a variety of tools and processes to analyze data from scientific investigations and engineering problems.
Using mathematics and computational thinking	Ability to use mathematics to represent physical variables, construct simulations, identify quantitative relationships, and make predictions in physical systems.
Constructing explanations and designing solutions	Ability to construct logically coherent explanations from evidence, consistent with current scientific understanding, and propose solutions that balance competing criteria and constraints.
Engaging in argument from evidence	Ability to use data to defend reasoning and explanations, and critique proposed design solutions.
Obtaining, evaluating, and communicating information	Ability to effectively communicate scientific findings and their implications, using a variety of forms, including oral, written, graphical representations, equations, and discussions.

Note: Descriptions summarized from *A Framework for K-12 Science Education: Practices, Crosscutting Concepts and Core Ideas*, pp. 50-53. National Research Council, 2012.

Crosscutting Concepts

The third dimension of the NGSS framework is composed of concepts that have application across all disciplines of science. These crosscutting concepts are ideas that provide "organizational frameworks for connecting knowledge from various disciplines into a coherent and scientifically-based view of the world" (National Research Council, 2012, p. 83). Table 2 details the seven CCCs that are incorporated into the framework.

Table 2.

Crosscutting Concepts in the NGSS Framework

Concept	Description
Patterns	Regularly occurring shapes, structures, or processes.
Cause and Effect	Causal connections between two or more events.
Scale, proportion, and quantity	Variations in size and quantities.
Systems and system models	Closely related, but distinguishable, parts of objects, organisms or entities which have boundaries, resources, flow, and feedback.
Energy and matter	Inputs, outputs, and conservation principles of energy and matter.
Structure and function	Complementary aspects of objects, organisms, and systems.
Stability and change	Changing and unchanging conditions, systems, or processes. Equilibriums, feedback loops, and cyclical processes.

Note: Descriptions summarized from *A Framework for K-12 Science Education: Practices, Crosscutting Concepts and Core Ideas*, pp. 85-100. National Research Council, 2012.

Understanding American-Curriculum International Schools

While much of the impetus driving science education reform has been the perceived deficit position of students in American schools, there is a subset of schools straddling a fuzzy line between being American and not American, and they too are grappling with similar reforms. Overseas schools using an American-curriculum, are a subcategory of a group of schools known collectively as international schools. The context of international schools has garnered much less interest from scholars than other areas of education and, though the literature is rapidly increasing, it remains an under-researched field (Hayden & Thompson, 2008; 2013).

The Nature of International Schools

A precise definition of what constitutes an international school is debated (Hayden & Thompson, 2013; Heyward, 2002; Joneitz & Harris, 2012; Terwilliger, 1972). Historically, international schools were recognized as those schools established outside of a home country for school-age children of internationally mobile professionals (Fertig & James, 2016). These schools were typically non-profit organizations established by expatriate community members (Hayden & Thompson, 2013). Such schools are likely to follow a national curriculum of a particular expatriate nationality that is different from the national curriculum of the country in which they are hosted. For example, a school in Cambodia catering to primarily British expatriates would offer the British national curriculum, one in Tanzania serving predominantly German expatriates would offer a German curriculum, and so on. It is common for these schools to offer instruction either in the language associated with the type of curriculum (e.g. in a French curriculum school, French would be the primary language of instruction) or in English. While it is usual for these schools to incorporate multicultural and global perspectives into their instruction, their primary goal is to approximate—albeit with a usually more culturally and ethnically diverse student body—a home-country education in an overseas location (Nagrath, 2011; Tate, 2016; Waterson, 2016). This type of international school has been classified as type-A by Hayden and Thompson (2013). In the case of American-curriculum type-A schools, they are often accredited by stateside regional agencies such as the Western Association of Schools and Colleges (WASC) or the Middle States Association Commissions on Elementary and Secondary Schools (MSA-CESS) (Ortloff & Escobar-Ortloff, 2001).

Though at one time most international schools were of this type-A variety, the number of international schools, their scope of mission, type of clientele, and nature of their governance has expanded substantially in the last half century (Bunnell, 2014; Hayden, 2011). In addition to the aforementioned type-A schools, Hayden and Thompson (2013) have identified two other broad categories of international schools. Unlike type-A schools that have formed to primarily serve the market need of a particular expatriate population, type-B international schools are characterized by having been formed with the purpose of promoting a non-national ideology (Hayden & Thompson, 2013). Type-B schools are exemplified by institutions such as the United World Colleges (UWC) system of schools founded in 1967 around the philosophy of Kurt Hahn. Eighteen UWCs, in 18 countries on four continents, operate with the expressed purpose of bringing together students from diverse backgrounds to "unite people, nations, and cultures for peace and a sustainable future." (Hayden & Thompson, 2013; United World Colleges, n.d.-a; n.d.-b). New York's United Nations International School, the Yokohama International School, and the International School of Geneva are other notable examples of schools established under a similar premise (Fabian, 2016; Walker, 2016).

In contrast to type-A and type-B schools, type-C international schools tend to be aimed at host-country nationals and operate on a more commercial basis than either of the former (Hayden & Thompson, 2013). This category of school is a relatively recent development and it is diverse in its composition. It includes satellites of prestigious schools such as the United Kingdom's Dulwich College and Harrow School, Canada's Branksome Hall, Vermont's St. Johnsbury Academy and California's Chadwick School. It also includes chains of schools such as those operated by GEMS Education, Cognita, and Nord Anglia Education, as well as a variety of smaller groups, and schools that operate as individual entities (Bunnell, 2014; Hayden & Thompson, 2013; Waterson, 2016).

As there is no single organization that governs international schools, reliable statistics of the field can be difficult to obtain (Hayden & Thompson, 2013). However, estimates of the numbers, types, and demographics of the schools can be gleaned from any of the 15 or more regional organizations in which many of the schools collaborate such as the Near East South Asia (NESA) association, East Asia Regional Council of Overseas Schools (EARCOS), Association of International Schools in Africa (AISA) and the European Council of International Schools (ECIS). Various other entities, such as International School Services (ISS) and Search Associates, both of which assist international schools with recruiting faculty, also compile large amounts of information about their clients (Ortloff & Escobar-Ortloff, 2001). From 1964 to 1995 the number of international schools worldwide grew from around 50 to about 1000 schools (Hayden & Thompson, 1995). In 2015, the International Schools Consultancy Group estimated there were over 7500 schools classifying themselves as "international" and predicted that by 2025 there would be over eight million students being served by more than 15,000 international schools worldwide (Brummitt & Keeling, 2013). Nearly all of this growth, however, has been in the more commercially focused type-C schools (Waterson, 2016). In contrast, the growth of type-A international schools has been significantly less. The state of American-curriculum international schools, in particular, can be estimated using the available information from the U.S. Department of State Office of Overseas Schools, which keeps information on this type of school in order to aid American diplomatic and expatriate families during relocations (U.S. Department of State, n.d.-a).

In 2019, the Office of Overseas schools indicated there were about 1100 schools outside of the United States that may be viable options for American expatriates[2]. Of those, there were 193 schools that the Department of State had relationships with by means of either direct or indirect support arrangements; these schools are classified by the Department of State as "assisted" schools (U.S. Department of State, n.d.-b). To be an assisted school, the institutions must demonstrate that they operate with an American educational philosophy and use relevant pedagogical approaches as well as promote international understanding (Mannino, 1992). Though the Department of State may provide various assistances to these schools, with few exceptions, they nonetheless operate as independent non-profit entities and are governed by school boards composed of parents and other community members (Gillies, 2001; James & Sheppard, 2014; Ortloff & Escobar-Ortloff, 2001). Accreditation is usually important to these schools,

[2] The U.S. Department of Defense operates approximately 160 American schools in 11 countries under the Department of Defense Education Activity. These schools, known as the Department of Defense Dependent Schools (DODDS) are typically excluded—as they are here—from discussions of international schools because access is restricted to military families and a very small number of other restricted groups. Consequently, the DODDS schools are not viable options for most expatriates (DODEA, n.d.; 2008; Ortloff & Escobar-Ortloff, 2001).

and it is common that they are accredited by stateside regional agencies such as the Western Association of Schools and Colleges (WASC) or the Middle States Association Commissions on Elementary and Secondary Schools (MSA-CESS) (Ortloff & Escobar-Ortloff, 2001). In some cases, where schools receive grants, or have other unique assistance agreements, one school board member may be appointed by the U.S. Ambassador to the host country.

Regardless of whether the school is assisted or not assisted, operating expenses for type-A schools are typically funded via student tuition with families commonly paying upwards of $20,000 per year, per student (Fertig & James, 2016; MacDonald, 2006). For example, tuition and fees for high school students at Singapore American School are approximately $32,000 annually, at South Korea's Seoul Foreign School they are more than $34,000, while at Frankfurt International School in Germany they are approximately $27,000 per year, and costs approach $50,000 per year at The American School in Switzerland[3] (Frankfurt International School, n.d.; Seoul Foreign School, n.d.; Singapore American School, n.d.-c; The American School in Switzerland, n.d.).

In any school, the ability of administrators to hire quality teachers is an important professional skill, but this is particularly true in an international school market that is increasingly competitive and where—in order to facilitate their mission—there is an expectation for teachers to be native English speakers and western-trained, and where the pool of candidates is restricted to educators who are already expatriates, or have a keen interest to live and work outside of their home country (Garton, 2000; Hayden &

[3] International schools charge tuition in local currency, U.S. dollars, or some combination of the two. Costs shown have been converted to USD using approximate exchange rates as of Feb 3, 2020.

Thompson, 1995; Marzano, 2007; Nagrath, 2011). In fact, the recruitment of international school teachers has become an industry in-and-of itself, with various recruiting fairs and placement services run by multiple agencies. Three of the most influential—the Council for International Schools, International School Services, and Search Associates—organize a number of recruiting fairs from around October through May of each year. Teachers often travel great distances to attend these fairs where they are able to meet school administrators and secure employment for the following academic year (Hayden, 2006). For example, a single three-day fair held in Thailand by Search Associates, in January of 2019, attracted administrators from 140 international schools and over 500 job-seeking teaching candidates (Search Associates, n.d.). Meanwhile, International School Services boasts of having assisted more than 50,000 educators secure employment in international schools since the organization's founding in 1955 (International School Services, n.d.).

The International School Teacher

To the extent that a majority of the faculty in the type-A, American-curriculum international schools are native English speakers and western-trained, their professional training bears similarities to their domestic counterparts. Teachers in these schools hold at least appropriate Bachelor's degrees from accredited universities and they have normally completed teacher preparation courses in their home countries (Nagrath, 2011). Accordingly, they hold state teaching credentials or, if not American, a teaching credential from the appropriate authorizing body of their home country. Therefore, they have the same formal pre-service educational training as what would be expected of domestic teachers. Likewise, the majority of teachers maintain their domestic credentials using processes and professional development activities approved by their domestic accrediting agency. An example of how this is facilitated outside of the United States can be found in EARCOS's arrangements with the State University of New York (SUNY), which offers workshops at their annual spring teachers' conferences; these workshops are eligible for graduate-level credits and they are commonly used by U.S. certified educators to meet their certification maintenance requirements (East Asia Regional Council of Overseas Schools, n.d.).

Thus, while pre-service training and the certification maintenance practices of type-A international school teachers are typically quite similar to their domestic counterparts, other characteristics of the population are distinct. By and large, type-A schools have policies to restrict the hiring of teachers to those with at least two years of teaching experience in their respective subjects (Nagrath, 2011). It is therefore rare to find a rookie teacher in one of these schools, and this policy contributes to maintaining a population of teachers in these schools that tends to have more formal education than U.S. public school teachers; compared to 56% of U.S. teachers holding advanced degrees, approximately 70% of teachers in the international schools surveyed for this study hold advanced degrees (Snyder, 2018; various school websites)[4].

The more educated nature of the international school teacher serves in part to address a common contextual component of the international school—high expectations

[4] The figure of 70% was approximated by reviewing information about faculty posted on the schools' websites, however demographic information obtained from the participants of this study suggest the number of teachers holding advanced degrees may be more than 80% (see Appendix K).

from the parent community. As previously detailed, parents are often making substantial investments in sending their children to an international school and the parent population, which overwhelmingly consists of educated and successful professionals seeking their children's admission into top universities, exerts a corresponding results-driven pressure (Mancuso, Roberts & White, 2010). A survey of these schools shows that it is common to find them boasting of college acceptance rates at or approaching 100%, with many schools seeing students accepted into prestigious universities in the United States and United Kingdom (various school websites).

In contrast to their experience in the schooling systems of their home countries, international school educators have also indicated social and professional stressors derived from their unique social and work context. In many instances, international schools function as a nucleus of social interaction for the expatriate community. Teachers in these schools attest to more tightly bonded school communities and often describe colleagues as more akin to a second family than just associates. Due to their more tight-knit nature, working in these settings is often described by teachers as working in a fishbowl or as claustrophobic (Zilber, 2005). The effect is further complicated because international schools are often keen to employ educators with spouses who are also teachers (i.e. "teaching couples") because of the financial expedience it affords the school[5]. Consequently, it is ordinary to find a significantly large number of family connections amongst faculty within any one school. This hiring practice also translates into an increased number of students who are the children of

[5] Many international schools provide housing benefit in addition to salary; the cost of housing a teaching couple is often substantially less than housing two single teachers in separate residences.

faculty. For example, at SAS during the 2018-2019 academic year, 12% of the researcher's own student assignment were the children of colleagues or supervisors. As a result, educators in international schools are more frequently managing unique stresses that come from teaching the children of close friends or colleagues (Zilber, 2005).

Another often-cited stressor is the socio-economic positioning of the international school teacher in comparison to the larger expatriate community. While teachers in these schools earn salaries comparable to, or better than, what they would earn in public schools in the United States, it is often appreciably less than other members of the communities they serve, such as members of the corporate world or foreign service who have generous salary and benefits packages (Zilber, 2005). Teachers employed with international schools also face stressors associated with contractual arrangements different than those they have experienced domestically. Significantly, these teachers subsist with less job security as initial contracts offered by international schools are ordinarily two years, after which extensions are only offered on an annual basis. Feelings of insecurity that might come from short term contracts can be further exacerbated by the lack of teachers' unions which could serve to protect educators or act as mediators during professional disputes (Hrycak, 2015).

Further, although these schools are billed as American international schools, they are decidedly multinational and multicultural in student composition. For example, the schools in this study, on average, had student bodies representing 49 countries with only 38% of the students holding passports from North America. Consequently, teachers in these schools need to be capable of managing the particular challenges sometimes associated with interactions between cultures, languages, and learning styles (Halicioglu, 2015).

To summarize, teachers in type-A American-curriculum international schools tend to be very qualified and experienced educators, bearing similarities to their U.S. counterparts in education and training, but operating in quite different social and professional contexts. While this section has introduced a breadth of international school types, it is particularly with the type-A American-curriculum international schools that this study is concerned.

Previous Cycles of Action Research

This study is a continuation of two previous cycles of action research that occurred at Singapore American School. The first cycle occurred during the fall of 2017 and was considered a reconnaissance cycle. The goal of this cycle was to develop an initial understanding of how SAS science teachers perceived prior professional development activities to have affected their understanding of the NGSS. It also sought to understand what SAS teachers had perceived to be barriers to NGSS implementation. The second cycle sought to further investigate the first and third themes that emerged from cycle one, in order to establish a deeper understanding of the problem of practice which would, in turn, inform the selection of the intervention to be used in this study.

First Cycle: Understanding SAS Teachers' Perceptions

The first cycle of action research involved interviewing a small number of science teachers in Singapore American School's middle school division. Examples of questions from the interview include, "What do you perceive to be the greatest barriers to implementation of the NGSS for middle school science teachers?" and "What aspects of your previously completed professional development were unhelpful?" (see Appendix A). Interviews revealed three themes. First, teachers experienced a level of discomfort with some aspects of pedagogy considered important to implementing the NGSS. Second, teachers perceived a value in the use of an outside expert to train teachers. Third, teachers were most receptive to PD activities that they felt could be immediately translated into their professional practice.

Teachers' discomfort with NGSS pedagogy included multiple practices, but comments were heavily weighted towards the use of modeling practices in instruction. This theme aligns with literature that indicates teachers' need for significant amounts of professional development, especially PD which targets the use of modeling in instructional practices, before they feel prepared or motivated to teach the NGSS (Haag & Megowan, 2015; Harris, Sithole, and Kibirige, 2017; Wilde, 2018). Emerging as a second theme from the cycle was teachers' perceived value of using an outside expert to provide assistance in the development of their NGSS-related PCK. During interviews, teachers frequently mentioned they felt time spent working with an external consultant was beneficial. The third insight gleaned from the cycle was teachers' tendency to associate the immediacy by which a PD activity could be translated into professional practice with its overall value. Teachers expressed frustration with PD activities they perceived to have little impact on day-to-day professional practice. For example, one teacher specifically identified a significant amount of time spent debating with colleagues about which NGSS performance expectations should be taught in each middle school grade—a decision he felt was better to be made by the school's director of teaching and learning rather than by classroom teachers.

Second Cycle: SAS Teachers' Practices and Further Perceptions

Cycle two was structured as a sequential multiple-methods study whereby quantitative data was collected via a survey instrument, analyzed, and then interviews were conducted to gain further insight. The Science Instructional Practices Survey (SIPS) instrument (see Appendix B) developed by Hayes, Lee, DiStefano, O'Conner, and Seitz (2016) was used for teachers to indicate the frequency with which they incorporated specific science practices into classroom instruction. Practices identified on SIPS are grouped into six factors, the first four of which align strongly with *A Framework for K-12 Science Instruction* (National Research Council, 2012): instigating an investigation, data collection and analysis, critique and argumentation and explanation, modeling, prior knowledge, and traditional instruction.

Data from the SIPS led to three inferences. Firstly, teachers indicated they were engaging with most NGSS-aligned instructional practices on an at least sometimes basis, and in many cases were engaging those practices more frequently. Second, though some teachers had indicated feeling uncomfortable with modeling instruction during cycle one interviews, by the time the SIPS was completed approximately 14 months later, a relatively high frequency of use signaled an increased level of comfort with the practice. Thirdly, the least utilized instructional practices were those within the "instigating an investigation" category. Instructional practices in this category correlate with student-driven scientific inquiry.

Following analysis of the SIPS data, an interview guide was devised, and a subset of the SIPS respondents was interviewed. Interview transcripts were subsequently coded using a grounded theory constant comparative method, resulting in the emergence of three primary themes. The first theme correlated well with prior analysis of the SIPS data; just as the SIPS data indicated teachers were engaged in several pedagogical approaches associated with the NGSS, interviewed teachers typically perceived themselves as knowledgeable and skilled in a number of NGSS-associated practices. Counterintuitively, however, this perceived strength sometimes gave rise to a confidence inspired aversion; that is to say teachers' perceptions that they were already adept with NGSS pedagogy, coupled with what they felt might turn out to be an educational fad, manifested in a reluctance to invest earnestly in the NGSS. Interviewees also expressed difficulty discontent, whereby some discontent with the NGSS was associated with stress caused by barriers to its implementation. These barriers were identified by participants as a perceived lack of support, lack of NGSS resources, and fatigue associated with integrating the NGSS framework with other school initiatives and structures. Table 3 presents the themes, their characterizations, and related inferences.

Table 3.

Emergent Themes from Action Research Cycle Two

Theme	Characterized By	Inferences
Confidence Inspired Aversion	Teachers possess previous knowledge which suggests new initiatives will fail or are unnecessary. Teachers are confident with their current practices and do not perceive a need for change. NGSS is not seen as significantly different than current practices.	Teachers are reluctant to fully embrace the NGSS because they believe they are already doing or accomplishing much of what the NGSS is intended to accomplish. Teachers are also reluctant to put forth additional effort when the initiative is seen as a potential passing fad.
Difficulty Discontent	Further movement towards change is discouraged by the challenges of early adopters. Physical or time constraints have proved a barrier to NGSS implementation. Interpersonal conflicts. Curricular misalignment.	Initial enthusiasm around NGSS implementation was dampened by a lack of professional development, limited resources (time, or physical), and inherent challenges of navigating change in a collaborative setting.
Cautious Optimism	Acknowledgment of some beneficial aspects of NGSS. Perceived growth in professional efficacy due to NGSS focus and related professional development.	Teachers acknowledge how NGSS focus and related professional development has improved their craft in some areas, and they are keen to continue growth. Optimism is cautious rather than enthusiastic.

In contrast to the prior themes which present as barriers to implementation, the third theme to emerge was a cautious optimism about the NGSS. Teachers typically indicated that, while they had reservations about the NGSS and were challenged by its implementation, they perceived the underlying framework and skills-based focus of the NGSS as being a worthwhile pursuit for science instruction, and they were hopeful that implementation would improve with time and continued efforts.

From Cycle Two to the Current Study

In reflecting on cycle two, the theme of teachers feeling discontented because of difficulties arising from initiative fatigue was especially poignant. Initiative fatigue is a phenomenon where educators—having finite time, energy and emotional resources—lose the ability to retain focus, prioritize tasks, or engage in monitoring practices as initiatives are layered one after another (Reeves, 2006; 2012; Schmoker, 2018). In addition to the feelings of initiative fatigue expressed by colleagues during cycle two interviews, the researcher had also participated in a number of informal collegial conversations around the topic, and was increasingly cognizant of his own experiences with initiative fatigue. A deeper understanding of the extent to which initiative fatigue was effecting colleagues provided motivation to identify an intervention for this study that would respect colleagues' already stressed time and emotional reserves.

Near the same time as cycle two was finalized, science consultant Paul Andersen introduced a screening tool to SAS faculty that was designed to evaluate the quality of NGSS-aligned assessments. Use of the screening tool could easily be incorporated into typical professional responsibilities, thereby mitigating the risk of adding to colleagues' already high levels of initiative fatigue by limiting additional time requirements placed on teachers. The selection of the assessment screening tool PD activity, as the intervention for this study, had the additional benefit that it was being implemented at a number of SAS peer institutions around the world and thus provided an opportunity to consider its impact not only within the immediate SAS context, but the context of international schools at large.

Retrospective Pretest-Posttest

This study takes advantage of retrospective pretest-posttest (RPP) design for quantitative data collection. In the literature, RPP has also been referred to as post-then-pre, post-then, and then tests (Klatt & Taylor-Powell, 2005). RPP is a method of administering a pretest after an event, asking respondents to consider their pre- condition in a retrospective fashion, followed by a post-assessment in typical fashion. RPP administration of self-measure surveys has been put to use as a method of data collection, particularly when there is risk for response-shift bias (Allen & Nimon, 2007; Bhanji, Gottesman, Grave, Steinert, & Winer, 2012; Chang & Little, 2018; Sibthorp, Paisley, Gookin & Ward, 2007). Response-shift bias is a threat to internal validity described as the tendency for respondents to change the way they evaluate themselves during self-report measures because of new understanding gained along the way (Bhanji, et. al., 2012; Chang & Little, 2018; Klatt & Taylor-Powell, 2005). This response shift bias occurs when the participants' frame of reference at the time of the pretest is different than at the time of the posttest, rendering responses difficult or impossible to compare (Chang & Little, 2018). This is to say that a respondent may not be aware of how much they do or do not know about a concept until after an intervention has occurred, so the metric they use to judge their understanding prior to an intervention may be different than the metric that they use afterward.

An early proposal to employ RPP design to mitigate response shift bias was made by Howard, Ralph, Gulanick, Maxwell, Nance, and Gerber (1979), who examined several studies and concluded that RPP design produced more reliable self-reported results, as gauged by objective measures, than traditional pretest-posttest design. Decades later, Klatt and Taylor-Powell (2005) conducted a review of literature relative to RPP. After reviewing 49 published articles, they concluded that (a) there was evidence for the existence of response shift bias (b) the existence of a difference between traditional pretest-posttest and RPP results had been substantiated by others, notably Pratt, McGuigan and Katzev (2000), Cantrell (2003) and Schmidt and Nübling (2005) and (c) there was evidence for the validity of RPP (Klatt & Taylor-Powell, 2005) as an appropriate method of data collection in some contexts.

Of the studies Klatt and Taylor-Powell (2005) deemed noteworthy, Cantrell's (2003) work is particularly relevant to the current study as it pertained specifically to traditional versus RPP methodology in measuring science teaching efficacy beliefs. The study sought to examine the difference in results obtained from traditional pretests with those obtained from retrospective pretests among preservice science teachers enrolled in a science methods course. In the study, 36 participants enrolled in the course were administered a traditional pretest at the beginning of the course, followed by a retrospective pretest and a posttest on the last day of the course. The measure used was the Science Teaching Efficacy Belief Instrument (STEBI) Form B. The STEBI consists of 23 Likert-type response items designed to gauge science teachers' perceived efficacy and is similar to the NFSE-STU used in this study (Enochs & Riggs, 1990). Cantwell (2003) found statistically significant differences between the two types of pretests administered, with students rating themselves higher on traditional pretests than they did using the retrospective pretest. Further, the use of retrospective pretests yielded greater statistical power than traditional pretests as measured by effect size. Interviews with students subsequently revealed evidence of response-shift bias present in traditional pretests. Cantwell (2003) wrote, "The interviews revealed that the students seemed to doubt the validity of their initial [traditional pretest] responses because at that time they did not have enough information upon which to base their beliefs" (p. 181).

More recently, inquiry into the application and validity of RPP methodology has continued. Bhanji, Gottesman, de Grave, Steinert and Winer (2012) concluded that pretests given retrospectively to medical students following a four-hour pediatric resuscitation course were at least as accurate as traditional pretests in measuring understanding, but were better at eliminating distractors (i.e. perceived changes in understanding which would not have occurred due to the intervention). In a study gauging the effects of a course preparing students for inclusive classrooms, Miller and Hinshaw (2012) found RPP methodology revealed changes in attitudes not evident using traditional pretests. Little, Chang, Gorrall, Waggenspack, Fukuda, Allen and Noam (2019) strongly argued for the use of RPP as a "psychometrically and practically strong alternative" to traditional pretest design in repeated measures studies of attitudes, skills, and values as it is "ideally suited to reduce bias and to capture true change effects" (p. 8). Given the body of literature supporting the use of RPP, the methodology was identified by the researcher as appropriate and useful for this study given the context of the intervention.

CHAPTER 3

METHOD

This chapter provides a brief review of the purpose and research question, followed by a detailed description of the context and the methodology used for the research.

Purpose of the Study

Adoption of the Next Generation Science Standards (NGSS) as a framework for science instruction requires educators to have a strong grasp of science content and practices, as well as the pedagogical implications of the way the dimensions are integrated (Bybee, 2014; Krajcik, 2015). This study seeks to improve science education in American-curriculum international schools by contributing to the understanding of how professional development impacts teachers' understanding of the NGSS.

Research Question

How does professional development mediated by the use of a screening tool enhance and/or challenge science teachers' understanding of the Next Generation Science Standards (NGSS) in American international schools?

Context of the Study

The setting for this study's intervention was a group of 15 American-curriculum international schools. These schools have adopted the NGSS as their framework for K-12 science instruction and have initiated professional development activities to help teachers successfully implement the standards, some of which has focused on the alignment of internal assessments with the constructs of the NGSS. The schools are of the category described by Hayden & Thompson (2013) as type-A, meaning those that primarily serve the expatriate communities in their respective host countries. They are all independent, non-profit organizations run by school boards composed largely of students' parents, and are accredited by a U.S.-based accrediting agency such as the MSA-CESS, WASC, or AdvancED. To provide a more detailed picture of the schools, Table 4 provides key characteristics of the participating schools' faculty and student bodies, gleaned from school websites and other public sources available in January 2019.

Table 4.

Faculty and Student Body Composition of Participating Schools

	Faculty		Students	
School	Total	U.S.	Total	Nationalities
American International School Vienna, Austria	100	69	777	54
American International School Dhaka, Bangladesh	99	78	766	45
Hong Kong International School, China	251	142	2585	15
International School of Beijing, China	192	97	1695	40
Colegio Nueva Granada, Columbia	341	93	1777	40
American School of Guatemala, Guatemala	220	78	1510	24
American International School of Budapest, Hungary	118	61	855	58
American Community School of Amman, Jordan	70	47	751	48
American School Foundation of Monterrey, Mexico	262	60	2409	15
Colegio Franklin Delano Roosevelt, Peru	174	62	1740	46
Singapore American School, Singapore	383	227	3946	56
American School of Barcelona, Spain	118	57	844	56
American School of Dubai, United Arab Emirates	173	102	1800	76
American Community School of Abu Dhabi, United Arab Emirates	133	72	1222	60
The American School in London, United Kingdom	211	142	1343	70
Total	2845	1387	24020	

Note: Data is aggregated from school websites and *Directory of International Schools*. (2016). Princeton, NJ: International School Services.

Participants

Participants in this study were elementary, middle, and high school teachers who were tasked with using the NGSS for science instruction. Participants were English-speaking and typically held teaching credentials from the United States or Canada. All but two participants provided daily science instruction to an ethnically and culturally diverse student body composed predominantly of expatriate students. Two participants were identified as instructional coaches who were not responsible for daily teaching activities, but were tasked with learning the NGSS framework and then training and otherwise assisting fellow educators with NGSS implementation at their schools. The two instructional coaches participated in the intervention alongside classroom educators during the same time period.

Science education consultant Paul Andersen led the intervention. Andersen, a former biology teacher and Montana state teacher of the year, is also recognized by YouTube as a top 10 YouTube Edu Guru thanks to his educational videos which have been viewed millions of times ("About Paul Andersen," n.d.). Andersen has been a speaker at NESA and EARCOS educators' conferences and has been hired by at least 29 international schools to assist with implementation of the NGSS, sometimes spending up to four weeks at a site during a given school year (P. Andersen, personal communication, January 31, 2019.; Near East South Asia Council of Overseas Schools, n.d.; East Asia Regional Council of Overseas Schools, 2017). Between August 2018 and June 2019, Andersen travelled to each of the 15 schools in this study to provide consultation and lead professional development activities, one being the intervention that is the focus of this study (P. Andersen, personal communication, January 31, 2019).

Role of the Researcher

In this action research study, I assumed the dual role of participant in and researcher of the intervention. Coupled with my employed capacity as an eighth-grade science teacher at SAS, my positionality in this study is that of an insider—a practitioner-researcher seeking to understand the outcomes of a program in his own context (Herr & Anderson, 2005). I participated in the intervention delivered at SAS, engaged in reflective conversations with other science teaching faculty, and translated new understanding from the intervention into my own professional practice. I collected pre- and post-intervention data from participants at each school, maintained a research journal, and also conducted formal interviews with participants from 13 of the 15 participating schools.

Intervention

3D-PAST: Systematically Analyzing NGSS Performance Expectation Assessments

Central to this study is the use of a screening tool dubbed (for the purposes of this study) 3D-PAST, an acronym for three-dimensional performance assessment screening tool. 3D-PAST is a practitioner-developed tool created by Andersen and Brosnick (n.d.) to aid in providing useful feedback to teachers in workshops they had hosted during the summer of 2018 (P. Andersen, personal communication, February 1, 2019). The tool is available publicly via Andersen's website (www.thewonderofscience.com) and is free to use and share via the Creative Commons Attribution-NonCommercial-ShareAlike 4.0

international license. The license allows others to "remix, tweak, and build upon" the work, so long as the original author is credited and subsequent works are licensed under identical terms (Creative Commons, n.d.). During the 2018-2019 academic year, Andersen led teachers through the use of 3D-PAST with at least 19 American-curriculum international schools (P. Andersen, personal communication, February 1, 2019).

3D-PAST consists of an 11-point checklist with each point identifying a key component of NGSS assessment design (see Appendix C). The first six points address one or more of the dimensions integrated into each NGSS PE. For example, the first point asks teachers to consider whether "the [assessment] prompts match the science and engineering practice and engage students in sense making" whereas the fifth point asks teachers to consider whether "the [assessment] prompts include the crosscutting concept." (Andersen & Brosnick, n.d.). The remaining five points seek to address other aspects of science instruction considered to be best practices, including the use of grade-appropriate language, graphic organizers, and scientifically accurate information (P. Andersen, personal communication, February 1, 2019). Accompanying the checklist is also a brief set of instructions and an explanation of pertinent vocabulary.

Intervention Procedure

The intervention in this action research study was a professional development activity engaging teachers with 3D-PAST as they worked to design assessments aligned with NGSS constructs. The PD activity was delivered to groups of teachers organized either by grade level or division (i.e., elementary, middle, or high school). The size of the group during any one instance of the intervention ranged from as few as three teachers to as many as 18 teachers. Led by Andersen, teachers progressed through a four-stage process. In the first stage, teachers were tasked with designing an assessment for one or more of the NGSS PEs. In the second stage, teachers exchanged assessments and then role-played being a student as they completed colleagues' assessments; this role-play activity was intended to help teachers build a deep familiarity with the way in which students might interact with the assessment. 3D-PAST was then used to evaluate alignment of the assessments with NGSS dimensions and best practices. As teachers discussed and critiqued assessments, they were encouraged to refer to other NGSS literature, such as NGSS evidence statements, to further clarify their understanding. NGSS evidence statements are typically one- or two-page documents that provide detailed descriptions of the NGSS performance expectations (see Appendix D). The final stage of the intervention asked teachers to make appropriate revisions to their assessments in accordance with new understandings developed in previous stages of the intervention.

Minor variations in the intervention's procedure were implemented in order to accommodate site-specific contexts. For example, time constraints may have required assigning participants to rewrite assessments independently instead of in collaboration with Andersen and other colleagues, or teachers may have critically reviewed their own assessments rather than exchanging them with colleagues. Importantly, however, the critical stage as it pertains to this study—use of 3D-PAST—had little variance from school to school. Table 5 summarizes the process of the intervention and minor variations that may have occurred.

Table 5.

Intervention Process and Minor Accommodations

Stage	Activity	Context Specific Accommodations
Design	Design NGSS assessment.	Example assessment provided.
Familiarization	Exchange assessment.	No exchange.
	Assume the role of student in completing assessment.	Critical, thorough, read of the assessment.
Critique (Use 3D-PAST)	Systematically check assessment against 3D-PAST.	None.
	Reflect, discuss with colleagues.	
	Critique assessment.	
Revision	Revise/rewrite assessment with Andersen.	Revise/rewrite assessment independently.

Intervention Timeline

The intervention was delivered to teachers at 15 schools between August 2018 and June 2019. The first teachers to participate in the intervention were employed with the American International School in Budapest, while the final group of teachers were employed with Colegio Franklin Delano Roosevelt in Peru. Six of the 15 schools, Colegio Franklin Delano Roosevelt, Colegio Nueva Granada, American School of Barcelona, Singapore American School, American International School Dhaka, and the American Community School of Amman were visited by Andersen on multiple occasions during the academic year; when multiple site visits occurred, either different groups of teachers participated in the 3D-PAST intervention, or teachers may have had a second experience with using the 3D-PAST. Table 6 details the specific periods within which the intervention occurred at each of the schools.

Table 6.

Timeframe of the 3D-PAST Intervention

		Timeframe (*dd/mm/yy*)	
School	Country	From	To
American International School of Budapest	Hungary	13/08/18	17/08/18
American School of Guatemala	Guatemala	31/08/18	01/09/18
Colegio Franklin Delano Roosevelt	Peru	04/09/18	06/09/18
Colegio Nueva Granada	Columbia	10/09/18	14/09/18
American School of Barcelona	Spain	17/09/18	21/09/18
Singapore American School	Singapore	24/09/18	05/10/18
International School of Beijing	China	08/10/18	11/10/18
American International School Dhaka	Bangladesh	15/10/18	19/10/18
American Community School of Amman	Jordan	02/12/18	05/12/18
American International School Dhaka	Bangladesh	20/01/19	24/01/19
Singapore American School	Singapore	28/01/19	01/02/19
Hong Kong International School	China	12/02/19	15/02/19
American Community School of Amman	Jordan	18/02/19	21/02/19
American School in London	United Kingdom	25/02/19	01/03/29
American School of Barcelona	Spain	04/03/19	12/03/19
American International School Vienna	Austria	14/03/19	15/03/19
American School Foundation of Monterrey	Mexico	25/03/19	29/03/19
Colegio Nueva Granada	Columbia	01/04/19	05/04/19
American School of Dubai	United Arab Emirates	22/04/19	25/04/19
American Community School of Abu Dhabi	United Arab Emirates	28/04/19	30/04/19
Colegio Franklin Delano Roosevelt	Peru	27/05/19	30/05/19

Data Collection

This study utilized a multiple-methods data collection process whereby both quantitative and qualitative data were collected. Multiple-method approaches seek to overcome limitations of a single data collection method and provide a deeper understanding of a phenomenon than either qualitative or quantitative methods may provide on their own (Cresswell, 2013; Clark & Creswell, 2010; Cresswell & Clark, 2017; Johnson, Onwuegbuzie & Turner, 2007). The multiple methods of data collection in this study included the use of the *New Framework for Science Education Survey of Teacher Understanding* (NFSE-STU) (see Appendix E), described in the subsequent section, and semi-structured interviews of a subsample of the NFSE-STU respondents.

Data collection began with distribution of the NFSE-STU. Administrators at each school were contacted to seek approval for, and then help facilitate, the instrument's distribution. The administrators subsequently forwarded, via email, a request to complete the NFSE-STU to faculty who had participated in the 3D-PAST intervention with Andersen. The email request (see Appendix F) contained a link to an online version of the NFSE-STU, as well as a link to submit a consent form if teachers were willing to further participate in a semi-structured interview. Both the online NFSE-STU and interview volunteer form were hosted by Qualtrics.

Approximately 450 teachers, across 15 schools, received the request to participate in the study. The American International School Vienna and the American International School of Budapest had the smallest number of teachers receiving the request, with each having only eight secondary school teachers who had participated in the intervention. In contrast, Singapore American School had the largest number of teachers receiving the request; administrators identified 103 faculty members who had participated in the intervention. Table 7 displays information provided by administrators at the schools, indicating the number of teachers who were forwarded the email invitation to participate in this study. Eight administrators provided a breakdown by division, which has been provided in the table when available.

Table 7.

Number of Teachers Participating in the 3D-PAST Intervention.

School	ES	MS	HS	Total
American International School Vienna		3	5	8
American International School Dhaka	14		6[i]	20
International School of Beijing	2	6	11	19
Colegio Nueva Granada	41	5	6	52
American School in London	0	7	9	16
American School of Guatemala				53
Hong Kong International School	0		19[i]	19
American International School of Budapest				8
American Community School of Amman				12
American School Foundation of Monterrey				12
Colegio Franklin Delano Roosevelt	65	5	8	78
Singapore American School	85	9	9	103
American School of Barcelona				10
American School of Dubai				8[ii]
American Community School of Abu Dhabi	28	8	3	39
Total	235	43	57	457

[ii]Estimated based on interview participant feedback.
[i]Reported MS and HS teachers as a single group.

For the interviews, participants self-identified from amongst the group of teachers who received the request to participate in the NFSE-STU. Contact information and informed consent from these volunteers were similarly collected via an online form (see Appendix G). Volunteers were then contacted by email to schedule interviews (see Appendix H), with interviews subsequently taking place at regular intervals over a period of approximately one month. The first interview occurred May 4, 2019, and the final interview was conducted on June 5, 2019.

Instruments

Quantitative Instrument

The quantitative instrument was the 31-item *New Framework of Science Education Survey of Teacher Understanding* (NFSE-STU) developed by Nollmeyer & Bangert (2017), reformatted to be delivered online and in a retrospective pretest-posttest fashion. Designed as a measure of teacher understanding of the NGSS, the instrument may also be considered a measure of personal PCK in the way that pPCK has been envisioned in the Refined Consensus Model of PCK for science instruction.

According to Nollmeyer and Bangert (2017), development of the instrument commenced following a thorough review of relevant literature and using procedures well-established in the field. The instrument was piloted and subsequently validated using exploratory and confirmatory factor analysis and internal consistency testing. Review of the analysis lead to the production of the final validated NFSE-STU (Nollmeyer & Bangert, 2015; 2017). Three of the survey's constructs mirror the NGSS's three dimensions (SEP, DCI, CCC) while a fourth evaluates teachers' understanding of the integration of the three dimensions, and the fifth evaluates teachers' understanding of best practices in science education. All five NFSE-STU constructs mirror the constructs of the 3D-PAST.

Qualitative Instrument

The qualitative instrument used in this study was a semi-structured interview. In conducting the semi-structured interview, comments were elicited from participants that would provide a deeper and more nuanced understanding of how teachers were impacted by the intervention and, in turn, more fully address the research question (Brinkmann & Kvale, 2015). Interviews were guided by 13 questions (see Appendix I). Nine of the questions correlated directly with the constructs of the NFSE-STU. For example, the question, "How do you feel that using 3D-PAST to evaluate your NGSS internal assessments changed the way you understood the NGSS or some aspect of what the NGSS requires of you as a teacher?" correlated with the NFSE-STU construct addressing the integration of the NGSS's three dimensions. The remaining four questions related to more general effects of the intervention. For example, the questions, "Do you feel that 3D-PAST is a useful tool?" and "Would you continue to use it?" were more general in nature and did not directly correlate with the NFSE-STU, but they were intended to solicit deeper conversation about the impact of the intervention.

While the interview guide served as a roadmap for the interview, the researcher asked additional probing questions to seek clarification and to further illuminate how teachers experienced knowledge transfer between domains of PCK. Occasionally, to avoid redundancy, some questions were skipped if the interview participant had provided information in prior questions that satisfactorily answered subsequent questions from the interview guide. Interviews ended with an opportunity for participants to speak freely about their general experiences with the intervention.

Teachers at Singapore American School were interviewed in person, while interviews with teachers at other sites were conducted using Zoom video conferencing software. Face-to-face interviews were audio recorded using two audio devices for redundancy, whereas interviews conducted via the Zoom platform were video recorded and audio recorded with an additional device for redundancy. Immediately after each interview, thoughts and impressions were recorded in a research journal. The interviews were then transcribed for later analysis.

Analysis Process

Quantitative Analysis

Quantitative data analysis sought to identify statistically significant changes in the way teachers perceived their understanding of five NGSS-related domains: science and engineering practices, disciplinary core ideas, crosscutting concepts, how the three NGSS dimensions are integrated with one another, and best practices in science instruction. Data from the pre- and post-intervention surveys were exported from Qualtrics and then imported into SPSS v25 statistical software for analysis. Respondents' scores for each of the constructs were calculated utilizing SPSS's compute-variable function by summing values of the response items correlated with each construct of the NFSE-STU, then dividing by the number of items within the construct. Table 8 shows the correlation of the response items with the constructs on the NFSE-STU instrument.

Table 8.

NFSE-STU Construct-Response Item Correlation

Construct	Response Items	Number of Items
Science and Engineering Practices	1 through 11	11
Disciplinary Core Ideas	12 through 18	6
crosscutting Concepts	19 through 24	6
Integration of the Three Dimensions	25 through 27	3
Best Practices in Science Education	28 through 31	4

Internal consistency of the pre- and post-intervention surveys were assessed using Cronbach alpha coefficient analysis, and then the normality of distribution was determined by evaluating skewness and kurtosis (see Appendix J). The statistical significance of changes between teachers' pre- and post-intervention construct scores were then determined using a paired samples t-test. Finally, effect sizes were calculated using Cohen's d and interpreted whereby $d \geq 0.2$ = small effect, $d \geq 0.5$ = medium effect, and $d \geq 0.8$ = large effect (Thalheimer & Cook, 2002).

Individual response items were also analyzed for statistical significance. As individual response items on the NFSE-STU are Likert-type items, thus ordinal in nature, statistical significance was determined using the Wilcoxon signed-rank test, which is appropriate for ordinal and non-normal data (Field, 2013). Effect sizes were then determined using the normal approximation z to r method for Wilcoxon signed-rank tests, where small, medium and large effect size correlate to $r \geq 0.1$, $r \geq 0.3$, and $r \geq 0.5$ respectively (Pallant, 2013).

Qualitative Analysis Process

Following transcription of the interviews, participant identities were protected by employing pseudonyms and replacing school names with randomly assigned code numbers. In instances where participant responses revealed identifying information, sections of text were omitted and marked appropriately for future reference. Transcripts were then returned to participants for review and verification of transcription accuracy.

Transcribed interviews were then imported into Dedoose research software version 8.3.x (2019). Dedoose (2019) was used to analyze the interviews using a hybrid process of a priori and open coding. A priori coding is a practice where codes are established by the researcher prior to analysis in order to make correlations with existing literature or theoretical frameworks (Elliot, 2018; Brinkmann & Kvale, 2015). For this study, initial codes of pPCK and ePCK were used, respectively, to indicate when teachers' statements suggested their personal knowledge had changed, or when their statements suggested actual changes to professional practice resulting from the intervention. A priori subcodes were also created for each of the five constructs of the NFSE-STU (i.e., where teachers' statements indicated changes in pPCK or ePCK, the statements were coded according to the particular construct within which those changes occurred). For example, the statement, "There was definitely an improvement on the understanding of the desired outcome. What do we want students to be? What kind of scientists do we want our students to be?" was tagged with the codes pPCK and also I3D—as I3D is a code created to refer to instances when teachers' statements suggested a change in the way they came to understand the three-dimensional nature of the NGSS

PEs. Weighting was also applied to pPCK and ePCK codes to indicate the relative extent to which the statements signified positive or negative effects on the teacher. For example, -2 would represent a sentiment that suggests a strongly negative impact on teachers' pPCK, whereas +2 would indicate a strongly positive impact on teachers' pPCK. In the previous example, the pPCK code was weighted +2 because the teacher was especially emphatic about the perceived enhancement of her understanding.

Cresswell (2013) recommended that when predetermined codes are used, researchers should still remain open to considering emergent codes that develop a more nuanced understanding of participant views. Accordingly, in addition to the a priori codes anchoring the qualitative analysis, emergent codes were developed using principles of grounded theory and the constant comparative method (Corbin & Strauss, 1990; Glaser, 1965; Glaser & Strauss, 1967). The open coding process advanced a more thorough understanding of the way in which teachers' PCK was enhanced or challenged by participation in the 3D-PAST intervention (Corbin & Strauss, 1990; Strauss & Corbin, 1990). As emergent codes were developed, they were compared with each other and with a priori codes, then either designated as subcodes for a priori codes or, in cases where relevant but distinct themes emerged, grouped into new categories. For example, many educators specifically referenced changes in their use of graphic organizers as a result of the intervention; accordingly, the subcode "graphic organizers" was created under the a priori code "BP," because the use of graphic organizers was considered best practice in this study's framework. Table 9 details the seven a priori concept-driven codes used in the analysis of interviews, whereas the codes emerging from the interview analyses are presented in chapter 5.

Table 9.

Concept Driven Codes Used in the Qualitative Analysis

Code	Theoretical Framework Correlation	Characterization (Participant's statement indicates...)
pPCK	Personal pedagogical content knowledge	A way in which teachers' personal pedagogical content knowledge was enhanced or challenged by participation in the 3D-PAST intervention.
ePCK	Enacted pedagogical content knowledge	A way in which teachers' enacted practice was enhanced or challenged by participation in the 3D-PAST intervention.
SEP	Science and engineering practices	Changes in understanding about science and engineering practices associated with the NGSS. These practices include asking questions and defining problems, developing and using models, planning and carrying out investigations, analyzing and interpreting data, using mathematics and computational thinking, constructing explanations and designing solutions, engaging in argument from evidence, and obtaining, evaluating and communicating information.
DCI	Disciplinary core ideas	Changes in understanding about NGSS disciplinary core ideas. DCI constitute scientific facts, theoretical knowledge and conceptual understandings that underpin the various disciplines of science.
CCC	Crosscutting concepts	Changes in understanding about NGSS identified crosscutting concepts. Crosscutting concepts are broad organizational frameworks that serve to connect knowledge from different disciplines.
I3D	Integration of the three dimensions	Changes in understanding about the way in which the three dimensions of the NGSS framework are intended to be cohesively integrated in instructional and assessment practices.
BP	Best practices	Changes in understanding about pedagogical approaches to teaching and assessing science which are not explicitly linked to the NGSS framework, including but not limited to use of graphical organizers, phenomena-based instruction, backwards-design practices, etc.

Note: Detailed description of each SEP can be found in Table 1.

Trustworthiness

Trustworthiness of the study was enhanced by the use of member checks and triangulation. Member checking served to verify the interview data and to mitigate researcher bias in interpretation (Morse, Barrett, Mayan, Olson & Spiers, 2002). Prior to coding of the transcripts, they were sent to participants to review for accuracy and then appropriate corrections were made, if requested. After coding, excerpts were then also sent to participants who were asked to comment if they felt codes had been applied inaccurately or in a way that misrepresented the participants' intended meaning. Participant feedback was considered and codes were adjusted accordingly prior to the final analysis.

In addition to member checking, the multiple methods of data collection allowed the use of triangulation to improve the study's trustworthiness; triangulation is the process of validating evidence by making comparisons across multiple sources (Cresswell & Clark, 2017). This study utilized survey responses, interviews, and a researcher journal. Each data set was analyzed individually and then cross-checked with the other data sets; they were also compared against pre-existing literature in the field. Rossman and Rallis (2016) also included appropriate human subject protocol as an important aspect of research trustworthiness. To ensure appropriate procedures, this study adhered to the methodology approved by Arizona State University's Institutional Review Board on April 17, 2019 (see Appendix K).

CHAPTER 4

FINDINGS

This chapter presents how the intervention, a professional development activity mediated by the use of a three-dimensional performance assessment screening tool (3D-PAST), enhanced or challenged science teachers' understanding of the Next Generation Science Standards (NGSS) in American international schools. Findings from the NFSE-STU data are presented first, followed by the themes emerging from the interviews, which are presented and correlated with the NFSE-STU data.

Quantitative Results

Data collected from 84 participants showed that teachers perceived the 3D-PAST intervention to improve their understanding in each of the five NGSS-related areas measured by the NFSE-STU: science and engineering practices, disciplinary core ideas, crosscutting concepts, integration of the NGSS's three dimensions, and best practices in science education. Of these, the intervention's greatest effect was found within the construct of science and engineering practices. Closer inspection of the SEP construct showed the intervention to be particularly useful in advancing teacher understanding in two areas: how to engage students with questions about scientific phenomena, and in deepening teacher understanding of how to engage students with scientific modeling practices. In addition, the intervention's effect on teacher understanding of the NGSS's three-dimensional nature was found to be large. Within the I3D construct, data suggests the intervention improved understanding of how teachers might encourage students' exploration of DCI by utilizing the other dimensions of the NGSS (i.e., by way of engaging students with scientific practices and by making connections through crosscutting concepts). Closely related to this, the data also suggests the intervention had a similarly positive, but slightly more modest, effect on how teachers understood the nature of the crosscutting concepts themselves. Of the five constructs measured by the NFSE-STU, data for the DCI and BP constructs showed the intervention to have the least effect, though still positive and appreciable.

The remainder of this section provides further detail regarding the analysis of the NFSE-STU data. Demographic information of the respondents is provided first, followed by descriptions of the tests that were used to determine that data's statistical significance and the intervention's effect size for each of the five constructs measured by the survey instrument.

Respondent Demographics

Of the approximately 450 teachers who received the request to participate in the survey, 85 completed the NFSE-STU via Qualtrics. One of the respondents—though he completed the survey—indicated he had not used the 3D-PAST, thus his responses were invalid for the purpose of the study and were removed from the data set. Consequently, data from 84 respondents were used in the final analysis. Demographic data were voluntarily provided by 76 of the respondents. Of these, Singapore American School faculty accounted for the largest number of participants, n = 30, while the International School Dhaka, American School in London, and Colegio Franklin Delano Roosevelt each only had one participant indicate employment with those schools. Also, of the respondents providing demographic data, 41% were elementary school teachers, 22%

were middle school teachers, and 37% were high school teachers. Approximately 82% of respondents held graduate degrees with 4% indicating they held terminal degrees. Respondents had, on average, nine years of experience teaching in an international school setting (see Appendix L for detailed demographic information of NFSE-STU respondents).

Internal Consistency

The NFSE-STU data were found to have good internal consistency as verified by Cronbach alpha calculations. Good internal consistency means the response items used to construct the scales of the survey are, indeed, all measuring the same construct (Streiner, 2003). Specifically, in both pre- and post-intervention surveys, all five constructs were found to have coefficient alphas where $\alpha \geq 0.8$, with the best practices construct for the post-intervention data found to have the lowest internal consistency, $\alpha = 0.815$, and the science and engineering construct for the pre-intervention data found to have the highest internal consistency, $\alpha = 0.935$ (see Appendix M). While there is not a universally accepted convention for qualifying coefficient alpha, the aforementioned values would typically be considered either “good,” “high,” or “excellent” (Taber, K. 2018). Having found the NFSE-STU data to show good internal consistency in this instance, they were deemed suitable for further analysis.

Descriptive Statistics

Descriptive statistics were then computed in order to summarize the distributions of data and for use in subsequent statistical significance tests. Mean scores were calculated for each construct from pre- and post-intervention responses. In each case, the mean scores were shown to increase from pre- to post-intervention, indicating that teachers perceived their understanding of each construct to have had a positive change resulting from the intervention. The largest changes occurred within science and engineering practices (SEP) and teachers' understanding of the way in which the NGSS's three dimensions are integrated (I3D). In both cases, teachers rated their understanding 13% higher following the intervention; for SEP the difference between the pre-intervention mean score ($M = 3.25$, $SD = 0.95$) and post-intervention mean score ($M = 4.02$, $SD = 0.73$), was 0.78 on the 6 point Likert scale, and for the I3D construct the difference was 0.76 (pre: $M = 2.98$, $SD = 1.13$; post: $M = 3.75$, $SD = 0.99$). Interestingly, though both the SEP and I3D constructs were rated by teachers as having similarly large changes, teachers indicated that their initial understanding of the I3D construct was the lowest of the five constructs measured by the NFSE-STU. In contrast, the smallest change of 7% ($diff = 0.43$) occurred in the best practices in science education (BP) construct, which had the second highest pre-intervention rating (pre: $M = 3.64$, $SD = 1.08$; post: $M = 4.07$, $SD = 0.89$). Also noteworthy, the DCI construct was rated by teachers as the area where they had the highest initial understanding. Figure 2 provides a graphical representation of the pre- and post-intervention levels of understanding for each construct. Table 10 shows the descriptive statistics and the differences in mean values for the pre- and post-intervention responses.

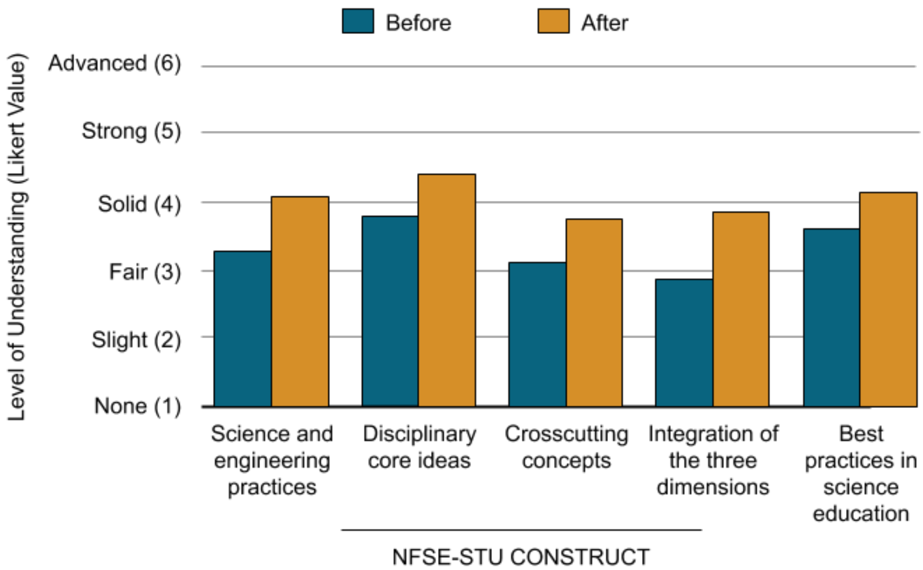

Figure 2. Comparison of pre- and post-intervention levels of understanding.

Table 10.

Descriptive Statistics for Pre- and Post- Intervention Constructs

	Pre-Intervention			Post-Intervention			Mean Difference	Percent Change
Construct	N	Mean	SD	N	Mean	SD		
Science and Engineering Practices	84	3.25	0.95	84	4.02	0.73	0.78	13
Disciplinary Core Ideas	81	3.81	0.96	82	4.29	0.73	0.49	8
Crosscutting Concepts	78	3.14	1.03	80	3.69	0.87	0.54	9
Integration of the Three Dimensions	77	2.98	1.13	78	3.75	0.99	0.76	13
Best Practices in Science Education	78	3.64	1.08	79	4.07	0.89	0.43	7

Paired Sample t-Tests

In order to verify whether the observed differences between pre- and post-intervention scores were statistically significant, paired sample *t*-tests were conducted using an $\alpha = 0.01$ level of significance. The tests were conducted for each of the five constructs and found to be statistically significant in every case. In other words, the *t*-tests indicated there is only a 1% chance that the differences observed in the data occurred merely by chance. Instead, there is a high level of confidence that the data reflects, on average, a true change in understanding as measured by the instrument and reported by participants. Table 11 shows the full results of the paired sample *t*-tests for each construct measured by the NFSE-STU.

Table 11.

Paired Sample t-test Results for NFSE-STU Data

Pair	M	SD	.99% CI Lower	Upper	*t*	df	*p* value (two tailed)	*r*
SEP Post - SEP Pre	0.78	0.64	0.77	0.78	11.05	83	0.00	0.74
DCI Post - DCI Pre	0.47	0.68	0.47	0.47	6.21	80	0.00	0.71
CCC Post - CCC Pre	0.55	0.65	0.55	0.56	7.50	77	0.00	0.78
I3D Post - I3D Pre	0.79	0.80	0.79	0.79	8.64	76	0.00	0.72
BP Post - BP Pre	0.46	0.82	0.46	0.46	4.97	77	0.00	0.66

Effect Size

Once the statistical significance of the pre- and post-intervention scores had been established, the intervention's effect size was calculated using Cohen's *d,* and then interpreted using guidelines ($d \geq 0.2$ = small effect, $d \geq 0.5$ = medium effect, $d \geq 0.8$ = large effect) reported by Thalheimer and Cook (2002). Accordingly, the intervention can be considered to have had a large effect on teacher understanding of science and engineering practices ($d = 1.21$), crosscutting concepts ($d = 0.85$), and the integration of the NGSS's three dimensions ($d = 0.98$), while it can be considered to have had a medium effect on teacher understanding of disciplinary core ideas ($d = 0.69$) and best practices in science education (d = 0.56). It is worth noting that the smallest effect sizes were found for the DCI and BP knowledge constructs, which were themselves rated highest in pre- and post-intervention understanding by teachers. In other words, the data suggests that the domains in which teachers perceived themselves to have had the greatest initial strengths were also where they felt the intervention to have had the least impact. Table 12 displays the results of the effect size calculations for each construct.

Table 12.

Effect Size of the Intervention by NFSE-STU Construct (Cohen's d)

Construct	M	SD	*d*	Effect size
Science and engineering practices	0.78	0.64	1.21	Large
Disciplinary core ideas	0.47	0.68	0.69	Medium
Crosscutting concepts	0.55	0.65	0.85	Large
Integration of the three dimensions	0.79	0.80	0.98	Large
Best practices in science education	0.46	0.82	0.56	Medium

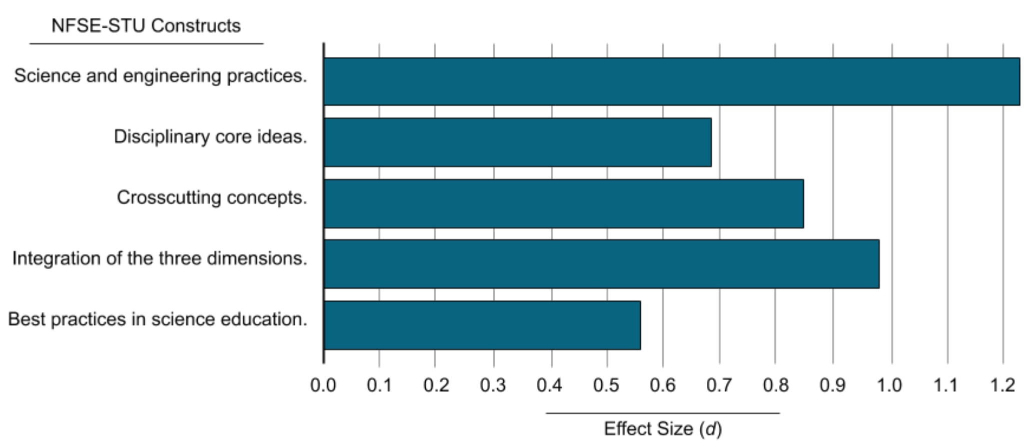

Figure 3. Effect size of the intervention on teachers' understanding of the NGSS by NFSE-STU construct.

Individual Response Items

In addition to the construct analysis, a higher resolution exploration was conducted by analyzing individual response item data from the NSFE-STU. Instead of paired sample *t*-tests, Wilcoxon signed-rank tests were used to determine statistically significant differences between pre- and post-intervention responses. In all instances, the difference between pre- and post-intervention mean scores for each of the 31 response items was found to be statistically significant, with effect sizes ranging from small to large.

Notably, within the SEP construct teachers indicated the intervention to have the largest effect on their understanding as it relates to engaging students with scientific phenomena and developing conceptual models. By contrast, they indicated the intervention's smallest effect to be related to developing students' computational thinking. Specifically, two items, "When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration" ($Z = 6.936$, $p < .000$, $r = 0.54$), and "When planning and teaching, educators have students develop and refine conceptual models to express their understanding of scientific phenomena," ($Z = 6.804$, $p < .000$, $r = 0.53$), showed the largest effect sizes. The smallest effect size occurred with a third item, "When planning and teaching, educators have students apply mathematical and computational thinking to investigate scientific questions and engineering problems" ($Z = 3.71$, $p < .000$, $r = 0.29$).

For the DCI construct, data for individual response items showed the intervention to have more modest effect; there were no items with a large effect size and two of the seven items comprising the construct had only a small effect size. The item, "When planning and teaching, educators include core ideas that relate to the interests and life experiences of students and societal concerns" ($Z = 3.575$, $p < .000$, $r = 0.28$), showed a small effect size, whereas the item, "When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students" ($Z = 3.455$, $p < .001$, $r = 0.27$) also showed small effect size. It is noteworthy that the last-mentioned item was the response with the smallest effect size of any item measured by the instrument. Together, the relatively smaller effect sizes for these items contributed to the overall DCI construct being shown as the area where the intervention had least impact on teachers' understanding.

Pertaining to the remaining items, the intervention was found to have a medium effect in all instances but two. Within the I3D construct the item, "When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts" ($Z = 6.117$, $p < .000$, $r = 0.5$), showed a large effect, whereas within the BP construct the item, "When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigation and engineering design" ($Z = 3.527$, $p < .000$, $r = 0.28$), indicated the intervention to have only a small effect. Figure 3 depicts the intervention's effect size on the seven specific areas described in this section. A table of the complete set of Wilcoxon signed-rank tests can be found in Appendix N.

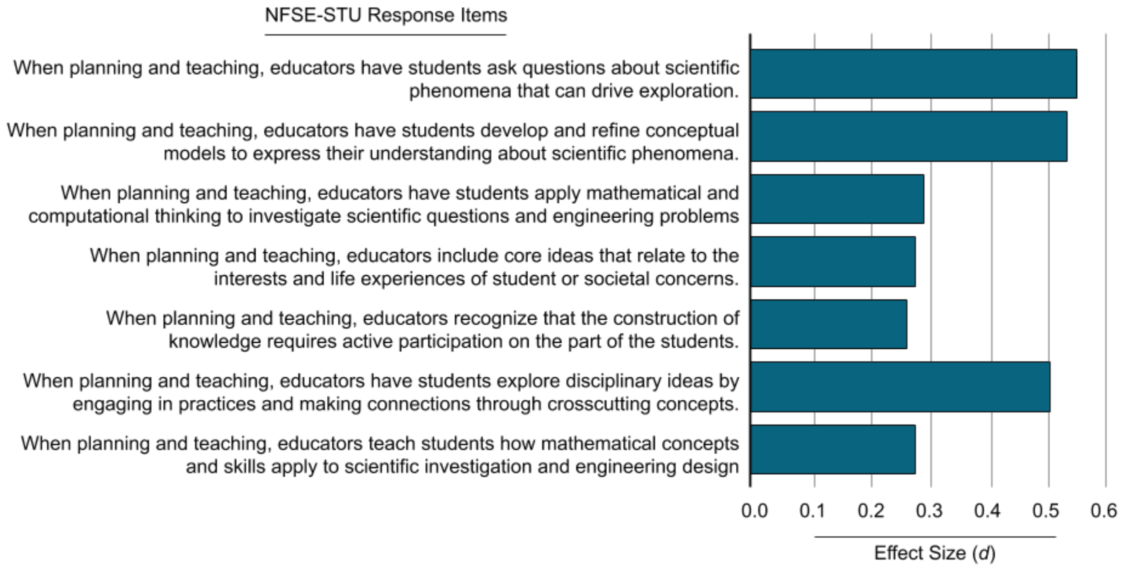

Figure 4. Effect size of the intervention for notable NFSE-STU response items.

Summary of Quantitative Results

Responses to the NFSE-STU suggest teachers perceived the professional development activity to have positively affected their understanding of the NGSS's science and engineering practices, disciplinary core ideas, crosscutting concepts, the way the NGSS's three dimensions are integrated, and best practices in science instruction. Of these, the two areas where the intervention was most effective were within the constructs of science and engineering practices and the integration of the three dimensions. Particularly, the data identifies the use of scientific phenomena and conceptual modeling with students as discrete areas of teacher understanding that were especially positively affected by the intervention. In contrast, the data suggests only a small positive effect on skills and practices related to mathematical applications as they relate to either the SEP or BP construct. Similarly, within the DCI construct, the data indicates teachers perceived the intervention to have had only a small effect on their understanding of how they might encourage students to construct knowledge through active participation.

Qualitative Results and Correlation with NFSE-STU Data

Semi-structured interviews were also conducted with a subsample of participants in order to provide a more nuanced understanding of how the intervention enhanced or challenged science teachers' understanding of the NGSS. Four themes emerged from an analysis of 29 interviews: (a) teachers perceived the use of the 3D-PAST as most valuable in the way it enhanced their understanding of the NGSS's three-dimensional nature and in the way it highlighted a persistent neglect of the NGSS's crosscutting concepts, (b) using the 3D-PAST elicited later changes in teachers' professional practice, (c) the use of the 3D-PAST helped to facilitate collaboration in professional practices, and (d) teachers' relationship and prior experiences with Andersen, the intervention moderator, influenced the way they perceived their experience with the intervention itself. Table 13 provides the key components and domain of effect for each theme. The remainder of this chapter articulates the development of the themes and triangulates the findings with those from the prior analysis of the NFSE-STU data.

Table 13.

Emergent Themes

Theme	Related Components	RCM Domain
Use of the 3D-PAST was perceived as most valuable in enhancing understanding of the integration of the NGSS three dimensions (I3D) and a neglect of the CCC dimension.	Teachers became aware of a misalignment of NGSS framework and assessment practices, particularly CCC dimension.	Personal pedagogical content knowledge (pPCK)
	Teachers became more familiar with "what an NGSS assessment looks like."	
	During the intervention teachers sought clarification of NGSS PE by referring to evidence statements.	
Participating in the intervention elicited changes in teachers' professional practice.	Teachers described an increased focus on SEPs and CCCs in assessment and instruction.	Enacted pedagogical content knowledge (ePCK)
	Teachers described increased use of some best practices in science education (i.e. graphic organizers, authentic data, use of common language, and backwards design principles).	
3D-PAST use helped facilitate collaborative professional practice	Teachers perceived the 3D-PAST to enhance impartial and impersonal professional critique.	
	Teachers perceived the 3D-PAST to encourage the use of common language.	
Teachers' prior experience with Andersen influenced their perceived experience with the intervention.	Teachers were challenged to disassociate the 3D-PAST intervention with other PD activities presented by Andersen.	Learning Context
	Teachers' perception of the value of the intervention correlated with their prior experiences and personal relationships with Andersen. This was true in both a positive fashion and also a negative fashion.	

Demographic Information of the Interviewees

Twenty-nine interviews were conducted with teachers from 12 of the 15 schools where the intervention took place; there were no teachers who volunteered to be interviewed from the American School in London, American School of Guatemala, or Colegio Franklin Delano Roosevelt. Elementary, middle, and high school levels were nearly equally represented with 9, 9, and 11 teachers respectively, as were genders with 14 male and 15 female interview participants. Approximately 11 hours of interviews were conducted, lasting between 11 and 41 minutes each, with a median length of approximately 21 minutes. Descriptives of the interview participants are provided in Table 14.

Table 14.

Descriptives of Interview Participants by School

		Grade Level Taught		
School	N	ES	MS	HS
American International School Vienna	2	0	1	1
American International School Dhaka	2	0	1	1
International School of Beijing	3	0	3	0
Colegio Nueva Granada	3	2	0	1
Hong Kong International School	2	0	0	2
American International School of Budapest	1	0	0	1
American Community School of Amman	2	2	0	0
American School Foundation of Monterrey	3	2	1	0
Singapore American School	8	3	2	3
American School of Barcelona	1	0	0	1
American School of Dubai	1	0	0	1
American Community School of Abu Dhabi	1	0	1	0
Total	29	9	9	11

Coding

The predetermined (a priori) codes—detailed in chapter 3—were applied to participant statements when they suggested ways in which the participants perceived the intervention to have impacted their personal pedagogical content knowledge (pPCK) or enacted pedagogical content knowledge (ePCK). Simultaneously, those statements were also coded as they correlated with the SEP, DCI, CCC, I3D, or BP constructs. Emergent codes were also created for instances when there were statements made by participants that expanded upon, or provided more detail for, the a priori codes.

For example, as a subcode of the SEP code, the code *modeling* was created to identify instances when educators referred to ways in which their understanding or use of scientific modeling had changed. Likewise, subcodes of *authentic data*, *graphic organizers*, *backwards design* and *common language* were developed as subcodes to the *BP* code, because a number of educators referred to these specific practices as ways that their understanding, or their professional practice, had changed as a result of the intervention. In some cases, new parent codes were also created when important ideas emerged that were not clearly aligned with established definitions of the a priori codes, but were nonetheless relevant to the way teachers experienced the intervention. For example, the codes *evidence statements* and *other PD* were created to identify statements, respectively, where teachers expressed value in using the NGSS evidence statements as part of the intervention, or they correlated their learnings from other PD activities with those they experienced during the study's intervention. Also, the parent code *peer protocol* was established when a number of teachers expressed that using the 3D-PAST improved collaborative practices. Finally, statements that illustrated the influence of Andersen's pre-existing relationship with teachers on the way they perceived the intervention warranted the development of the parent code *mentor relationship* with subcodes *accountability* and *character influence*. As the emergent codes were developed, they were compared and cross-checked against existing codes, then subsequently refined and aggregated into the final set of codes that informed theme development. Table 15 details the emergent codes resulting from the interview analyses.

Table 15.

Codes Emerging from Open Coding of Interviews

Emergent Codes		
Parent	Sub code	Teacher's statement indicates:
Peer Protocol		Effect on collaborative professional practice.
Mentor Relationship		Influence of pre-existing relationship with Andersen on perceived experience of the 3D-PAST intervention.
Mentor Relationship	Accountability	Feelings of accountability due to a relationship with Andersen.
Mentor Relationship	Character Influence	Personal sentiments towards Andersen.
SEP	Modeling	Changes in use/understanding of modeling instruction.
BP	Graphic Organizers	Change in the use/understanding of graphic organizers in instruction/assessment.
BP	Backwards design	Lesson planning influenced by considering alignment with 'end goal' assessments.
BP	Common Language	Alignment of language to reflect NGSS.
Other PD		Learnings from professional development activities other than the intervention.
Evidence Statements		Use of evidence statements to aid NGSS understanding.

Note: Codes in italic are existing a priori codes

Enhancing Understanding of NGSS's Three-Dimensional Nature.

The first theme to emerge from interviews was that teachers perceived the intervention to be valuable in furthering their understanding of how the NGSS integrates the SEP, DCI, and CCC dimensions; particularly with regard to the CCC, the PD activity highlighted the dimension's persistent neglect among teachers in their professional practice. Interview elements that contributed to the theme included the following: (a) teachers expressing new awareness of how assessment and teaching practices were not well aligned with all three dimensions of the NGSS framework, (b) teachers describing how they felt once they had become more familiar with "what an NGSS assessment looks like," and (c) teachers recounting clarifying events that occurred during the intervention. Teachers' perceptions that the intervention positively impacted their understanding of the NGSS correlated well with findings from the NFSE-STU data that similarly indicated a large effect within the I3D construct (refer to Table 13). Participant reports of the intervention's tendency to highlight the neglected CCC dimension are also consistent with data from the NFSE-STU that likewise exhibited a large effect within the same construct. The thrust of the theme entails teachers' developing a new sense of understanding about the nature of NGSS-associated content and/or practices or, in other words, the development of new pPCK; this point will be elaborated more extensively in chapter 5.

Sentiments expressed by at least 24 of the 29 teachers interviewed supported the development of the theme. Participants made either general proclamations to the effectiveness of the intervention's ability to be useful in developing understanding, or more clearly articulated discrete changes in their understanding. Representative of the more generalized statements, Christine, a high school chemistry teacher, succinctly stated, "I think it's a really great tool because it made me aware of a lot of aspects of a really good quality NGSS assessment that . . . I didn't know what I didn't know!" An elementary school teacher, Craig, similarly provided a more general response as to the effect of the intervention. When asked if using 3D-PAST helped to clarify his understanding of the NGSS three-dimensionality, Craig stated, "I wish we had [3D-

PAST] much earlier though, because we would build these units and then if we could now get this, this makes more sense, this is a little more clear cut compared to what we were trying to do before."

Other teachers made statements that began to hone-in on mechanisms through which the intervention enhanced their understanding of NGSS dimensions and pedagogy. For example, Gerald, a high school physics teacher, described how going through the professional development activity helped him to more clearly envision what three-dimensional assessments were supposed to look like, and how he could use the tool as a "skeleton" on which he could hang different parts of an assessment as he created them. Sandra, a high school chemistry teacher, said using the 3D-PAST helped her to "cross the finish line," that is to say, it helped her to ensure the assessment tasks she created addressed each of the NGSS dimensions and were constructed in a way consistent with best practices. She stated,

> It wasn't that the assessment screening tool, in some ways, didn't present anything new to us, but it did, in some ways, reemphasize and illustrate and point out where we, kind of, weren't crossing the finish line all the way and so that was . . . beneficial for us and for many of the people here at the school.

Arthur, a high school teacher, communicated a similar thought but drew specific attention to the CCC domain. He highlighted how direct instruction and assessment of the CCC dimension is historically neglected and how using a tool like the 3D-PAST served as a mechanism of accountability. He stated,

> I thought it was a really useful exercise because, again, I think it highlights the difficulty in assessment . . . You know, because again, that's the one that's not really something that we've done explicitly as science teachers over the years. Obviously, we've done content, we've done SEPs, but we've never [explicitly taught CCCs] . . . so just to go to some of the sessions and like, "Okay, yeah, we're going to have a checklist where on this question we're saying, 'okay, here's some data, what are some patterns you notice?'"

Arthur went on to describe how, in addition to highlighting the neglected CCC dimension, participating in the intervention led to a further realization about the philosophical rationale of including each of the three dimensions in assessments. He said, "Previously, we had discussed patterns, and I think also Paul made a comment at some point which was, 'Well, if you truly believe [the NGSS] is thirds, thirds, thirds, then shouldn't your question to end scoring be third, third, third?'"

The theme was further corroborated by educators like Stephen, who was a high school teacher responsible for providing instruction in a number of subjects, including Advanced Placement courses as well as lower level science courses utilizing the NGSS framework. Similar to Arthur's statement, Stephen described how the CCC dimension was rarely addressed. At the same time, he also indicated how using the 3D-PAST shed light on his and his colleagues' more general struggles to develop assessments true to the three-dimensional framework. He stated,

> Rarely do we focus on crosscutting concepts across disciplines. Certainly. So, in that sense, a real three-dimensional assessment is just not a thing yet that we're

> really looking at. So that's definitely, definitely helped to kind of crystallize that, I know, for a lot of folks who maybe hadn't done that work up to now and kind of, kind of seen what [three-dimensional assessment] looks like.

A middle school science teacher, Diana, described how the intervention motivated substantive dialogue about the challenges she and her colleagues experienced trying to address the CCC dimension:

> And then we talked tons about how you incorporate CCCs into assessments without just over-assessing everything because it's hard enough to try to bring everything in for the NGSS, we find. So, what we do with CCCs is just try to make sure that our language and the prompts reflect the CCCs.

Louise, a high school biology teacher, similarly described how participating in the intervention led to new realizations about NGSS-aligned pedagogy. When asked how she perceived the intervention to have affected her understanding of NGSS approaches to instruction, she referenced changes in intentionality towards using scientific phenomena, scientific questioning, and crosscutting concepts, "One of the big things was being more intentional about 'Aha!' moments and being more intentional about the length of time you question, and the length of time you spend identifying those crosscutting patterns."

One of the most common clarifying events cited by teachers was the need to refer to other NGSS literature—specifically NGSS evidence statements—in order to better understand the PEs as they were working through the points of the screening tool; 19 of the 29 teachers interviewed mentioned how they benefited from using the evidence statements in conjunction with 3D-PAST. Jennifer, a middle school science teacher, was especially articulate in describing this effect when she stated,

> I realized [the 3D-PAST] was actually forcing me to actually look at those evidence statements more. Because in going through that checklist I was realizing that maybe, well, what is the skill and how is that described in the NGSS? What is the DCI? What is [the] crosscutting concept and how was it described in the NGSS? But that was actually, in a way, [what] forced me to look at the NGSS more fully, and in whatever standard that I was looking at, I actually found that I understood what the standard was getting at better by using this to evaluate my assessment.

The effect was also articulated well by Sean, an elementary teacher, who described how during the intervention he used the evidence statements in conjunction with the 3D-PAST to more clearly define assessment boundaries:

> It really gave me some clarity on the boundaries and sort of a lot deeper understanding. Because on first read, sometimes the disciplinary core ideas can read maybe like they're dealing with one thing and then if you get into the evidence statement and really dig deeper, it's a bit different in my experience than what my initial thoughts would be.

Eliciting Changes in Professional Practice

Another theme to emerge from the interviews was that participating in the PD activity elicited later changes in the teachers' professional practice. Leading to the development of this theme were statements made by 23 of 29 interviewed teachers suggesting that, subsequent to the intervention, they increased their use of scientific modeling (8 participants) or their use of scientific phenomena to engage students (5 participants), or they increased both (10 participants). These statements lend validity to findings from the NFSE-STU data that similarly showed large effect size in these areas; For example, the response items "When planning and teaching educators have students develop and refine conceptual models to express their understanding of scientific phenomena" and "When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration" both were found to have large effect sizes (see Figure 3).

Interviews also revealed changes that are not strictly associated with the framework of the NGSS, but rather with what might be considered general best practices in instruction. For example, participants mentioned increased use of graphic organizers (9 participants) and authentic data in assessments (8 participants), use of common NGSS-aligned language (9 participants), and thoughtful incorporation of backwards-design principles (22 participants). Indeed, only one participant could not attest to some subsequent change in at least one of these practices. A hurried review would make this component of the theme seem at odds with the NFSE-STU data that showed the intervention's effect to be smallest for the BP domain (refer to Table 13). However, this is reconciled by two important issues: Firstly, the aforementioned practices were not explicitly referenced by any response items on the NFSE-STU and, secondly, the terms used in the survey instrument were designed to obtain a measure understanding rather than of active practice. Therefore, rather than contradicting findings from the NFSE-

STU data, this theme expands upon the findings by providing insight about these other areas that the quantitative instrument, itself, was unable to measure.

Representative of the sort of statements underlying the development of this theme are those like that of Kenneth, who teaches science at the sixth-grade level. Kenneth stated that after using the 3D-PAST he began incorporating more SEPs into assessments because such activities were necessitated by changes in his assessment format. He said,

> If I'm going to give them a set of data to analyze and interpret so they can give me a claim, I need to do it in the classroom as well. So not only do you need to be able to identify a chemical reaction, but also measure temperature or measure whatever so that they are able to use that data. So, that has led us in the classroom, and it's helped me have that conversation with my teaching partner as well. That it's not only presenting the content, but also [students] need to have the skills to be able to answer these questions that are [aligned with] the screening tool.

Kenneth was particularly passionate about this backwards-design thinking and continued to describe how the intervention led to a "big shift" in his understanding of the NGSS's three-dimensional nature and subsequently forced him to reevaluate his teaching practices. In particular, he recognized that if the assessments he used were written to address the three dimensions of the NGSS, then he would need to appropriate adequate instructional time to those dimensions. He stated,

> Yeah. That's where the big shift was. . . . not only assessing the disciplinary core ideas. Right? That kind of clicked for me in terms of three-dimensional teaching in general. Right? As like, "Oh! These are equal weight, right?" The process is just as important. You need a tool for assessing that. Your assessment needs to be designed with that in mind and students should have opportunities to understand and do that before they get to the assessment.

Jennifer also made statements of this backwards-design variety. She described how she felt she was creating more effectively structured units with well-aligned formative assessments, because using the screening tool helped her to have more clear end goals when planning. She said,

> I finally have learned, okay, you've got to pull that [3D-PAST] out at the very beginning of unit planning, and that's just not the way I've ever worked in 20 years. I have never written my test first. I know I'm supposed to. I always know where I want [students] to go. Do you know what I mean? If I'm being honest, I've never fully written a brand-new assessment for a brand-new unit first. And it is really helpful to do that because then of course it makes it a lot easier to write the formative, and then everything else falls into place.

An increased use of graphic organizers in assessments was also commonly cited. For example, Jill, a middle school teacher, expressed how even though she was not yet wholly confident with the quality of her assessments, she had nonetheless begun incorporating graphic organizers (point 7 on the 3D-PAST) to help students better articulate their thoughts. About these changes she stated, while holding a 3D-PAST for the researcher to see,

> I think [graphic organizers] are really helping the students, but I'm not sure they're true assessments. Even with this, maybe I haven't used this enough to know. I'm not sure they're valid. That's my biggest issue . . . when is it a valid assessment? [The 3D-PAST] is helpful.

Diana also made statements that indicated this effect to be more widespread. Though she felt that her own use of graphic organizers hadn't increased in any meaningful way, she explained how she had observed a significant increase in the practice among colleagues who had also participated in the intervention with her. She made the claim,

> My issue with them is I would want to modify them quite a bit because I really struggle with anything, with just taking something that somebody else has made and using it. So, I just haven't taken the time to modify it in a way that I would like. But I know that the cause and effect one, a lot of people are using. Yeah, they're using them.

She then paused for a moment before reemphasizing the extent to which graphic organizers had been adopted, suggesting how some of her colleagues had been so enthusiastic about the graphic organizers that they had converted them into posters to use as teaching aids. She chuckled, and then said exuberantly, "They're all over people's walls and stuff!"

Facilitating Collaborative Practice.

The third theme to arise from the interviews was that teachers came to understand how using the 3D-PAST facilitated professional collaboration by way of focusing efforts and mitigating interpersonal conflicts. The observation that using the screening tool could serve such a function was outside the scope of the NFSE-STU, so it could not be triangulated with the quantitative data set. The idea that the tool might aid collaborative practice was also not considered in the initial formation of the semi-structured interview guide, so comments in this vein were usually put forth by teachers without explicit prompting. It is significant, then, that over a third (11 of 29) of the interviewed teachers made comments to this effect, and even more significant considering that it is primarily relevant in the larger school settings where professional learning communities (PLC) are utilized. In other words, in the smaller schools where only one teacher is tasked with teaching a subject there is far less expectation to work collaboratively on assessment design, so the effect of the 3D-PAST to either focus efforts on collaboration or mitigate interpersonal conflicts would not have occasion to be experienced by teachers.

Contributing to this theme were statements similar to those made by Ruby who is an elementary school teacher at one of the larger schools. She described how when working in PLCs of up to a dozen colleagues, conversation readily digressed from an agreed upon task, slowed completion, and caused frustration due to feelings of "wasted time." She explained how the 3D-PAST helped alleviate the problem by refocusing conversations in those scenarios:

> Every meeting we're in there's, depending on how many [educational assistants] there are in a classroom, you know, there could be 12 people in a room, right? If not more. So, you need a protocol so that everybody stays on the same page and doesn't go off into . . . "well, what I think should happen . . ." No, it's not what *you* think should happen, you use *the tool sheet*. Talk about the assessment. That's what we're doing.

Christine, who like Ruby worked in one of the larger schools in the study and regularly cooperated with colleagues as part of a formal PLC structure, described how using the screening tool expedited assessment design through focusing conversation. Christine stated,

> I love having a good list that just helps me systematically evaluate something. . . . I think it's really going to help the conversations with peers to say, "We're good on all these except for number eight and nine, that's where we should probably have another look."

and she further iterated, "So I think it can save us a lot of spinning wheels during meeting times."

The use of the screening tool was also suggested as a method for mitigating interpersonal conflicts during professional collaboration. Teachers expressed appreciation for the tool's ability to serve as an impartial and impersonal medium by which meaningful and focused critique of colleagues' work could be made. When asked to articulate more thoroughly on the idea, Roy, a middle school teacher, presented an example of how rather than engaging in a back-and-forth with colleagues over whose preferences were more valid, conversations could be redirected towards the impersonal 3D-PAST. Pointing at the 3D-PAST, Roy simulated a redirection of this nature:

> Yeah, like, "Hey we've agreed this is what we're measuring all of our assessments against. And, you know, if you have any reason to argue whether we should be doing this or not, that's separate from the assessment. We're just saying, 'Does the assessment actually do this?'" [as he pointed again to the 3D-PAST].

Roy then described the screening tool as a proxy for an expert who can intervene in disputes:

> And by Paul [Andersen] providing this we already, kind of, have a third point of an expert saying, "Here's what the NGSS is suggesting." And we give some respect to someone that's kind of given their life and full time to it.

The unique context of international schools made this function of the screening tool especially salient for Jennifer who, like Roy, works as one member of a formalized PLC. She articulated the importance of protocol in international schools where an inordinate number of close relationships between faculty members are found. She stated,

> The first thing that I like the most about it is that it gives a very impersonal tool for PLCs to look at each other's work. I think that's really important because everyone is—especially in international schools—we are friends with each other. We see each other at the pool. You need really good working relationships here with your team, particularly your PLC. So if somebody wrote something, and I'm not thinking of [my] particular PLC, but others that I've been a part of, if someone wrote something and they kick it off with, "Well I spent three hours writing this

assessment, what do you think?" it kind of tees you up to be like, "Oh, I better be really careful!" The screening tool takes a lot of that out of it, and so hopefully everybody's using it.

Influence of the Intervention Moderator

The final theme to emerge was the strong influence of Andersen's relationship with teachers on the way they perceived the value of the 3D-PAST intervention; those teachers who felt a personal affinity for Andersen tended to express positive experiences with the intervention, whereas teachers with whom Andersen had less rapport tended towards more moderate—or in a few cases, negative—evaluations of the intervention's impact. Also, for those teachers who had previously engaged with Andersen, there was sometimes a conflation of the learning experiences, meaning teachers were not able to isolate the learning that occurred as a result of the intervention from learnings that occurred as a result of previous Andersen-led PD experiences. As with the prior theme, the influence of the intervention moderator is a theme that cannot be triangulated with the NFSE-STU data as the survey instrument was not designed to capture a phenomenon of this kind. Accordingly, this theme likewise further expands upon findings from the quantitative data and, itself, warrants special consideration that will be addressed in chapter 5 with more detail.

There was a spectrum of this moderator influence displayed across all 29 interview participants. At one end of the spectrum were teachers who readily expressed their adoration for Andersen, with one even starting her interview with the proclamation, "I want to start, I love Paul, I adore him. I think he's so great in so many ways and I've learned a ton from him." Another teacher, when asked to expound on the way 3D-PAST impacted her teaching team, interjected the praise, "I think it started with Paul. Love Paul. He's coming back again. It's going to be great." In so doing, she declared a personal connection with Andersen that was predetermining the value of subsequent PD experiences led by him. Other statements such as, "Paul's great, really knowledgeable, has really driven science, and sort of inspired a lot of teachers to teach science in a really thoughtful way here," and, "Yeah, it was great. A wonderful experience and Paul Andersen was fantastic," and yet another, "So great for Paul that you're doing this" (as if the study was intended to be a marketing tool for Andersen) and a variety of other similar sentiments were commonplace. However, one middle school teacher, Patrick, was especially articulate in describing the personal impact Andersen had on him:

> To be honest with you, in my 17 years of teaching, and including my years of schooling, I've never had a mentor before and Paul is that, first. He's that guy I look at and I go like, "That. That guy's modeling what I want to be doing in my classroom."

Another example of this effect came from Emily, an elementary teacher. She described being influenced as much by Andersen's inspirational qualities as by the particular content of the PD sessions. She stated,

> I'm just going to start by saying his personality, his gregariousness just really gets people excited. When he comes to speak, if he does a keynote, of course he's doing a lot of the speaking, but when he's in a smaller group setting with just teachers, I love that he's like, "I don't know, what do you think?" Kind of like you

> could see how he would really keep his audience engaged by just turning that question right around. And I think the teachers appreciate that because he's very upfront. I don't know much about NGSS and yet somehow he slides his professional excellence into us, by making us kind of feel like, "Oh yeah! I've got that!" I like that about professional development, that he's reaching in and kind of helping us remember what we already know and making us feel successful and not afraid of something that our school kind of put on the back burner.

In each of these cases, the participants were very positive about the 3D-PAST intervention, even when they were challenged at times to communicate specific ways that it had actually benefited their understanding of the NGSS, or ways that it had elicited changes in their teaching practice.

Conversely, the relatively small number of teachers who were not especially positive, or were even dismissive, about their experience with the intervention were less enthusiastic about Andersen in general. For example, when asked if she'd continued to use 3D-PAST of her own accord following the intervention, Deborah, a high school teacher, bluntly proclaimed, "Paul kind of left a really bad taste in my mouth. Maybe if someone else had introduced it, I might have been more inclined to use it. But yeah, I wasn't terribly thrilled with him." Later in the interview, Deborah then made the unequivocal statement, "Well, but to be perfectly honest with you, I don't actually really like Paul Andersen very much. So, that might also color my perception." Joan, a middle school science teacher, began a description of her experience with the 3D-PAST activity stating, "It was fine." However, after a short pause she shifted her reply, "It was difficult. It was difficult, yeah. Because if you're going through the entire thing, yeah, I just, it was difficult." She continued to explain details of how she progressed through the intervention and ended her account by disclosing how her feelings about Andersen had changed recently, "Paul's all about Paul, Wyatt. I used to be a big Paul fan. I'm not so sure I'm a big Paul fan anymore."

In another case, William, a high school teacher, declared that he did not see any value in using the 3D-PAST; he also explained that his prior experiences with Andersen had left him generally unimpressed. These sentiments seemed compounded by his feelings toward administrators who were unresponsive to feedback about PD activities. William expressed,

> Personally, I didn't see any need for Paul Andersen to come back and they were like, "Paul Andersen's coming back!" [mockingly] And the principals are never even having a discussion with the science team, asking what our viewpoints are, they're not following up with how our experience has been.

There was also evidence of teachers conflating prior experiences with Andersen with learnings from the intervention that is the focus of this study. All but five of the interviewed teachers had participated in prior PD sessions led by Andersen because their schools had hired him to provide NGSS training sessions over multiple years. The prior PD sessions included topics like curriculum sequencing, using scientific phenomena to engage students, and effective practices in scientific modelling. The conflation of things learned from prior PD with things learned from the intervention was clearly evident when teachers were asked very direct questions about the use of the 3D-PAST, but their replies included references to previous PD activities with Andersen. Sometimes shifts between these learnings occurred mid-response. Take for example the following exchange in which an elementary school teacher, Craig, described specific practices of modeling instruction (a topic of a different PD session) after being asked unambiguously about points on the 3D-PAST:

> Researcher: You mentioned that these, seven through eleven [referring to points on the screening tool], which are related to more, a little more, closely to practices in the classroom.
>
> Craig: Correct.
>
> Researcher: That's actually shifted a little bit . . .
>
> Craig: [talks over] No that's . . .
>
> Researcher: . . . what you're doing in the classroom, because of [3D-PAST]?
>
> Craig: Absolutely. Well it's helped to guide. And one of the main messages, or things that we walked away with, working with Paul, was that the main things we need to teach grade five, we need to make sure that they have a clear understanding of how to model properly, how to have a very specific zoom in to show more at the molecular level what's actually happening.

Craig then further elaborated on various aspects of scientific modeling that were presented in one of the previous Andersen-led PD activities, rather than effects which could be more readily and reasonably construed as resulting directly from the 3D-PAST intervention, itself.

Summary of the Findings

Data from participant responses to the NFSE-STU indicate that teachers perceived the study's intervention to have had a positive effect on understanding in each of the five constructs measured by the instrument: science and engineering practices, disciplinary core ideas, crosscutting concepts, integration of the three dimensions of the NGSS, and best practices in science instruction. Further exploration of the intervention's effect, by way of interviews with a subsample of teachers, saw the development of four themes—the first of which supports the NFSE-STU data, while the remaining three themes reflect phenomenon beyond the scope of the NFSE-STU and so can be considered to expand upon the findings from the quantitative data.

The first theme was that teachers perceived the intervention to enhance their understanding of the NGSS's three dimensions, and that the intervention highlighted the neglected CCC domain. This theme is consistent with findings from the NFSE-STU, particularly the instrument's I3D and CCC constructs that similarly showed large effect sizes.

A second theme emerged from the interviews indicating that teachers have increased the extent to which they incorporated scientific phenomena, modeling instruction, and graphic organizers into instruction. Interviews also provided evidence for teachers' increased use of backwards-design thinking as they planned instruction.

The third theme to emerge was the perceived value of the 3D-PAST to improve collaborative practices. Given the unique contexts of international school teachers, this theme presents as particularly promising and relevant.

Finally, the fourth theme revealed the influence of the intervention's moderator on the way the intervention, itself, was perceived to have value. This fourth theme carries weight for interpretation and implications of the findings. The significance of these findings and their relation to the extant literature will be further elaborated in the next chapter.

CHAPTER 5

DISCUSSION

The purpose of this study was to explore how a professional development activity in which teachers used a tool to evaluate the alignment of student assessments with the NGSS framework either enhanced or challenged their understanding of the standards themselves. The PD activity was led by a science consultant at 15 type-A American international schools in 13 countries. This final chapter provides a discussion of the findings in relation to the theoretical frameworks guiding the study, the limitations of the study, the implications of the study in relation to practice and research, and closing thoughts.

Complementarity of the Quantitative and Qualitative Data

This study utilized the *New Framework of Science Education Survey of Teacher Understanding* (NFSE-STU) developed by Nollmeyer and Bangert (2017) as an instrument to collect quantitative data from a large sample of teachers who participated in the intervention; it also used semi-structured interviews, with a subsample of the NFSE-STU respondents, to develop a more nuanced understanding of the way teachers were impacted by the intervention. Taken independently of each other the quantitative or qualitative data collected provide partial understanding of the professional development activity's effect on teachers, but taken together the evidence from both methods provide deeper insight than either could have developed in isolation.

In particular, data produced from the NFSE-STU is limited in that the design of the instrument focused solely on teacher understanding. The way teachers subsequently transform understandings into new professional practices are not within the scope of the NFSE-STU, but interviews were able to shed light on this aspect. Interviews also provided insight into contextual aspects that influenced the intervention's effect on teachers.

Results in Relation to the Extant Literature and Theories

This study was guided by the Refined Consensus Model of pedagogical content knowledge (RCM). Understanding measured by the NFSE-STU is a component of a teacher's larger reservoir of subject-specific content and pedagogical knowledge, and so aligns with the conceptualization of personal pedagogical content knowledge as envisioned in the RCM (Carlson & Daehler, 2019). Consequently, the statistically significant positive changes found in the analysis of the NFSE-STU data suggest that the professional development activity did, indeed, enhance teachers' pPCK, and these findings were supported by the interviews.

The NFSE-STU, however, is limited in scope in that it measures merely teachers' perceived understanding. The RCM, in contrast, makes a more holistic effort to distinguish realms of PCK. Lying outside of and apart from pPCK, the RCM describes collective pedagogical content knowledge as the vast body of knowledge extant in the field; such knowledge may be from various studies, literature, knowledge held by professional groups, or knowledge from other informed sources. By this definition, the NGSS sit firmly within the realm of cPCK; they are a set of standards that are the product of numerous studies and professional collaborations. According to the RCM, one way in which a teacher's pPCK grows is when knowledge from the cPCK domain is transferred into the pPCK domain. Consequently, in so much as the intervention was found to grow teachers' understanding of the NGSS, it may also be considered to be a mechanism through which there was an effective knowledge transfer between the two domains of cPCK and pPCK.

This may not be the entire story, however. It may not have always been the case that the intervention actually grew pPCK in the sense that entirely new pPCK understandings were transferred from the cPCK domain. This was indicated by Sandra when she said,

> It wasn't that the assessment screening tool, in some ways, didn't present anything new to us, but it did, in some ways, reemphasize and illustrate and point out where we, kind of, weren't crossing the finish line all the way.

In fact, there a number of instances when statements made by teachers gave the impression that they felt they were more being reminded of things they already knew, rather than actually learning new things altogether. In these instances, the intervention may have served rather to *awaken* existing, but dormant, pPCK more so than to actually *transfer* cPCK—as new—into a teachers' existing bank of pPCK. As interesting as this distinction between transferring and awakening may be in a theoretical discussion, in the conception of the RCM it is a distinction that is, nonetheless, largely irrelevant because the RCM posits that pPCK only serves students when it is acted upon in the form of ePCK.

The concentrically-nested domains of the RCM (see Figure 1) position enacted pedagogical content (i.e., the things teachers actually do with students and in their other instruction-related activities) at the model's center. Such positioning is appropriate for multiple reasons. First, it should be self-evident that some knowledge must first be activated before it can be acted upon; this activation may be either new pPCK recently gained via transfer, or old pPCK recently awakened. Second, ePCK is arguably the heart of teaching itself. If the purpose of formal education is chiefly to build students' understanding, then what purpose does the professional educator serve if not to interact and to take action in ways that elicit new understanding in students? By the same token, what good is developing some aspect of a teacher's pPCK if the teacher cannot, or will not, ever act upon it in a way that impacts student learning? While the study set out to explore a possible mechanism by which teachers experience growth of their pPCK, the interviews revealed that the intervention not only served this purpose, but it apparently also was effective in eliciting changes in ePCK as well. That is to say, it elicited changes that really matter to students; it elicited changes in what teachers actually do.

The RCM goes further still. More than just distinguishing discrete realms of pedagogical content knowledge, it recognizes that the development of pPCK (and so, in turn, ePCK) always occurs within the context of a learning environment. In the case of this study, then, it must not be forgotten that any changes brought about by the intervention were achieved within the unique context of the international school setting. The third and fourth themes to develop from the interviews especially speak to the primacy of context when considering the impacts and implications of the intervention.

The third theme was teachers' perceived value of the tool to mitigate interpersonal conflict and expedite meetings. While, surely, a tool that serves to mitigate conflict and speed along meetings would be deemed valuable by teachers in most any context, the unique context of the international school likely made this point markedly more salient for the teachers in this study. As detailed in chapter 2, though the professional training and qualifications of the typical international school teacher are similar to their domestic counterparts, the conditions in which they operate are quite unique. Navigating the web of personal-professional relationships is different in international schools than in domestic schools as high amongst the differences is the especially tight-knit communities in which these teachers live and work. Recall Jennifer, who reflected on using the screening tool:

> I think [using an impersonal tool is] really important because everyone is — especially in international schools — we are friends with each other. We see each other at the pool. You need really good working relationships here with your team, particularly your PLC.

She and others who expressed similar sentiments affirmed previous research suggesting that the international school teacher faces unique stressors relating to the tight knit communities of these types of schools (Hayden, 2006; Zilber, 2005).

The fourth theme to develop similarly speaks to the importance of context in not merely the acquisition of pPCK but to teachers' willingness to activate it (i.e., to make it become ePCK by transforming it into meaningful learning experiences for students or into other professional practices). One component of this theme found teachers frequently challenged to differentiate between the things they had learned during previous instances of PD and the understandings they had gleaned directly from the intervention. Teachers, tended to situate the new pPCK within the old in such a way that the boundaries were blurry; it was unclear where one set of PD started and the other one ended. The primacy of context was further illuminated by the influence exerted by Andersen himself. The teachers' relationships with Andersen, sometimes developed over multiple years, proved a strong predictor of their willingness to acquire new pPCK through the intervention or to make changes to their professional practice as a result of it.

Limitations

The contextual influences of the third and fourth themes also hint at what may be the most significant limitations of this study. Related to the third theme, the international school teachers who participated in this study operate in unique social and employment contexts, which may amplify the relational aspect of the intervention impact. Participants spoke to the particular benefit of using the tool to navigate interpersonal relationships, a point that stood out strongly because of the atypically tight work/social relationships that exist in the schools of this type. Also, as noted in chapter 2, these type-A schools are unique in that they are often better funded than their U.S. private and public-school counterparts, and these schools typically have parent communities that expect the highest quality teachers. This combination of resources and expectations allows type-A international schools to develop systematic and well-resourced professional development programs for their teachers; it is not a typical domestic school that would have the impetus or the resources to bring an individual halfway around the world to provide training, and then also give entire groups of teachers time out of their classrooms to participate in the training. Consequently, the findings of this study are limited in that their scope is within this unique international school aggregation of work and social context, and the findings are more difficult to apply to other populations of educators who do not operate in similar environments.

The fourth theme also alluded to a context-related limitation of the study. The study found that Andersen's character and the interpersonal relationships he established significantly influenced the way teachers perceived the value of the intervention. The way this limits the study hinges upon whether the moderator is considered a part of the intervention itself or if he is not. If the goal is to examine the value of the intervention in a way separate from Andersen himself (i.e., to merely examine the process of using the three-dimensional screening tool to review assessments), then the development of this theme could be interpreted as indicative of a large experimenter effect, which would necessitate the use of caution when interpreting and applying nearly all of the findings of the study. However, this issue can be mitigated substantially if the moderator is considered to be an integral component of the whole intervention. These limitations should be kept in mind while considering the implications of the study.

Implications

Implications for Practice

It may be presumed that the end-goal of any regimen of teacher professional development is to create changes in teachers which then have a net positive effect on student learning. Consequently, the findings from this study strongly support the recommendation that international school leaders seeking to maximize the impact of science teacher professional development to (a) consider PD activities that train teachers in the consistent use of aids—such as 3D-PAST—while developing student assessments that hold potential to elicit changes in pPCK and ePCK, (b) seek to identify, implement, and evaluate the impact of PD as it pertains to enacted pedagogical content knowledge with particular regard to the alignment of NGSS assessments and (c) give careful consideration to the contextual components of PD that strongly influence the extent and nature of the pedagogical content knowledge that is assimilated, transferred and acted upon by teachers.

With regard to the recommendation of using of aids, the analysis conducted in this study suggests that teacher use of the three-dimensional performance assessment screening tool achieves the goal of improving student learning by two means. Firstly, the use of 3D-PAST was shown to enhance teacher understanding of the NGSS, particularly with regards to the use of science and engineering practices and the way in which the three-dimensionality of the NGSS is understood and addressed with students. As a result, teachers are better prepared to guide students using the pedagogical approaches the NGSS considers most effective at developing the science literacy necessary for student success in the world's increasingly science and technology-driven economies. Secondly, the use of a simple tool like the 3D-PAST was shown as an effective means of creating change in professional practice that had a direct impact on teacher's planning and interactions with students or, in other words, *enacted* pedagogical content knowledge. This is important because it is in the domain of ePCK that teachers bring about meaningful learning opportunities for students.

To the second point, this study especially demonstrated how teachers' intentional and methodical efforts to align assessments with the NGSS constructs further elicited changes in other practices that are even more closely connected to student learning experiences, such as increased uses of modelling instruction, scientific phenomenon, and graphic organizers. Even though the schools in this study, and those like them, are typically well resourced and may be able to be able to spend more on PD opportunities for their faculty than their domestic counterparts, given ePCK's direct relation to student learning, international school administrators should nonetheless be active in identifying and selectively implementing PD experiences that bring about changes in ePCK as efficiently as possible. This study revealed the changes in teacher ePCK through interviews, a method which in practice—especially in the larger schools—would not likely be an expedient way for school leaders to gauge the efficacy of PD as it pertains to effecting ePCK, nor would the NFSE-STU utilized in the study be effective in this regard as its scope of measurement is solely pPCK. However, the Science Instructional Practices Survey (Hayes, Lee, DiStefano, O'Conner, & Seitz, 2016) utilized in cycle two of this action research study is easy to distribute and uses language more consistent with measurement of the ePCK domain, so may be a worthy consideration as a tool for school leaders to use in evaluating the quality of NGSS-related professional development.

Finally, but importantly, school administrators need to be cognizant of the primacy of context in the way professional development activities are experienced by teachers and the extent to which pPCK is assimilated by teachers. While this is likely to be true in any setting, the findings of this study suggest context is especially influential in the international school environment where there are uniquely heightened stresses associated with working relationships and other working conditions, including initiative fatigue[6]. Particularly, this study found that the way in which the moderator of the intervention was able to develop relationships with teachers was an essential component of the interventions' ultimate success, not only in growing teachers' pPCK but in eliciting changes in ePCK. Consequently, in seeking to maximize the efficacy of PD for science teachers, international school leaders are wise to consider how they might encourage relationship-building between teachers and the leaders of professional development that encourages greater PCK assimilation. To this end, this study showed that the character qualities of the PD moderator, as well as opportunities to develop relationships over multiple PD engagements and extended periods of time, are important factors to consider.

Implications for Research

The implications of this study for educational research are that (a) the study contributes to the nascent literature utilizing the Refined Consensus Model of pedagogical content knowledge to situate studies of teacher pedagogical content knowledge, (b) the study fills a void in the literature concerning professional development practices in international school settings, and (c) raises questions that may serve to spawn additional cycles or areas of action research.

With regard to the first implication, though pedagogical content knowledge has been used as a construct for educational researchers to study teacher knowledge and

[6] Initiative fatigue was discussed in chapter two as a stressor identified in the second the action research study. It was a key factor contributing to the ultimate selection of the 3D-PAST professional development activity as the intervention to be explored for the dissertation.

practices since it was originally presented by Shulman (1986), the RCM itself is a more recent formulation that seeks to reconcile challenges presented by a number of models of PCK that have been used in the interim (Carlson & Daehler, 2019). As new as the model is at the time of this writing, there is not a wealth of published literature demonstrating application of the RCM; a cursory review of the literature revealed a handful of studies to date. Consequently, the present study contributes to this new body of literature by providing an example of how the framework can be used to situate teacher development of PCK; in this instance, the RCM proved useful in making important distinctions between the effects of the intervention on teachers' understanding of the NGSS as well as the effects of the intervention on teachers' practices. It also aided in illuminating the importance of context and situating it as an influential lens that colors the way domains of PCK are developed by teachers.

This closing chapter has emphasized the importance of context in interpreting the findings of this study and, as yet, it has done so to encourage consideration of the way in which the person leading the PD has an effect on teachers' perceptions. However, there is another facet of context, that must not be forgotten: This study examined a particular type of PD activity, focused on a specific set of science standards, led by a single individual, at 15 international schools in 13 countries, over a period of just 10 months. That such a study could even take place indicates a sort of homogeneity across geographically disparate institutions that is quite remarkable, and reinforces the understanding of type-A international schools as a group.

The second implication concerns the field of research related to the context of international schools. Research in the field is increasing at a steady pace, but it still significantly lags behind research in other fields (Hayden & Thompson, 2008; 2013). This is notable because international schools serve a significant, and growing, number of students. Thus, the study adds to the relatively less abundant, but increasingly needed, body of literature related to the international school context. Further, the study's investigation into a particular science-teacher and NGSS focused PD activity in these schools may be one of the first of its kind. Given these unique characteristics, the study fills a particular void within what is already a relatively sparse body of literature.

Finally, findings from this study evoke some important issues that may be starting points for future action research. The analysis conducted in this study revealed the importance, and validated the position, of context within the Refined Consensus Model. Chiefly, the nature of international schools and the relationship between the intervention moderator and the teachers he worked with were shown to be important contextual components impacting teachers' assimilation of PCK; new cycles of action research could seek to investigate, more specifically, the aspects of moderator-teacher relationships that are most consequential to increasing PD efficacy or, conversely, attempt to isolate more completely the teachers' perception of the 3D-PAST to elicit important changes in teacher pPCK and ePCK from the moderator influence. Additionally, as the tool is designed for the discipline of science—and specifically the NGSS—there is potential to investigate if similar tools could be designed and implemented for other teachers, in other disciplines, to produce similar effects.

Concluding Thoughts

Educational research is difficult. Each study is an exploration of a unique, multifaceted, and dynamic context. When quantitative instruments are used to evaluate perceptions or understanding, researchers ultimately accept margins of error that are intrinsic to an attempt to quantify things that are inherently not quantifiable. When qualitative measures are used, educational researchers do their best to explore phenomena that frustrate categorization and defy clear explanations. Regardless of whether quantitative or qualitative, or both, methods are used to build our understanding of what, how, and why humans learn what we learn, our understanding is never quite complete. A new context, a new person, an unexplored avenue, or a different technology, brings new insight that may expand or deepen what we know, but in doing so, often also twists or complicates it. Rather than making things seem cleaner, time and again, educational research seems to make things more messy.

In an age when big data and personalized, but impersonal and emotionless, algorithms seem to infiltrate so many aspects of our lives, there seems to similarly be a growing trend towards the trumpeting of personalized but impersonal programs in education, as if just selecting the right set of standards, model for grading and reporting, ways to respond and intervene with students, professional collaboration structure, or professional development regimen will finally tidy things up. But the messiness exists because education is a human endeavor and humans are messy beings. We, each, have our own emotions, memories, compulsions, proclivities, aversions, mannerisms, senses of humor, and values that aggregate into a unique, but constantly changing, cognitive imprint that leaves its mark on the interpretation of every stimulus we receive.

In these terms, this study set out to explore how a personalized but impersonal tool—personalized because it targets a specific practice of a specific kind of teacher, and impersonal because it is little more than a checklist of things to include on an assessment—might change the cognitive imprint of the teachers who were exposed to it. While analysis of the data suggests that the 3D-PAST was effective in bringing about positive changes in the way teachers understood the NGSS, the ways they conducted their practices with students, and even the ways they engaged in professional collaboration, it also highlighted the inescapable messiness of humanity with which education is entangled; it exposed the significant influence of human relationship and human interaction in the way the impersonal tool was received. In the deep and uniquely human complexities of the educational endeavor, there surely is a need for systems, protocols, tools and programs to guide what we do. Yet as we—educators and administrators—seek to change the cognitive imprints of our students and faculty, we are wise to never lose sight of the mark that our own imprint leaves on others. For in the way that we interact with our students and colleagues, and to the extent that we build rapport with them, so also do we increase or diminish the value of the things we ask of them, and the prospect that any related changes will be lasting.

REFERENCES

Abell, S. K. (2007). Research on science teacher knowledge. In S. K. Abell, & N. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 1105–1149). Mahwah, NJ: Lawrence Erlbaum Associates.

About Paul Andersen. (n.d.) Retrieved March 05, 2019, from http://www.bozemanscience.com/about

Allen, J. M., & Nimon, K. (2007). Retrospective Pretest: A Practical Technique for Professional Development Evaluation. *Journal of Industrial Teacher Education*, *44*(3), 27-42.

American Association for the Advancement of Science. (1990). *Science for all Americans*. New York, NY: Oxford University Press.

American Association for the Advancement of Science. (1994). *Benchmarks for science literacy*. New York, NY: Oxford University Press.

Andersen, P., & Brosnick, L. (n.d.). *3-Dimensional Performance Assessment Screening Tool* [Pdf]. Paul Andersen. Retrieved January 9, 2018 from https://thewonderofscience.com/s/3D-Screening-Tool-8lxx.pdf

Anderson, M., & Freebody, K. (2012). Developing communities of praxis: Bridging the theoretical practice divide in teacher education. *McGill Journal of Education*, 47(3), 359- 377.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133-180. https://doi.org/10.3102/0002831209345157

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59(5), 389–407. https://doi.org/10.1177/0022487108324554

Bentley, M. L., Ebert, E. S. II, & Ebert, C. (2007). *Teaching constructivist science: Nurturing natural investigations in the standards-based classroom*. Thousand Oaks, CA: Sage Publications.

Bhanji, F., Gottesman, R., de Grave, W., Steinert, Y., & Winer, L. R. (2012). The retrospective pre–post: A practical method to evaluate learning from an educational program. *Academic Emergency Medicine*, *19*(2), 189-194. https://doi.org/10.1111/j.1553-2712.2011.01270.x.

Bianchini, J. A., & Kelly, G. J. (2003). Challenges of standards-based reform: The example of California's science content standards and textbook adoption process. *Science Education*, *87*(3), 378-389. https://doi.org/10.1002/sce.10064

Brinkmann, S., & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Brummit, N., & Keeling, A. (2013). Charting the growth of international schools. *International Education and Schools: Moving Beyond the First 40 Years* (pp. 25-36). London: Bloomsbury Academic. https://doi.org/10.5040/9781472553034

Brunsell, E., Kneser, D. M., & Niemi, K. J. (2014). *Introducing teachers and administrators to the NGSS: A professional development facilitator's guide.* Arlington, VA: National Science Teachers Association.

Bunnell, T. (2014). *The changing landscape of international schooling: Implications for theory and practice*. New York, NY: Routledge.

Bybee, R. W. (2013). The next generation science standards and the life sciences. *Science and Children*, *50*(6), 7.

Bybee, R. W. (2014). NGSS and the next generation of science teachers. *Journal of Science Teacher Education*, 25(2), 211-221.

California Department of Education. (2015, October 15). NGSS for California Public Schools, K-12. Retrieved May 17, 2016, from http://www.cde.ca.gov/pd/ca/sc/ngssstandards.asp.

Cantrell, P. (2003). Traditional vs. retrospective pretests for measuring science teaching efficacy beliefs in preservice teachers. *School Science and Mathematics*, *103*(4), 177-185. https://doi.org/10.1111/j.1949-8594.2003.tb18116.x

Carlson, J. & Daehler, K.R. (2019). The Refined Consensus Model of pedagogical content knowledge in science education. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 77-92). Springer Nature. https//doi.org/10.1007/978-981-13-5898-2

Chan, K. K. H. & Hume, A. (2019). Towards a consensus model: literature review of how science teachers; pedagogical content knowledge is investigated in empirical studies. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (pp. 3-76). Singapore: Springer Nature. https//doi.org/10.1007/978-981-13-5898-2

Chang, R., & Little, T. D. (2018). Innovations for evaluation research: Multiform protocols, visual analog scaling, and the retrospective pretest–posttest design. *Evaluation & The Health Professions*, *41*(2), 246-269. https://doi.org/10.1177/0163278718759396

Clark, V. L. P., & Creswell, J. W. (2010). *Understanding research: A consumer's guide*. Upper Saddle River, NJ: Merrill/Pearson Educational.

Commission on Mathematics and Science Education (U.S.). (2009). *Opportunity equation: Transforming mathematics and science education for citizenship and the global economy.* Carnegie Corporation.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3-21.

Covay Minor, E., Desimone, L., Caines Lee, J., & Hochberg, E. D. (2016). Insights on how to shape teacher learning policy: the role of teacher content knowledge in explaining differential effects of professional development. *Education Policy analysis archives*, *24*(61), n61. https://doi.org/10.14507/epaa.24.2365

Creative Commons. (n.d.). *About The Licenses.* Retrieved February 03, 2019, from https://creativecommons.org/licenses/

Cresswell, J.W. (2013). Data analysis and representation. *Qualitative inquiry and research design: choosing among five approaches.* Los Angeles, CA: Sage Publications, 15-41.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Los Angeles, CA: Sage Publications.

DeBoer, G. (1991). *A history of ideas in science education: Implications for practice.* New York, NY: Teachers College Press.

Dedoose [Computer software]. (2019). Retrieved from: http://www.dedoose.com

DODEA, (n.d). *About DoDEA*. *U.S. Department of Defense Education Activity.* Retrieved October 10, 2018 from https://www.dodea.edu/aboutDoDEA/

DODEA. (2008). *Regulation 1342.13*. Retrieved from: https://www.dodea.edu/Offices/PolicyAndLegislation/upload/1342_13010.pdf

East Asia Regional Council of Overseas Schools. (n.d.). *EARCOS Teachers' Conference 2017*. Retrieved November 21, 2018, from http://www.earcos.org/etc2017/etc-credit.php

East Asia Regional Council of Overseas Schools. (2017). *15th Annual EARCOS Teachers' Conference 2017, Connecting Global Minds.* Retrieved from https://www.earcos.org/etc2017/ETC2017-program_small_res.pdf

Elliott, V. (2018). Thinking about the coding process in qualitative data analysis. *The Qualitative Report, 23*(11), 2850-2861. Retrieved January 01, 2020 from https://nsuworks.nova.edu/tqr/vol23/iss11/14

Enochs, L.G., & Riggs, I. M. (1990). Further development of an elementary science teaching efficacy beliefs instrument: A preservice elementary scale. *School Science and Mathematics, 90*, 694-706.

Fabian J. (2016). A Pedagogy for International Education. In M. Hayden, & J. Thompson (Eds.), *International Schools: Current Issues and Future Prospects.* (pp. 85-104). Oxford: Symposium Books Ltd.

Fertig, M. & James, C. (2016). The Leadership and Management of International School: very complex matters. In M. Hayden, & J. Thompson (Eds.), *International Schools: Current Issues and Future Prospects.* (pp. 105-128). Oxford: Symposium Books Ltd.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA: Sage Publications.

Frankfurt International School. (n.d.). *Tuition and Fees for Academic Year 2018-2019*. Retrieved January 06, 2019, from https://www.fis.edu/page.cfm?p=945

Garton, B. (2000). Recruitment of teachers for international education. In M. Hayden, & J. Thompson (Eds.), *International Schools and International Education: Improving Reaching, Management and Quality* (pp. 85-95). London: Kogan Page.

Gillies, W. D. (2001). American international schools: Poised for the twenty-first century. *Education*, *122*(2), 395-401.

Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social Problems*, *12*(4), 436-445.

Glaser, B. & Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York, NY: Aldine De Gruyter.

Haag, S., & Megowan, C. (2015). Next generation science standards: A national mixed-methods study on teacher readiness. *School Science and Mathematics*, *115*(8), 416-426. https://doi.org/10.1111/ssm.12145

Halicioglu, M. L. (2015). Challenges facing teachers new to working in schools overseas. *Journal of Research in International Education*, *14*(3), 242-257. https://doi.org/10.1177147524091561150

Harris, K., Sithole, A., & Kibirige, J. (2017). A needs assessment for the adoption of next generation science standards (NGSS) in k-12 education in the United States. *Journal of Education and Training Studies*, *5*(9), 54-62.

Hayden, M. (2006). *Introduction to international education: International schools and their communities*. London: Sage Publications.

Hayden, M. (2011). Transnational spaces of education: The growth of the international school sector. *Globalisation, Societies and Education*, *9*(2), 211-224.

Hayden, M., & Thompson, J. (1995). International schools and international education: A relationship reviewed. *Oxford Review of Education*, *21*(3), 327-345. https://doi.org/10.1080/14767724.2011.577203

Hayden, M., & Thompson, J. (2008). *International schools: Growth and influence*. Paris, France: United Nations Educational, Scientific and Cultural Organization.

Hayden, M., & Thompson, J. (2013). International Schools: Antecedents, current issues and metaphors for the future. In R. Pearce (Ed.), *International Education and Schools: Moving Beyond the First 40 Years* (pp. 3-24). London: Bloomsbury Academic. https://doi.org/10.5040/9781472553034

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*(2), 137-164. https://doi.org/10.1007/s10972-016-9488-5

Herr, K., & Anderson, G.L. (2005). The continuum of positionality in action research. In *The Action Research Dissertation: A Guide for Students and Faculty* (pp. 29-48). Thousand Oaks, CA: Sage Publications.

Heyward, M. (2002). From international to intercultural: Redefining the international school for a globalized world. *Journal of Research in International Education*, *1*(1), 9-32. https://doi.org/10.1177/147524090211002

Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher*, *31*(5), 3-15. https://doi.org/10.3102/0013189X031005003

Howard, G. S., Ralph, K. M., Gulanick, N. A., Maxwell, S. E., Nance, D. W., & Gerber, S. K. (1979). Internal invalidity in pretest-posttest self-report evaluations and a re-evaluation of retrospective pretests. *Applied Psychological Measurement*, *3*(1), 1-23. https://doi.org/10.1177/014662167900300101

Howell, P. B., Cook, C. M., Miller, N. C., Thompson, N. L., Faulkner, S. A., & Rintamaa, M. F. (2018). The Complexities of Middle Level Teacher Credentialing: Status Report and Future Directions. *RMLE Online*, *41*(4), 1-12. https://doi.org/10.1080/19404476.2018.1456840

Hrycak, J. (2015). Home and away: An inquiry into home-based and overseas teacher perceptions regarding international schools. *Journal of Research in International Education*, *14*(1), 29-43. https://doi.org/10.1177/1475240915573136

Hume, A., Cooper, R., & Borowski, A. (2019). *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science.* Singapore: Springer Nature. https//doi.org/10.1007/978-981-13-5898-2

International School Services. (n.d.). *About ISS*. Retrieved January 06, 2019, from https://www.iss.edu/who-we-are/about

James, C., & Sheppard, P. (2014). The governing of international schools: The implications of ownership and profit motive. *School Leadership & Management*, *34*(1), 2-20. https://doi.org/10.1080/13632434.2013.813457

Johanningmeier, E. V. (2010). A nation at risk and Sputnik: Compared and reconsidered. *American Educational History Journal*, *37*(1/2), 347.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112-133. https://doi.org/10.1177/1558689806298224

Joneitz, P.L. & Harris, D. (2012). World yearbook of education 1991: International schools and international education. Routledge. https://doi.org/10.4324/9780203080344

Kang, E. J., Donovan, C., & McCarthy, M. J. (2018). Exploring elementary teachers' pedagogical content knowledge and confidence in implementing the NGSS science and engineering practices. *Journal of Science Teacher Education*, *29*(1), 9-29. https://doi.org/10.1080/1046560X.2017.1415616

Klatt, J., & Taylor-Powell, E. (2005). Synthesis of literature relative to the retrospective pretest design. Panel presentation for 2005 Joint CES/AEA Conference, Toronto, ON.

Klieger, A., & Yakobovitch, A. (2011). Perception of science standards' effectiveness and their implementation by science teachers. *Journal of Science Education and Technology*, *20*(3), 286-299.

Kind, V. (2009). Pedagogical content knowledge in science education: perspectives and potential for progress. *Studies in Science Education*, *45*(2), 169-204. https://doi.org/10.1080/03057260903142285

Krajcik, J. (2015). Project-based science: Engaging students in three-dimensional learning. *The Science Teacher, 82*(1), 25.

Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the next generation science standards. *Journal of Science Teacher Education, 25*(2), 157–175. https://doi.org/10.1007/s10972-014-9383-2

Krepf, M., Plöger, W., Scholl, D., & Seifert, A. (2018). Pedagogical content knowledge of experts and novices—what knowledge do they activate when analyzing science lessons? *Journal of Research in Science Teaching*, *55*(1), 44-67. https://doi.org/10.1002/tea.21410

Lam, S. F., Cheng, R. W. Y., & Choy, H. C. (2010). School support and teacher motivation to implement project-based learning. *Learning and Instruction*, *20*(6), 487-497.

Lederman, N. G., & Gess-Newsome, J. (1992). Do subject matter knowledge, pedagogical knowledge, and pedagogical content knowledge constitute the ideal gas law of science teaching? *Journal of Science Teacher Education*, *3*(1), 16-20.

Little, T. D., Chang, R., Gorrall, B. K., Waggenspack, L., Fukuda, E., Allen, P. J., & Noam, G. G. (2019). The retrospective pretest-posttest design redux: On its validity as an alternative to traditional pretest-posttest measurement. *International Journal of Behavioral Development*, 44(2), 1-9. https://doi.org/10.1177/0165025419877973

Loveland, T. (2004). Technology Education Standards Implementation in Florida. *Journal of Technology Education*, 16(1), 40-54.

Mancuso, S. V., Roberts, L., & White, G. P. (2010). Teacher retention in international schools: The key role of school leadership. *Journal of Research in International Education*, *9*(3), 306-323. https://doi.org/10.1177/1475240910388928

Mannino, E. N. (1992). *Special activities of the office of overseas schools.* Washington, DC: Department of State, Office of Overseas Schools.

MacDonald, J. (2006). The international school industry: Examining international schools through an economic lens. *Journal of Research in International Education*, *5*(2), 191-213. https://doi.org/10.1177/1475240906065618

Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA: ASCD.

Miller, M., & Hinshaw, R. E. (2012). The retrospective pretest as a gauge of change. *Journal of Instructional Psychology*, *39*(3-4), 251-259.

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods*, *1*(2), 13-22. https://doi.org/10.1177/160940690200100202

Nagrath, C. (2011, August 28). What Makes a School International? Retrieved September 26, 2017, from https://www.tieonline.com/view_article.cfm?ArticleID=87

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. *The Elementary School Journal*, *84*(2), 113-130.

National Education Association (1918). *Cardinal principles of secondary education: A report of the commission on the reorganization of secondary education.* (U.S. Bureau of Education Bulletin No. 35). Washington, DC: U.S. Government Printing Office

National Research Council. (1996). *National science education standards*. Washington, DC: The National Academies Press. https://doi.org/10.17226/4962

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

National Society for the Study of Education. (1932). *A program for teaching science: Thirty-first yearbook of the NSSE*. Chicago, IL: University of Chicago Press.

National Society for the Study of Education. (1947). *Science education in American schools: Forty-sixth yearbook of the NSSE*. Chicago, IL: University of Chicago Press.

Near East South Asia Council of Overseas Schools. (n.d.). *Events: Unlocking the Power of the NGSS.* Retrieved March 04, 2019 from https://www.nesacenter.org/events/event-archives/winter-training-institute2017/paul-andersen

Neumann, K., Fischer, H. E., & Kauertz, A. (2010). From PISA to educational standards: The impact of large-scale assessments on science education in Germany. *International Journal of Science and Mathematics Education*, *8*(3), 545-563. https://doi.org/10.1007/s10763-010-9206-7

NGSS Lead States. 2015. *NGSS Evidence Statement:MS-PS2-2*. Washington, DC: The National Academies Press. Retreived January 7, 2019 from: https://www.nextgenscience.org/sites/default/files/evidence_statement/black_white/MS-PS2-2%20Evidence%20Statements%20June%202015%20asterisks.pdf

Nollmeyer, G.E., & Bangert, A.W. (2015) Assessing K-5 elementary teachers understanding and readiness to teach the new framework for science education. *The Researcher, 27*(2), 7-13.

Nollmeyer, G. E., & Bangert, A. W. (2017). Measuring elementary teachers' understanding of the NGSS framework: An instrument for planning and assessing professional development. *Electronic Journal of Science Education*, *21*(8), 20-45.

NSTA. (n.d.). *About the next generation science standards.* Retrieved November 16, 2019, from https://ngss.nsta.org/About.aspx

Ortloff, W. G., & Escobar-Ortloff, L. M. (2001). *Professional development needs of American international schools overseas: An opportunity for service.* Southern Regional Council on Educational Administration. Retrieved March 11, 2019 from https://files.eric.ed.gov/fulltext/ED468742.pdf

Pallant, J. (2013). *SPSS survival manual*. UK: McGraw-Hill Education.

Pratt, H. (2013). *The NSTA reader's guide to the next generation science standards*. Arlington, VA: NSTA press.

Pratt, C. C., McGuigan, W. M., & Katzev, A. R. (2000). Measuring program outcomes: Using retrospective pretest methodology. *American Journal of Evaluation, 21*(3), 341-349. https://doi.org/10.1177/109821400002100305

Pruitt, S. L. (2014). The next generation science standards: The features and challenges. *Journal of Science Teacher Education*, *25*(2), 145-156. https://doi.org/10.1007/s10972-014-9385-0

Reeves, D. (2006). Leading to change/pull the weeds before you plant the flowers. *Educational Leadership*, *64*(1), 69-90.

Reeves, D. B. (2012). *Transforming professional development into student results*. Alexandria, VA: ASCD.

Reiser, B. J. (2013). *What professional development strategies are needed for successful implementation of the Next Generation Science Standards?* Invitational Research Symposium on Science Assessment. September 2013. Educational Testing Services.

Rossman, G. B., & Rallis, S. F. (2016). *An introduction to qualitative research: Learning in the field*. Thousand Oaks, CA: Sage Publications.

Rovegno, I. (1995). Theoretical perspectives on knowledge and learning and a student teacher's pedagogical content knowledge of dividing and sequencing subject matter. *Journal of Teaching in Physical Education*, *14*(3), 284-304. https://doi.org/10.1123/jtpe.14.3.284

Schmidt, J., & Nübling, R. (2005). *Assessment of the Outcome-Quality of Inpatient Psychosomatic Rehabilitation: A Comparison Between Different Methods of Change Measurement.* In A. Beauducel, B. Biehl, M. Bosnjak, W. Conrad, G. Schönberger, & D. Wagener (Eds.), Multivariate Research Strategies (pp. 261-281.). Auchen: Shaker.

Schmoker, M. (2018). *Focus: Elevating the essentials to radically improve student learning*. Alexandria, VA: ASCD.

Search Associates. (n.d.). *Bangkok-January.* Retrieved January 15, 2019, from https://www.searchassociates.com/international-teaching-job-fairs/bangkok-january/

Seoul Foreign School. (n.d.). *Tuition and Fee Schedule*. Retrieved February 3, 2020 from https://www.seoulforeign.org/admissions/tuition-fees/

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4-14. https://doi.org/10.3102/0013189X015002004

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-22. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Sibthorp, J., Paisley, K., Gookin, J., & Ward, P. (2007). Addressing response-shift bias: Retrospective pretests in recreation research and evaluation. *Journal of Leisure Research*, *39*(2), 295-315. https://doi.org/10.1080/00222216.2007.11950109

Singapore American School. (n.d.-a). *Academics.* Retrieved September 27, 2017, from https://www.sas.edu.sg/academics

Singapore American School. (n.d.-b). *By the numbers.* Retrieved February 2, 2020, from https://www.sas.edu.sg/about-us/the-sas-story/by-the-numbers

Singapore American School. (n.d.-c). *SAS Fee Schedule and Policy*. Retrieved February 3, 2020 from https://www.sas.edu.sg/fs/resource-manager/view/0fc4cbf2-f7b6-4a12-b820-2ca6c42476c8

Snyder, T.D. (2018). *Mobile Digest of Education Statistics, 2017* (NCES 2018- 138). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Strauss, A. & Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques.* Newbury Park, CA: Sage Publications.

Stevenson, H., & Stigler, J. W. (1994). *Learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. New York, NY: Simon and Schuster.

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, *80*(1), 99-103. https://doi.org/10.1207/S15327752JPA8001_18

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Tate, N (2016). What Are International Schools For? *International Schools: Current Issues and Future Prospects*. In M. Hayden, & J. Thompson (Eds.), *International Schools: Current Issues and Future Prospects.* (pp. 17-36). Oxford: Symposium Books Ltd.

Terwilliger, R. I. (1972, March). International schools: Cultural crossroads. *The Educational Forum, 36*(3), 359-363. https://doi.org/10.1080/00131727209338990

Thalheimer, W., & Cook, S. (2002, August). *How to calculate effect sizes from published research articles: A simplified methodology.* Retrieved Dec 30, 2019 from http://work-learning.com/effect_sizes.htm.

The American School in Switzerland, (n.d.). *Tuition and Fees.* Retrieved February 11, 2020 from https://www.tasis.ch/page.cfm?p=736

United World Colleges. (n.d.-a). *Schools and Colleges.* Retrieved February 03, 2020 from https://www.uwc.org/schools

United World Colleges. (n.d.-b). *What is UWC?*. Retrieved January 06, 2019, from https://www.uwc.org/about

U.S. Department of State. (n.d.-a). *Office of Overseas Schools.* Retrieved November 7, 2018 from https://www.state.gov/m/a/os/

U.S. Department of State. (n.d.-b) *Schools Worldwide*. Retrieved April 10, 2019 from https://www.state.gov/m/a/os/c1684.htm

Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of Educational Psychology, 103*(4), 952. https://doi.org/10.1037/a0025125

Walker, G. (2016). International Schools and International Curricula. In M. Hayden & J. Thompson (Eds.), *International Schools: Current Issues and Future Prospects* (pp. 37-52). Oxford: Symposium Books Ltd.

Ward, P., Tsuda, E., Dervent, F., & Devrilmez, E. (2018). Differences in the content knowledge of those taught to teach and those taught to play. *Journal of Teaching in Physical Education*, *37*(1), 59-68. https://doi.org/10.1123/jtpe.2016-0196

Waterson, M. (2016). The Corporatisation of International Schooling. In Hayden & Thompson (Eds.), *International Schools: Current Issues and Future Prospects* (pp. 185-214). Oxford: Symposium Books Ltd.

Wells, J. G. (2019). *STEM education: The potential of technology education*. Council on Technology and Engineering Teacher Education. Retrieved November 5, 2019 from https://vtechworks.lib.vt.edu/bitstream/handle/10919/93963/Wells_Ch10_STEM-Potential-of-TechEd-2019YrBk.pdf?sequence=1&isAllowed=y

Windschiti, M. A., & Stroupe, D. (2017). The three-story challenge: Implications of the next generation science standards for teacher preparation. *Journal of Teacher Education*, *68*(3), 251-261. https://doi.org/10.1177/0022487117696278

Wilde, C. L. (2018). How teachers are making sense of the next generation science standards in secondary schools: A mixed-methods study (Doctoral dissertation). Retrieved August 10, 2019 from https://csusm-dspace.calstate.edu/bitstream/handle/10211.3/205262/WildeChristina_Summer2018.pdf?sequence=1

Zilber, E. (2005). International school educators and their children: Implications for educator-parents, colleagues and schools. *Journal of Research in International Education*, *4*(1), 5-22. https://doi.org/10.1177/1475240905050288

APPENDIX A

ACTION RESEARCH CYCLE 1 INTERVIEW PROTOCOL

Action research cycle 1 interview protocol:

Perceived Barriers

1. For first- to eighth-grade teachers, what are the perceived barriers that inhibit implementation of the NGSS?

    a. What do you perceive to be the greatest barriers to implementation of the NGSS for middle school science teachers?

    b. What do you perceive to be the greatest barriers to implementation of the NGSS for elementary school science teachers?

Perceived Benefits

2. What types of NGSS-related professional development have you conducted that has been most beneficial to you?

    a. What aspects of your previously completed professional development helpful?

    b. What aspects of your previously completed professional development were unhelpful?

    c. Other than the NGSS-related professional development that you've conducted, what other sorts of NGSS-related professional development do you think would be valuable to you?

Resources

3. What kinds of other resources (coaching, classroom support, etc.) will be needed to assist teachers to implement the NGSS?

APPENDIX B

SCIENCE INSTRUCTIONAL PRACTICES SURVEY AND SCORING GUIDE

Science Instruction Practices Survey

This survey is adapted from the work of Hayes, K. N., Lee, C. S., DiStefano, R., O'Connor, D., & Seitz, J. C. (2016).

Reference: 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*(2), 137-164.

How often do your students do each of the following in your science class?	Never	Rarely (a few times per year)	Sometimes (Once or twice per month)	Often (once or twice per week)	Daily or almost daily
1. Generate questions or predictions to explore	1	2	3	4	5
2. Identify questions from observations or phenomena	1	2	3	4	5
3. Choose variables to investigate (such as in a lab setting)	1	2	3	4	5
4. Design or implement their own investigations	1	2	3	4	5
5. Make and record observations	1	2	3	4	5
6. Gather quantitative and qualitative data	1	2	3	4	5
7. Organize data into charts or graphs	1	2	3	4	5
8. Analyze relationships using charts or graphs	1	2	3	4	5
9. Analyze results using basic calculations	1	2	3	4	5
10. Explain the reasoning behind an idea	1	2	3	4	5
11. Respectfully critique each other's reasoning	1	2	3	4	5
12. Supply evidence to support a claim or explanation	1	2	3	4	5
13. Consider alternative explanations.	1	2	3	4	5

14. Make an argument that supports or refutes a claim.	1	2	3	4	5
15. Create physical model of a scientific phenomenon.	1	2	3	4	5
16. Develop a conceptual model based on data or observations (model is not provided by textbook or teacher)	1	2	3	4	5
17. Use models to predict outcomes.	1	2	3	4	5
18. Use models to predict outcomes. How often do you do each of the following in your science instruction.	1	2	3	4	5
19. Provide direct instruction to explain science concepts	1	2	3	4	5
20. Demonstrate an experiment and have students watch	1	2	3	4	5
21. Use activity sheets to reinforce skills or content	1	2	3	4	5
22. Go over science vocabulary	1	2	3	4	5
23. Apply science concepts to explain natural events or real-world situations	1	2	3	4	5
24. Talk with your students about things they do at home that are similar to what is done in science class	1	2	3	4	5
25. Discuss students' prior knowledge or experience related to the science topic or concept	1	2	3	4	5

SIPS Survey Scoring Guide

This survey scoring guide is adapted from the work of Hayes, K. N., Lee, C. S., DiStefano, R., O'Connor, D., & Seitz, J. C. (2016).

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*(2), 137-164.

Composite Factor	NGSS SE Practice	Survey Item
Instigating an Investigation	Questioning; Planning and Carrying Out Investigations	1. Generate questions or predictions to explore
		2. Identify questions from observations or phenomena
		3. Choose variables to investigate (such as in a lab setting)
		4. Design or implement their own investigations
Data Collection and Analysis	Planning and Carrying Out an Investigation; Analyze and Interpret Data; Using Mathematical and Computational Thinking	5. Make and record observations
		6. Gather quantitative and qualitative data
		7. Organize data into charts or graphs
		8. Analyze relationships using charts or graphs
		9. Analyze results using basic calculations
Critique, Argumentation and Explanation	Constructing Explanations; Engaging in Argument from Evidence	10. Explain the reasoning behind an idea
		11. Respectfully critique each others' reasoning
		12. Supply evidence to support a claim or explanation
		13. Consider alternative explanations.
		14. Make an argument that supports or refutes a claim.
Modeling	Developing and Using Models	15. Create physical model of a scientific phenomenon.
		16. Develop a conceptual model based on data or observations (model is not provided by textbook or teacher)
		17. Use models to predict outcomes.

Traditional Instruction	18. Use models to predict outcomes. How often do you do each of the following in your science instruction.
	19. Provide direct instruction to explain science concepts
	20. Demonstrate an experiment and have students watch
	21. Use activity sheets to reinforce skills or content
	22. Go over science vocabulary
Prior Knowledge	23. Apply science concepts to explain natural events or real-world situations
	24. Talk with your students about things they do at home that are similar to what is done in science class
	25. Discuss students' prior knowledge or experience related to the science topic or concept

APPENDIX C

3D-PAST: THREE-DIMENSIONAL PERFORMANCE ASSESSMENT SCREENING TOOL

The Three-Dimensional Performance Assessment Screening Tool (3D-PAST) is typically distributed to teachers as a two-sided, laminated card, with the images show below printed on either side. Teachers use this tool in conjunction with a pre-made assessment and a detailed description of the relevant Next Generation Science Standard (NGSS) Performance Expectation (PE).

Side 1: Description of the process for using the tool and definitions of key vocabulary.

Side 2: Checklist of key aspects of assessments that are properly aligned with the Next Generation Science Standards three-dimensional performance expectations.

Performance Assessment Screening Tool

1. **Read** or **take** the entire assessment
2. Apply the **checklist** in order (1,2,3,...11)
3. Give **feedback** on missing elements

Performance Expectation (PE) - the entire standard (e.g. MS-LS1-1)

Disciplinary Core Idea (DCI) - the content (e.g. F=ma)

Science and Engineering Practices (SEP) - elements of scientific inquiry and engineering design (e.g., Modeling).

Crosscutting Concepts (CCC) - interdisciplinary thinking strategies (e.g. Patterns)

Phenomenon - fact or situation that is observed

Problem - a need or desire that can be solved

Stimuli - information (e.g. data, text, etc.) required for the prompts

Prompts - questions

☐ 1. The **prompts** match the Science and Engineering Practice (SEP) and engage students in sense making.

☐ 2. The **stimuli** have the required information needed to utilize the SEP. (e.g., data for analysis)

☐ 3. The **stimuli** have multiple and sufficient information needed open up the SEP. (a rich task)

☐ 4. The **prompts** elicit observable understanding of the Disciplinary Core Idea. (DCI)

☐ 5. The **prompts** include the Crosscutting Concept. (CCC)

☐ 6. The **prompts** include language (i.e., bullets) from grade appropriate progressions. (DCI)(CCC)(SEP)

☐ 7. The **prompts** include graphic organizers.

☐ 8. The **entire assessment** contains information that is scientifically accurate and properly attributed.

☐ 9. The **prompts** points in the direction of explaining the phenomenon or designing a solution.

☐ 10. The **phenomenon** or **problem** is authentic, interesting, and requires students to figure something out.

☐ 11. The **phenomenon** or **problem** is novel to show the transfer of knowledge. (e.g., not in the unit)

Source:

Andersen, P., & Brosnick, L. (n.d.). *3-Dimensional Performance Assessment Screening Tool* [Pdf]. Paul Andersen. Retrieved January 9, 2018 from https://thewonderofscience.com/s/3D-Screening-Tool-8lxx.pdf

APPENDIX D

NEXT GENERATION SCIENCE STANDARDS PERFORMANCE EXPECTATION EVIDENCE STATEMENT EXAMPLE

Associated with each Next Generation Science Standard student performance expectation is an evidence statement document which describes each of the three dimensions integrated into the standard, and grade-level appropriate observable features. Below is a partial image of the evidence statements document for MS-PS2-2, a physical science standard for the middle school level.

MS-PS2-2 Motion and Stability: Forces and Interactions

Students who demonstrate understanding can:

MS-PS2-2. **Plan an investigation to provide evidence that the change in an object's motion depends on the sum of the forces on the object and the mass of the object.** [Clarification Statement: Emphasis is on balanced (Newton's First Law) and unbalanced forces in a system, qualitative comparisons of forces, mass and changes in motion (Newton's Second Law), frame of reference, and specification of units.] *[Assessment Boundary: Assessment is limited to forces and changes in motion in one-dimension in an inertial reference frame and to change in one variable at a time. Assessment does not include the use of trigonometry.]*

The performance expectation above was developed using the following elements from the NRC document *A Framework for K-12 Science Education:*

Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Planning and Carrying Out Investigations Planning and carrying out investigations to answer questions or test solutions to problems in 6–8 builds on K–5 experiences and progresses to include investigations that use multiple variables and provide evidence to support explanations or design solutions. • Plan an investigation individually and collaboratively, and in the design: identify independent and dependent variables and controls, what tools are needed to do the gathering, how measurements will be recorded, and how many data are needed to support a claim. - ***Connections to Nature of Science*** **Scientific Knowledge is Based on Empirical Evidence** • Science knowledge is based upon logical and conceptual connections between evidence and explanations.	**PS2.A: Forces and Motion** • The motion of an object is determined by the sum of the forces acting on it; if the total force on the object is not zero, its motion will change. The greater the mass of the object, the greater the force needed to achieve the same change in motion. For any given object, a larger force causes a larger change in motion. • All positions of objects and the directions of forces and motions must be described in an arbitrarily chosen reference frame and arbitrarily chosen units of size. In order to share information with other people, these choices must also be shared.	**Stability and Change** • Explanations of stability and change in natural or designed systems can be constructed by examining the changes over time and forces at different scales.

Observable features of the student performance by the end of the course:		
1		Identifying the phenomenon to be investigated
	a	Students identify the phenomenon under investigation, which includes the change in motion of an object.
	b	Students identify the purpose of the investigation, which includes providing evidence that the change in an object's motion is due to the following factors:
		i. Balanced or unbalanced forces acting on the object.
		ii. The mass of the object.
2		Identifying the evidence to address the purpose of the investigation
	a	Students develop a plan for the investigation individually or collaboratively. In the plan, students describe*:
		i. That the following data will be collected:
		1. Data on the motion of the object.
		2. Data on the total forces acting on the object.
		3. Data on the mass of the object.
		ii. Which data are needed to provide evidence for each of the following:
		1. An object subjected to balanced forces does not change its motion (sum of F=0).
		2. An object subjected to unbalanced forces changes its motion over time (sum of F≠0).

Source:

NGSS Lead States. 2015. *NGSS Evidence Statement:MS-PS2-2*. Washington, DC: The National Academies Press. Retrieved January 7, 2019 from: https://www.nextgenscience.org/sites/default/files/evidence_statement/black_white/MS-PS2-2%20Evidence%20Statements%20June%202015%20asterisks.pdf

APPENDIX E

NEW FRAMEWORK OF SCIENCE EDUCATION SURVEY OF TEACHER UNDERSTANDING

This instrument was distributed in electronic form, via Qualtrics. As of March 20, 2020, it is hosted at: https://asu.co1.qualtrics.com/jfe/form/SV_a4wbAG4QmxaMQLj

New Framework for Science Education: Survey of Teacher Understanding Developed by Nollmeyer & Bangert (2015)													
		Before using 3D-PAST						After using 3D-PAST					
		Level of understanding…						Level of understanding…					
		None	Slight	Fair	Solid	Strong	Advanced	None	Slight	Fair	Solid	Strong	Advanced
Construct 1: Science & Engineering Practices													
1.	When planning and teaching, educators have students participate in practices used by scientists and engineers in the real world.	1	2	3	4	5	6	1	2	3	4	5	6
2.	When planning and teaching, educators have students ask questions about scientific phenomena that can drive exploration.	1	2	3	4	5	6	1	2	3	4	5	6
3.	When planning and teaching, educators have students ask questions to define engineering problems that can drive design.	1	2	3	4	5	6	1	2	3	4	5	6
4.	When planning and teaching, educators have students develop and refine conceptual models to express their understanding about scientific phenomena.	1	2	3	4	5	6	1	2	3	4	5	6
5.	When planning and teaching, educators have students develop models to visualize and refine an engineered design.	1	2	3	4	5	6	1	2	3	4	5	6
6.	When planning and teaching, educators have students plan and carry out investigations to gather data about scientific phenomena and engineering problems.	1	2	3	4	5	6	1	2	3	4	5	6
7.	When planning and teaching, educators have students apply mathematical and computational thinking to investigate scientific questions and engineering problems.	1	2	3	4	5	6	1	2	3	4	5	6

8.	When planning and teaching, educators have students construct evidence-based explanations to describe phenomena that incorporate their understandings about science.	1	2	3	4	5	6	1	2	3	4	5	6
9.	When planning and teaching, educators have students design and refine solutions that meet the needs of an engineering problem.	1	2	3	4	5	6	1	2	3	4	5	6
10.	When planning and teaching, educators have students engage in evidence-based argumentation about scientific explanations and engineered designs.	1	2	3	4	5	6	1	2	3	4	5	6
11.	When planning and teaching educators have students communicate ideas clearly and persuasively through words, images, and other media.	1	2	3	4	5	6	1	2	3	4	5	6
Construct 2: Teaching Disciplinary Core Ideas													
12.	When planning and teaching, educators focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.	1	2	3	4	5	6	1	2	3	4	5	6
13.	When planning and teaching, educators recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.	1	2	3	4	5	6	1	2	3	4	5	6
14.	When planning and teaching, educators use a learning progression approach by building from prior knowledge and working towards future sophistication.	1	2	3	4	5	6	1	2	3	4	5	6
15.	When planning and teaching, educators include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.	1	2	3	4	5	6	1	2	3	4	5	6
16.	When planning and teaching, educators include core ideas that are important in investigating more complex ideas and solving problems.	1	2	3	4	5	6	1	2	3	4	5	6
17.	When planning and teaching, educators include core ideas that relate to the interests and life experiences of students or societal concerns.	1	2	3	4	5	6	1	2	3	4	5	6

18.	When planning and teaching, educators recognize that the construction of knowledge requires active participation on the part of the students.	1	2	3	4	5	6	1	2	3	4	5	6
Construct 3: Crosscutting Concepts													
19.	When planning and teaching, educators have students consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.	1	2	3	4	5	6	1	2	3	4	5	6
20.	When planning and teaching, educators have students develop an understanding that phenomena work differently at different scales.	1	2	3	4	5	6	1	2	3	4	5	6
21.	When planning and teaching, educators have students use systems thinking when investigating scientific phenomena.	1	2	3	4	5	6	1	2	3	4	5	6
22.	When planning and teaching, educators have students consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.	1	2	3	4	5	6	1	2	3	4	5	6
23.	When planning and teaching, educators have students investigate phenomena in terms of structure and function as a means of sense-making.	1	2	3	4	5	6	1	2	3	4	5	6
24.	When planning and teaching, educators have students identify what aspects of a system remain stable over time and what aspects undergo patterns of change.	1	2	3	4	5	6	1	2	3	4	5	6
Construct 4: Integration of the Three Dimensions													
25.	When planning and teaching, educators have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.	1	2	3	4	5	6	1	2	3	4	5	6
26.	When planning and teaching, educators intentionally select practices and concepts that best facilitate student sense-making for particular core ideas.	1	2	3	4	5	6	1	2	3	4	5	6
27.	When planning and teaching, educators have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.	1	2	3	4	5	6	1	2	3	4	5	6

Construct 5: Best Practices in Science Education														
28.	When planning and teaching, educators use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.	1	2	3	4	5	6	1	2	3	4	5	6
29.	When planning and teaching, educators have students engage in sustained investigations accompanied by necessary teacher support.	1	2	3	4	5	6	1	2	3	4	5	6
30.	When planning and teaching, educators teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.	1	2	3	4	5	6	1	2	3	4	5	6
31.	When planning and teaching, educators teach students how mathematical concepts and skills apply to scientific investigation and engineering design.	1	2	3	4	5	6	1	2	3	4	5	6

APPENDIX F

EMAIL REQUEST FOR PARTICIPATION

The email below was forwarded to appropriate faculty via administrators (typically individuals serving in the role of Director of Teaching and Learning or similar) or lead faculty at each of the participating schools. Included in the email were links to the digital version of the NFSE-STU and a link to volunteer to be interviewed, both hosted by Qualtrics.

- - - - - - - - - -

Dear International School Colleague,

My name is Wyatt Wilcox. I am a science teacher at Singapore American School, and also a doctoral candidate in Arizona State University's Doctor of Education in Leadership and Innovation program. My dissertation research seeks to better understand the effect of professional development activities which target assessment design for the Next Generation Science Standards.

As a faculty member at an international school that has worked with Paul Andersen on NGSS assessment design during the past school year, I'm reaching out to ask if you would be willing to complete a short online survey and/or engage in an interview about the way the activities have impacted your understanding of the NGSS.

If you're willing to participate in this study, please follow the appropriate link below, where you will be presented with additional information about the study and the opportunity to consent.

Note: Survey responses are completely anonymous. If you choose to engage in an interview, the researcher is not able to link the survey data to the interview.

Thank you, so much, for your consideration!

> > SURVEY ONLY or BOTH < <
You will be directed to the survey. If after submitting the survey, you would be willing to participate in an interview, you will be directed to a separate, unlinked, contact form

which will allow you to submit contact information without associating it with your survey responses.

> > INTERVIEW ONLY < <
You will be directed to submit your contact information. This will allow me to contact you for the purpose of arranging a convenient time for the interview.

Warm regards,

Wyatt Wilcox
Doctoral Candidate, Ed.D. Leadership and Innovation
Arizona State University

Middle School Science Teacher
Singapore American School

APPENDIX G

CONSENT TO PARTICIPATE IN INTERVIEW

Shown below is the informed consent for participation in the semi-structured interview. This informed consent was distributed via Qualtrics. Participants were required to acknowledge and consent to participation prior to providing their contact information which would be subsequently used by the researcher to schedule the interview.

- - - - - - - - - - -

Dear International School Colleague,

My name is Wyatt Wilcox. I am an 8th-grade science teacher at Singapore American School, and also a doctoral candidate in Arizona State University's Doctor of Education in Leadership and Innovation program. My dissertation study seeks to better understand the effect of professional development activities which target assessment design for the Next Generation Science Standards.

As a faculty member at an international school that has worked with Paul Andersen on NGSS assessment design during the past school year, I am seeking your participation in an online survey and/or to engage in an interview about the way the activities have impacted your understanding of the NGSS.

You must be at least 18 or older to participate in this study.

There are no foreseeable risks or discomforts to your participation. There are no direct benefits to you. Results from the study may serve to improve future professional development activities for international school educators.

Participation in either or both parts of the study is voluntary. If you choose not to participate or to withdraw from the study, there will be no penalty.

The results of this study may be used in reports, presentations, or publications but your name will not be used. Your identity will remain anonymous. All data from the study will be maintained in a secure location.

Interview Participation:

The interview is expected to require approximately 30 minutes. It will be conducted by an online video conference, arranged at a time that is convenient for you. I am also asking your permission to audio record the interview. Only the research team will have access to the recordings. The recordings will be deleted immediately after being transcribed and any published quotes will be anonymous. To protect your identity, please refrain from using names or other identifying information during the interview. At any time, you may request the recording to be stopped, and it will be stopped.

If you have any questions concerning this research study, you may contact me at wwilcox1@asu.edu, or Dr. Gustavo Fischman at fischman@asu.edu. Additionally, if you have any questions about your rights as a participant in this research, or if you feel you have been placed at risk, you may contact the Chair of the Human Subjects Institutional Review Board, through the Arizona State University Office of Research Integrity and Assurance, at 480.965.6788.

If you are willing to participate in this study, please indicate your consent below.

Sincerely,

Wyatt Wilcox

Doctoral Candidate, Ed.D. Leadership and Innovation
Arizona State University

You must consent to participate in the interview:

- ❑ Yes. I have read the conditions of the study and consent to participate in an interview.

Please provide your contact information. This information will be used by the researcher (Wyatt Wilcox) solely for the purpose of scheduling the interview.

Name:
Email:

APPENDIX H

INVITATION TO SCHEDULE INTERVIEW

The following email was sent to interview volunteers following the researcher's receipt of their contact information.   The email includes a link to a Google Appointment Calendar which aids scheduling of interview appointments across the varying time zones in which participants were located.

- - - - - - - - - -

Hello <<Interview Participant>>

Thank you for indicating that you are willing to be engaged for my study examining the impact of NGSS professional development in international schools. As a fellow international school educator, I know how busy we all are at this time of year, so your participation is very much appreciated.

I'd like to arrange an interview with you at a time that is convenient.   Due to the time zone differences, this may be challenging.   However, I have created a Google Appointment Calendar that has appointment slots indicating times that I can be available. If you are signed-in to a Google account, it should display the times in your local time.

Please take a look at this appointment slot calendar. You can either select a spot yourself or email me with a time that works for you.

I expect the interviews to last no more than 30 minutes, and I will use a video-chat software called Zoom, which can be run through your internet browser.

If you have any questions or concerns, please don't hesitate to ask.

Regards,

Wyatt Wilcox
Doctoral Candidate, Ed.D. Leadership and Innovation
Arizona State University

APPENDIX I

SEMI-STRUCTURED INTERVIEW PROTOCOL

Semi-Structured Interview Protocol

1. Welcome & Introduction

   Hello, my name is Wyatt Wilcox. First, let me thank you for your time and willingness to participate in this interview about your use of 3D-PAST. I expect this interview to last about 30 minutes.

   This interview will be recorded so that I can recall the information and points that were discussed.

   Is this acceptable to you?

2. Review of Consent Form & Secondary Verbal Agreement

   *consent form will be distributed and returned electronically, prior to the interview.

   Thank you for completing the consent form and returning it to me. Do you have any questions about the consent form, and also can you reconfirm verbally that you've given your consent to participate in this research?

3. Ground Rules

   I will be asking you a series of questions about your use of 3D-PAST and the way it has affected your understanding of the NGSS. These questions form the framing of the interview, and I encourage you to elaborate on these points, but you are also welcome to talk about things that I haven't asked a direct question about. The point of this interview is to elaborate and extend the quantitative survey that you completed earlier. Your responses will be used to inform the study which I am completing.

4. Participant Introduction
   a. *Please tell me a little about yourself and your professional context.*
      i. *What is your school, position, and number of years that you've taught a the school?*
      ii. *Where is your school in the processes of implementing the NGSS?*

iii. *Can you describe your experience with 3D-Past prior to this activity?*

5. 3D-PAST & Construct Questions

   a. Science and Engineering Practices
      i. *Can you describe your experience with teaching science and engineering practices, prior to using 3D-PAST, as they relate to the NGSS?*
      ii. *Do you feel that using 3D-PAST to evaluate your NGSS internal assessments changed the way you understood the NGSS or some aspect of what the NGSS requires of you as a teacher with regard to teaching science and engineering practices?*

   b. Teaching Disciplinary Core Ideas
      i. *Can you describe your experience with teaching the NGSS's disciplinary core ideas, prior to using 3D-PAST, as they relate to the NGSS?*
      ii. *How do you feel that using 3D-PAST to evaluate your NGSS internal assessments changed the way you understood the NGSS or some aspect of what the NGSS requires of you as a teacher with regard to teaching disciplinary core ideas?*

   c. Crosscutting Concepts
      i. *Can you describe your experience with teaching the NGSS's crosscutting concepts, prior to using 3D-PAST, as they relate to the NGSS?*
      ii. *How do you feel that using 3D-PAST to evaluate your NGSS internal assessments changed the way you understood the NGSS or some aspect of what the NGSS requires of you as a teacher?*

   d. Integration of the Three Dimensions
      i. *Can you describe your experience, prior to using 3D-PAST, with integrating the three dimensions (science and engineering practices, crosscutting concepts, and disciplinary core ideas) as they relate to the NGSS?*
      ii. *How do you feel that using 3D-PAST to evaluate your NGSS internal assessments changed the way you understood the NGSS*

or some aspect of what the NGSS requires of you as a teacher?

e. Best Practices in Science Education
    i. *How do you feel that using 3D-PAST to evaluate your NGSS internal assessments changed the way you understood the NGSS or some aspect of what the NGSS requires of you as a teacher with regards to best practices in science education?*

f. General Questions Related to 3D PAST Usage:
    i. *Prior to using 3D-PAST, what has been your experience with designing assessments that are aligned to the NGSS?*
    ii. *Do you feel that 3D-PAST is a useful tool? Would you continue to use it?*
    iii. *In using 3D-PAST, did you find yourself referring to the detailed NGSS PEs? In what ways were those useful?*

g. Non-Specific
    i. *Are there any other points related to using 3D-PAST that you would like to elaborate on?*

6. Closing
    a. *Those are all of the questions that I have for you. I want to thank you again for your participation in the study and also for your time during this interview today. Your contributions have provided me with useful information.*

    Thank you.

APPENDIX J

SKEW AND KURTOSIS OF NFSE-STU CONSTRUCT SCORES

Skewness and Kurtosis for Mean Differences Within Constructs

Construct	N	Skewness	Std. Error	Kurtosis	Std. Error
SEP Mean Differences	84	1.074	0.263	1.537	0.52
DCI Mean Differences	81	0.931	0.267	1.716	0.529
CCC Mean Differences	78	1.879	0.272	4.029	0.538
I3D Mean Differences	77	0.665	0.274	0.462	0.541
BP Mean Differences	78	0.953	0.272	3.919	0.538

APPENDIX K

INSTITUTIONAL REVIEW BOARD APPROVAL

EXEMPTION GRANTED

Gustavo Fischman
Division of Educational Leadership and Innovation - Tempe
480/965-5225
fischman@asu.edu

Dear Gustavo Fischman:

On 4/17/2019 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Improving Science Education in International Schools Through Professional Development Targeting Next Generation Science Standards Assessment Design
Investigator:	Gustavo Fischman
IRB ID:	STUDY00010047
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	• Wilcox_Interview Guide, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions); • Wilcox_Consent_Form_Interview, Category: Consent Form; • Wilcox_Interview_Contact_Form_Qualtrics, Category: Recruitment materials/advertisements /verbal scripts/phone scripts; • Wilcox_IRB_Supplemental, Category: Other (to reflect anything not captured above); • Wilcox_Social_Behavioral_Protocol, Category: IRB Protocol; • Wilcox_Consent_Form_Survey, Category: Consent Form; • Wilcox_Recruitment Email, Category: Recruitment Materials; • Wilcox_Recruitment Email, Category: Recruitment materials/advertisements /verbal scripts/phone scripts;

	• Wilcox_Survey Instrument Qualtrics Form, Category: Measures (Survey questions/Interview questions /interview guides/focus group questions);

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (2) Tests, surveys, interviews, or observation on 4/17/2019.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,

IRB Administrator

cc: Wyatt Wilcox
Wyatt Wilcox

APPENDIX L

DEMOGRAPHIC INFORMATION OF NFSE-STU RESPONDENTS

Demographic Data for NFSE-STU Respondents by School

School	N[i]	Highest Degree Held			Grade Level Taught			Years of Experience in International Schools		
		BA	MA	PhD	ES	MS	HS	Min	Max	Avg
American International School Vienna	3	1	2	0	0	2	1	3	11	8
American International School Dhaka	1	0	1	0	0	0	1	15	15	15
International School of Beijing	5	1	4	0	0	2	3	3	14	8
Colegio Nueva Granada	9	2	7	0	5	0	4	1	16	8
The American School in London[ii]	1	-	-	-	-	-	-	-	-	-
American School of Guatemala	4	1	3	0	3	1	0	2	20	9
Hong Kong International School	3	1	2	0	0	1	2	1	7	4
American International School of Budapest	2	0	1	1	0	1	1	2	7	4.5
American Community School of Amman	2	0	2	0	1	0	1	10	10	10
American School Foundation of Monterrey	7	0	7	0	4	1	2	7	21	12
Colegio Franklin Delano Roosevelt	1	0	1	0	1	0	0	9	9	9
Singapore American School	30	5	24	1	17	6	7	1	27	13
American School of Barcelona	3	0	2	1	0	1	2	2	8	4
American School of Dubai	3	1	2	0	0	1	2	15	17	16
American Community School of Abu Dhabi	2	1	1	0	0	1	2	4	8	6
Total:	76	13	59	3	31	17	28	-	-	-
Average:	-	-	-	-	-	-	-	5	14	9

[i]*The total N of survey respondents was 84, however only 76 respondents provided demographic data.*
[ii]*A single respondent indicated employment with the American School in London, however opted out of providing further demographic data.*

APPENDIX M

INTERNAL CONSISTENCY FOR NFSE-STU

Cronbach Alpha Analysis for Pre- and Post- Intervention Constructs

		Coefficient Alpha	
Construct	Survey Items	Pretest	Posttest
Science and Engineering Practices	1-11	0.935	0.91
Disciplinary Core Ideas	12-18	0.932	0.884
Crosscutting Concepts	19-24	0.925	0.881
Integration of the Three Dimensions	25-27	0.87	0.875
Best Practices	28-31	0.884	0.815

APPENDIX N

WILCOXON SIGNED-RANK TESTS FOR NFSE-STU RESPONSE ITEMS

Wilcoxon Signed-Rank Tests for All NFSE-STU Response Items

Construct	Response Item			N	Mdn	Z	Sig. (2-Tailed)	*r*	
Science and Engineering Practices									
	When planning and teaching, educators have students...								
	1	participate in practices used by scientists and engineers in the real world.	Pre	83	3	6.32	.000	0.49	Medium
			Post	84	4				
	2	ask questions about scientific phenomena that can drive exploration.	Pre	83	3	6.93	.000	0.54	Large
			Post	84	4				
	3	ask questions to define engineering problems that can drive design.	Pre	82	3	6.07	.000	0.47	Medium
			Post	83	4				
	4	develop and refine conceptual models to express their understanding about scientific phenomena.	Pre	81	3	6.80	.000	0.53	Large
			Post	84	5				
	5	develop models to visualize and refine an engineered design.	Pre	78	3	5.59	.000	0.45	Medium
			Post	82	4				
	6	plan and carry out investigations to gather data about scientific phenomena and engineering problems.	Pre	84	4	5.64	.000	0.44	Medium
			Post	84	4				
	7	apply mathematical and computational thinking to investigate scientific questions and engineering problems.	Pre	81	3	3.71	.000	0.29	Small
			Post	83	4				

8	construct evidence-based explanations to describe phenomena that incorporate their understandings about science.	Pre	80	3	6.81	.000	0.49	Medium
		Post	84	4				
9	design and refine solutions that meet the needs of an engineering problem.	Pre	80	3	4.71	.000	0.37	Medium
		Post	82	3				
10	engage in evidence-based argumentation about scientific explanations and engineered designs.	Pre	81	3	5.87	.000	0.46	Medium
		Post	83	4				
11	communicate ideas clearly and persuasively through words, images, and other media.	Pre	82	4	5.18	.000	0.40	Medium
		Post	83	4				
Disciplinary Core Ideas								
	When planning and teaching, educators...							
12	focus on a few core ideas instead of a large number of topics so that students can achieve greater depth in their understanding.	Pre	78	4	4.99	.000	0.40	Medium
		Post	81	4				
13	recognize that the development of student understandings of disciplinary core ideas is a progression that takes place over years.	Pre	80	4	4.91	.000	0.39	Medium
		Post	82	4				
14	use a learning progression approach by building from prior knowledge and working towards future sophistication.	Pre	79	4	4.16	.000	0.33	Medium
		Post	81	4				

15	include core ideas that have broad importance across multiple disciplines or are key organizing principles within a discipline.	Pre Post	79 78	4 4	5.33	.000	0.42	Medium
16	include core ideas that are important in investigating more complex ideas and solving problems.	Pre Post	79 79	4 4	4.71	.000	0.37	Medium
17	include core ideas that relate to the interests and life experiences of students or societal concerns.	Pre Post	79 79	4 4	3.58	.000	0.28	Small
18	recognize that the construction of knowledge requires active participation on the part of the students.	Pre Post	79 81	5 5	3.48	.001	0.27	Small
Crosscutting Concepts								
	When planning and teaching, educators have students...							
19	consider issues of cause and effect when questioning and discussing scientific phenomena or engineering designs.	Pre Post	76 80	3 4	5.39	.000	0.44	Medium
20	develop an understanding that phenomena work differently at different scales.	Pre Post	76 79	3 4	4.82	.000	0.39	Medium
21	use systems thinking when investigating scientific phenomena.	Pre Post	76 79	3 4	5.29	.000	0.43	Medium

22	consider that since energy and matter are conserved, much can be determined by studying their flow into and out of systems.	Pre	77	3	4.71	.000	0.38	Medium
		Post	79	4				
23	investigate phenomena in terms of structure and function as a means of sense-making.	Pre	77	3	4.72	.000	0.38	Medium
		Post	78	4				
24	identify what aspects of a system remain stable over time and what aspects undergo patterns of change.	Pre	77	2	5.27	.000	0.42	Medium
		Post	78	4				
Integration of the Three Dimensions								
	When planning and teaching, educators...							
25	have students explore disciplinary ideas by engaging in practices and making connections through crosscutting concepts.	Pre	75	2	6.12	.000	0.50	Large
		Post	77	4				
26	intentionally select practices and concepts that best facilitate student sense-making for particular core ideas.	Pre	76	3	5.36	.000	0.43	Medium
		Post	78	4				
27	have students use the crosscutting concepts when engaging in practices about disciplinary core ideas.	Pre	74	3	5.79	.000	0.48	Medium
		Post	78	4				

Best Practices in Science Education									
		When planning and teaching, educators...							
	28	use both teacher-led and student-led strategies to facilitate student understanding of science and engineering content.	Pre	77	4	3.72	.000	0.30	Medium
			Post	78	4				
	29	have students engage in sustained investigations accompanied by necessary teacher support.	Pre	77	4	3.80	.000	0.31	Medium
			Post	79	4				
	30	teach students how to present their scientific ideas and engineering solutions with clarity through both the written and spoken word.	Pre	76	4	4.54	.000	0.37	Medium
			Post	79	4				
	31	teach students how mathematical concepts and skills apply to scientific investigation and engineering design.	Pre	78	4	3.53	.000	0.28	Small
			Post	78	4				