8-2024

Elementary Science Teacher Self-Efficacy in the Context of NGSS-aligned Curriculum Enactment: An Explanatory Sequential Mixed Methods Study

Wesley P. Sliger
University of Nebraska-Lincoln

ELEMENTARY SCIENCE TEACHER SELF-EFFICACY IN THE CONTEXT OF NGSS-ALIGNED CURRICULUM ENACTMENT: AN EXPLANATORY SEQUENTIAL MIXED METHODS STUDY

by

WESLEY P. SLIGER

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Educational Studies

Under the Supervision of Professor Lawrence Scharmann

Lincoln, Nebraska

August, 2024

ELEMENTARY SCIENCE TEACHER SELF-EFFICACY IN THE CONTEXT OF NGSS-ALIGNED CURRICULUM ENACTMENT: AN EXPLANATORY SEQUENTIAL MIXED METHODS STUDY

Wesley Patrick Sliger, Ph.D.

University of Nebraska, 2024

Advisor: Lawrence Scharmann

A confluence of circumstances in educational policy and teacher preparation have placed a low priority on elementary science teacher preparation. Historically, a lack of elementary science teachers' subject matter knowledge and pedagogical knowledge has negatively impacted science teaching self-efficacy, which is linked to many teacher and student outcomes. Meanwhile, the past decade has ushered in the Next Generation Science Standards (NGSS) and the modern era of reform-based science teaching. The emergence of self-proclaimed NGSS-aligned resources has outpaced academic examination regarding professional development (PD), curriculum implementation, and curriculum enactment. The purpose of this study was to use social cognitive theory to assess whether the implementation of a phenomena-based science curriculum was related to science teaching self-efficacy and outcome expectations of elementary science teachers. A mixed methods explanatory sequential case selection variant design was used. The initial quantitative strand used the STEBI-A tool to gather information about participants' science teaching self-efficacy and outcome expectancy. A subset of the initial participants was selected for a qualitative interview based on their range of STEBI-A scores. The cases revealed that teachers with higher science teaching self-efficacy and outcome expectancies enjoyed science, prioritized science, and recounted positive student

engagement, interest, and motivation more frequently than teachers with lower STEBI-A scores. Teachers involved in a curriculum selection process or those that received coaching sessions from a science coordinator reported developing skills and confidence in teaching science. All teachers reported increased comfort in teaching while using NGSS-aligned curriculum materials over time, but also lacked time to teach science during the school year and referenced their limited experience with professional development. Participants held a wide range of beliefs about the nature of inquiry in science. Teachers' use of science curriculum support materials was varied as well. From a PD standpoint, more time must be apportioned to developing reform-oriented practices in elementary science teachers, and a teacher leadership model could be used to accomplish these goals when contractual time is limited.

ACKNOWLEDGEMENTS

I could not have taken this journey without the unrelenting support of my wife, Kelly. You have been my loudest, most persistent, and most loyal cheerleader. Words cannot express the gratitude I have for the stress you endured and the late nights and long weekends you sacrificed to take care of our home and our family. Thank you to my daughter, Edie, who reminded me that it's human to take a break from work on this "sisseration" to color mermaids and eat snacks. I'm thankful for the rest of my family, friends, and coworkers who encouraged me and showed an interest in my work, even if every conversation was cut short by them desperately checking the time.

I'd like to thank my middle school students and my preservice teachers, who taught me thousands of lessons about science, teaching, and learning. I'm grateful for Mr. Thompson, whose mother handmade a blanket embroidered with "Sliger, Ph.D." four years too early. I really didn't want to live with the shame of owning that blanket without finishing the job, so thanks for the motivation!

Major thanks to my doctoral committee for providing support and guidance. Dr. Males gave thorough and thoughtful rough draft feedback. Dr. Gatti gave me hope that the COVID-19 pandemic wouldn't unravel democracy or my sanity. Dr. Zeleny provided endless warmth and support that inspired creativity and autonomy. Dr. Scharmann gave me the gift of mentorship. Having a mentor in the world of science education with a sense of pragmatism has been a blessing, but your genuine investment in my professional and academic development holds immeasurable value to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Introduction

A confluence of circumstances in educational policy and teacher preparation have placed a low priority on elementary science teacher preparation. Legislation in the United States has supported an environment for high-stakes testing, placing an emphasis on literacy and math while leaving science on the fringe (No Child Left Behind [NCLB], 2002; Lontok et al. 2015). Elementary teachers have recently demonstrated lower preparedness in science compared to both their secondary peers and compared to their own preparedness in math (Banilower et al., 2018). Seven percent of preservice elementary teachers take no science courses during their teacher preparation program and only twenty-three percent take a single science course (Banilower et al., 2018).

Meanwhile, the past decade has ushered in the modern era of reform-based science teaching, which is billed at minimum as an evolution of practice (Shapiro & Kraus, 2023) and to some as a "true paradigm shift" (Metz, 2013). The *Framework for K-12 Science Education* and subsequent development of the *Next Generation Science Standards (NGSS)* ambitiously laid out the knowledge and skills that students needed to acquire, but left little guidance to support teachers in their science teaching practices (National Academy of Sciences, Engineering, and Medicine [NASEM], 2015). The shifts proposed by these documents were large enough that teachers' understanding of them remain lacking to this day (Lowell & McNeill, 2023).

Historically, a lack of elementary science teachers' subject matter knowledge and pedagogical knowledge has negatively impacted a host of individual factors such as science teaching identity (Kane & Varelas, 2016) as well as attitudes, beliefs, and self-efficacy (Mulholand & Wallace, 1996; Rice & Roychoudhury, 2003). A resultant lack of science teaching efficacy presents further problems such as low effort and persistence (Gibson & Dembo, 1984), lower commitment to the profession (Coladarci, 1992), and lower willingness to innovate instruction (Allinder, 1994). These teacher factors impart their ill-effects onto student attitudes, beliefs, and achievement (Moore & Esselman, 1992; Ross, 1992; Watson, 1991; Woolfolk, Rosoff, & Hoy, 1990).

As educational policymakers, states, districts, schools, teachers, and other stakeholders in science education have attempted to keep abreast of science education reforms, curriculum publishers have rushed in to advertise new textbooks, lab kits, curriculum packages, and professional development opportunities. Almost overnight, publishers made claims that their materials were NGSS-aligned (NSTA, 2013). At the time, anyone could make this claim and slap an “NGSS-aligned” sticker on their materials with impunity. Consumers may have taken these claims at face value, especially without a clear system for school districts to vet materials or share evidence of the effectiveness of these curricula (NASEM, 2015). A literature review by Davis and colleagues (2016) focusing on science curriculum use gives reason for skepticism surrounding emerging curriculum materials, finding that teachers tend to fall back on prior knowledge and experience, which may or may not align with reform-based science teaching.

The emergence of self-proclaimed NGSS-aligned resources has outpaced academic examination regarding professional development, curriculum implementation strategies, and curriculum enactment (NASEM, 2015), creating gaps in existing literature. All the recommendations offered from current and future studies must eventually be filtered through the lens of teachers who undergo professional development, make curricular decisions, and actually enact curriculum. Science teacher efficacy is a key characteristic that must be considered in the reform-based science teaching era.

Elementary science teachers must not only believe that there is value in reform-based science teaching, but have the confidence to implement sound learning experiences in science (Love, Napoli, and Lee, 2023).

Although there are hundreds of examples where science teaching efficacy has been measured with respect to other variables (Deehan, 2016), few of these studies have been conducted during the NGSS era with a specific focus on NGSS-aligned practices and curriculum materials at the elementary level. The Science Teaching Efficacy Belief Instrument A (STEBI-A) is a valid and reliable instrument developed by Riggs and Enochs (1990) that measures the science teaching efficacy and outcome expectancy of in-service science teachers. I performed a literature review to examine contemporary uses of STEBI to understand literature on elementary science teaching self-efficacy. I used the following keyword terms: “STEBI,” STEBI NGSS,” “STEBI implementation,” “STEBI fidelity,” “science self-efficacy,” and “NGSS self-efficacy.”

Two of the reviewed studies used quantitative methods. A study on the effects of a writing methods course using poetry with science themes did not yield significant changes in overall self-efficacy (Love, Napoli, & Lee, 2023). Another study on in-service elementary science teachers showed that the act of unpacking NGSS standards improved science teaching self-efficacy (Robertson, 2022).

Five of the reviewed studies using the STEBI used a mixed-methods approach Two of the five studies focused on pre-service elementary science teachers. A study using an inquiry-embedded physics content course improved science teaching efficacy and subject matter knowledge (Menon & Sadler, 2016). A study using a science methods course focusing on engineering yielded improved science teaching efficacy (Yesilyurt et al., 2021). The other three studies focused on in-service elementary science teachers and used professional development programs that lasted from one to three years in length. All of the professional development programs focused on student-centered practices and resulted in an increase in science teaching self-efficacy (Kang et al., 2019; Mentzer et al., 2017; Sandholtz and Ringstaff, 2014).

I did not find any studies describing the relationship between any specific NGSS-aligned science curriculum enactment and science teaching self-efficacy. A research gap exists because prior research has not adequately compared cases in terms of quantitative differences in science teaching self-efficacy due to specific NGSS-aligned curriculum use. There is a need to use quantitative data and especially qualitative data to explain the results in detail through the voices and perspectives of participants. A study employing these methods would benefit the field and practitioners working with preservice elementary science teachers to navigate the NGSS era of science standards. Policymakers and administrators should be able to use this study to inform decision-making on curriculum and professional development. Teachers may use this study to understand the perspectives of their peers and to reflect on their own practice.

Purpose Statement

The purpose of this study is to use social cognitive theory to assess whether the implementation and enactment of a phenomena-based science curriculum relates to science teaching self-efficacy and outcome expectations of elementary science teachers. An explanatory sequential case selection variant mixed methods design will be used that involves collecting quantitative data first and then explaining the quantitative results with in-depth qualitative data. In the first, quantitative phase of the study, I collected Science Teaching Efficacy Belief Instrument A (STEBI-A) data from 3rd to 5th grade elementary school teachers that teach science at a midwestern, urban school district to assess their current science teaching self-efficacy and outcome expectations. In the second phase, I conducted qualitative semi-structured interviews in a multiple case study as a follow-up to help explore the experiences of four 3rd to 5th grade elementary science teachers representing different combinations of science teaching self-efficacy and outcome expectations.

Research Questions/Hypothesis

The research questions guiding this study are written in the style of Creswell and Plano Clark (2018) to reflect the mixed methods methodology used in this study. The research questions are as follows:

1. **Quantitative Research Question:** How do elementary teachers rate themselves on science teaching self-efficacy and outcome expectancy during the implementation of an NGSS-aligned science curriculum?
2. **Qualitative Research Question:** How do elementary teachers describe their science teaching self-efficacy and outcome expectancy during their curriculum implementation experiences?
3. **Mixed Methods Research Question:** In what ways do the interview findings with elementary science teachers help to explain the quantitative results on science teaching self-efficacy and outcome expectancy?

Theoretical Framework

This study operates under the worldview of pragmatism. Pragmatism focuses on the research questions asked rather than the methods employed by the researcher. Taking a real-world, practice-oriented approach values both objective and subjective knowledge. In this worldview, the false dichotomy between postpositivism and constructivism creates barriers between fields of study that are unhelpful artifacts of unnecessarily partitioned thought (Tashakkori & Teddlie, 2010).

Bandura's (1982) work on social cognitive theory provides the theoretical framework for the current study. Social cognitive theory tells us much about the environment's relationship with an individual's self-efficacy, that is, an individual's level of confidence in performing a specific behavior. Where self-efficacy represents the expectations surrounding a person's ability, outcome expectancy represents the expected consequences of performing a task at a particular level of ability. The resulting interaction of these two constructs results in a predictable matrix of affective reactions and behaviors as shown in Table 1 (Bandura, 1982). For example, a strong sense of self-efficacy paired with a responsive environment that rewards achievement will result in generally "assured, active responsiveness." In contrast, the interaction between low self-efficacy and any outcome expectation tends to lead toward undesirable affective states.

Table 1. Self-Efficacy Judgment vs. Outcome Judgment

	Low Outcome Judgment	High Outcome Judgment
High Self-Efficacy Judgment	Social Activism Protest Grievance Milieu Change	Assured, Opportune Action
Low Self-Efficacy Judgment	Resignation Apathy	Self-Devaluation Despondency

Note. From "Self-Efficacy Mechanism in Human Agency", by A. Bandura, 1982, *The American Psychologist*, *37*(2), p. 140 (https://doi.org/10.1037/0003-066X.37.2.122).

When individuals are efficacious, but not able to achieve the desired responsiveness from their environment, their behaviors and efforts will not necessarily screech to a halt. However, efficacious individuals will seek out control of their environment or may attempt to modify the social systems that, in part, control outcomes (Bandura, 1982).

Bandura's stance was that self-efficacy causally influenced outcome expectancy, but not vice versa, stating that it would be peculiar if "…outcomes that flow from actions are made to precede the actions. (Bandura, 1997)" Some studies (Kirsch, 1982; Corcoran & Rutledge, 1989) have shown that outcome expectancies can indeed influence self-efficacy, contrary to Bandura's arguments. Kirsch (1982) offered snake-fearful college students hypothetical monetary incentives to interact with a snake, and self-efficacy increased as a result of increasing incentives. Corcoran and Rutledge (1989) found similar effects with smokers promised hypothetical monetary incentives that increased with longer periods of time without smoking. These effects are stronger for behaviors that involve regulation of behavior rather than physical skills.

Furthermore, outcome expectancies may influence self-efficacy regardless of the skills an individual actually possesses. In a study among chronic pain patients asked to perform functional tasks, the patients shared expected pain and fear of injury as reasons for their self-efficacy ratings (Council, Ahern, Follick, & Kline, 1988). Rhodes and Blanchard (2007) found similar results when asking college students their reasons for self-efficacy ratings for exercise, with the students reporting expectations of improved health as an influencing factor. Williams (2010) has found that Bandura's view of outcome expectancy remains the status quo in self-efficacy research and calls for more serious efforts by researchers to "disentangle self-efficacy from the influence of expected outcomes."

Definitions of Key Terms

To facilitate the understanding of this study, different terms are defined herein.

Curriculum. "Consciously directed training experiences that schools use for completing and perfecting the unfoldment" of abilities (Bobbitt, 1918).

Feasibility. "The likelihood that a project, program, or intervention can be successfully implemented in a school or other institution. (Hauk & Kaser, 2020)"

Curriculum Efficacy Study. Studies aiming to understand whether an intervention produces a desired effect, at least under extremely ideal circumstances.

Fidelity of Implementation. An alignment between planned program components and a teachers' enactment of said components (O'Donnell, 2008).

Fidelity to Goal. An alignment between classroom practices and the overall goal of the curriculum (McNeill, 2018).

Educative Curriculum Materials. Materials included in a curriculum with the intent of supporting teachers' pedagogical content knowledge and subject matter knowledge.

Subject Matter Knowledge. The knowledge a teacher has in a specific discipline (e.g., science).

Pedagogical Content Knowledge. The knowledge of teaching (Shulman, 1987).

Reciprocal Determinism. The dynamic relationship between a person, their environment, and their behavior in Bandura's social cognitive theory.

Self-Efficacy. An individual's level of confidence in performing a specific behavior.

Outcome Expectancy. The expected consequences of performing a task at a particular level of ability.

Teacher Efficacy. A teacher's confidence in their ability to help students achieve specific educational outcomes.

Summary

The reform-based science teaching movement has given academics, policymakers, and educators hope that learning science through inquiry can be accomplished in a way NGSS's forebearers had envisioned. Although stakeholders in education such as teacher preparation programs and school districts have made strides in addressing the shifts posed in science teaching practices, they still fall short of supporting elementary science teachers in teaching students to use authentic science practices.

Through this study, I plan to investigate how elementary science teachers rate their science teaching efficacy in the context of NGSS-aligned curriculum implementation. I hope to examine the supports that teachers receive and how teachers use reform-based science teaching materials.

CHAPTER II: REVIEW OF THE LITERATURE

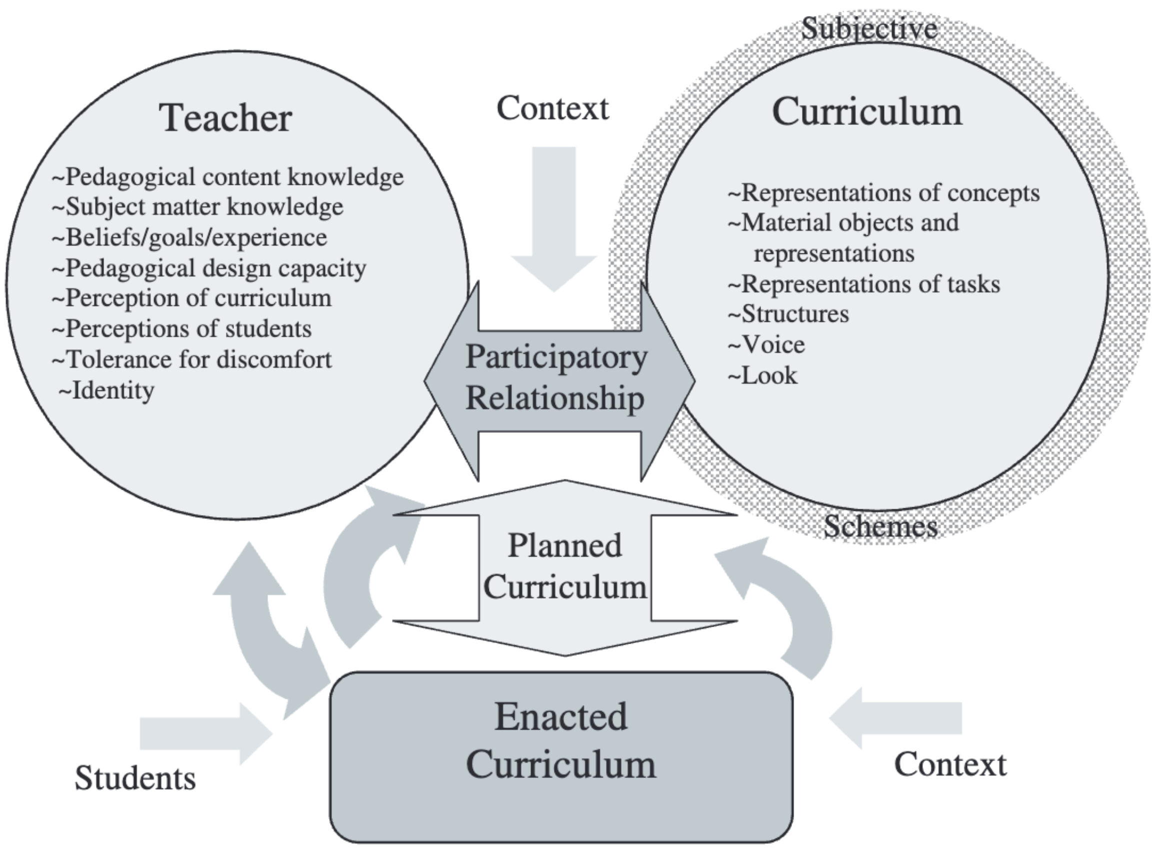
Curriculum as a Construct

Historically, the concept of "curriculum" has been loosely defined for centuries (Hamilton, 2014). In *The Curriculum*, Bobbitt (1918) defines curriculum as "consciously directed training experiences that schools use for completing and perfecting the unfoldment" of abilities. Bobbitt also acknowledged that some of these "training experiences" are undirected, that is, taught through socialization. Most experts in the field of curriculum theory don't agree on a definition for curriculum (Wiles, 2008), and work to parse out the ambiguities of this conception. Curricula can be broadly classified by their aims such as skills development, learner-centered practices, social justice goals, and subject-specific academic acculturation (Yaşar & Aslan, 2021). The parties that interact with curriculum are usually easily identifiable, but the same cannot be said for the purposes and aims of curriculum. Decisions regarding the selection of knowledge to be learned, how that knowledge should be obtained, and questions around cultural, moral, and ethical perspectives abound the landscape of curriculum.

Once a curriculum developer considers the weight of these philosophical conundrums, then they can begin to work through the tension between theory and practice. Remillard's (2005) framework of the participatory relationship between teachers and curriculum provides a useful visual that uses nodes called "constructs" to demonstrate the interplay between the two. In this framework, the curriculum construct is narrowed down to well-defined characteristics such as materials, representations, and tasks. A sampling of the complexity of individual characteristics of teachers is listed in Figure 1. Individual teacher factors such as beliefs and knowledge have been shown to be influential in curriculum use (Darling-Hammond et al., 2002; Lewis et al., 2021).

A series of interpretations and analyses both shaped and constrained by a teacher's experience transforms the vision of the curriculum developer into the "planned curriculum." This interpretation is further transformed into the curriculum that is enacted in the classroom. The bi-directional arrows between the dimensions of this model show reciprocal change. One relevant example to this study is that when teachers use educative curriculum materials, their beliefs and practices may change (Marco-Bujosa et al., 2017).

Figure 1. Framework of Components of Teacher-Curriculum Relationship



Note. From "Examining key concepts in research on teachers' use of mathematics curricula" by J. T. Remillard, 2005, *Review of Educational Research, 75*(2), p. 235. (https://doi.org/10.3102/00346543075002211).

Curriculum Implementation

Curriculum transforms theory into actionable learning processes and materials. Although curriculum design is a variable process, the intervention must be feasible at minimum. Hauk and Kaser (2020) provide a useful description of feasibility, which is "the likelihood that a project, program, or intervention can be successfully implemented in a school or other institution." They further categorize 18 factors associated with successful program implementation into four components of feasibility: technical, organizational, support, and usability. Technical factors are physical materials of the curriculum, such as hardware and software, assignments, and lab equipment. Organizational factors include establishing commitment and trust from involved stakeholders, managing potentially conflicting commitments, and ensuring standards alignment. Support factors include professional development opportunities and ongoing technical support. Usability factors describe the ease of use and the quality of materials, but also include contextual factors. These contextual factors include a willingness for change by the target users and an overall program fit.

The internal validity of an intervention is established in efficacy studies. Feasibility is considered one focus of efficacy studies, which aim to understand whether an intervention produces a desired effect, at least under extremely ideal circumstances. In efficacy studies, fidelity is carefully controlled as if in a laboratory setting. At this stage, if a curriculum does not reliably obtain its expected results, then the blame rests on the program's underlying theory or model. The external validity, or generalizability, of an intervention occurs during effectiveness studies, which occur in field settings. Efficacy studies are a gatekeeper to moving forward with effectiveness studies, because a failure at the efficacy stage demonstrates failure of theory rather than failure of implementation (O'Donnell, 2008).

Calls for the measurement of the fidelity of curriculum implementation have led to overlapping terminologies in the literature (National Research Council, 2004). Fidelity of implementation (FOI) is a construct loosely defined as an alignment between planned program components and a teachers' enactment of said components (O'Donnell, 2008). Subcategories of FOI include fidelity to procedure (i.e. number and order of methods), fidelity to structure (i.e. adherence, duration), and fidelity to process (i.e. quality of delivery) (McNeill et al., 2018). The increased attention to fidelity is not without merit. Increasing the FOI of an intervention has shown to increase program effectiveness (Blakely et al., 1987) and student achievement (Kurz et al., 2010).

Given this knowledge about fidelity, questions immediately arise concerning teachers' use of curricular materials. From a curriculum designer's perspective, attempts to make materials "teacher proof" (Bolin, 1987; Krajcik et al., 2000; Welch, 1979) almost seem reasonable. This slight to practitioners can understandably lead to pushback. The field of education walks the line between science and art. Some of the "art" of teaching is precipitated by an educator's agency, knowledge, and expertise. These valuable assets may not overlap with those of curriculum authors.

A strict adherence to curriculum delivery may seem unpalatable to practitioners who must meet the unique needs of their students and contend with the complex and competing demands of their stakeholders. Teachers may adapt curriculum to maximize students' access to the material, scaffolding materials in context-specific ways a curriculum author may not anticipate. Experienced teachers tend to follow a trajectory of effective curriculum use by moving from using curriculum materials as-is, then creating new materials, followed by adapting materials (Taylor, 2013). Using existing curriculum materials or creating new materials is far less cognitively demanding than assessing the relevance of existing materials and modifying them in a deliberate manner that encompasses both the curriculum's approach and the teacher's approach.

A construct called fidelity to goal provides a meaningful compromise that focuses on aligning classroom practices with the overall goal of the curriculum (McNeill, 2018). Deviating from prescribed components such as activity duration and provided assessments becomes less of an issue so long as the rationale behind curricular modifications are faithful upon teacher enactment (Blakely et al., 1987; Ben-Peretz, 1990). However, Taylor (2013) points out that not only is making curricular modifications a cognitively demanding task, but it is also difficult to do well, that is, in a way that aligns the goals, style, and philosophy of the curriculum with the beliefs and practices of the teacher.

Next Generation Science Standards and Inquiry-based Instruction

In 2012, the National Research Council convened a multidisciplinary committee of scientists, educators, and policy makers to develop the *Framework for K-12 Science Education*. The *Framework* drew on scientific research on science learning to articulate learning expectations for science students. The *Framework* laid the foundation for the development of the *Next Generation Science Standards (NGSS)* in 2013. The NGSS are standards, not curriculum (NGSS Lead States, 2013). The standards describe the knowledge and skills that students need to acquire but leave curricular decisions such as the methods and manner of instruction to teachers, schools, districts, and states. NGSS organized three-dimensional science standards that diverged from prior standards in that they included so-called science and engineering practices (SEPs) and crosscutting concepts (CCCs). The first two dimensions are SEPs which represent skills and practices that scientists and engineers use, and CCCs that provide a lens through which the nature of science can be explored. The third dimension is the disciplinary core ideas (DCIs), which are the content to be learned. NGSS has been adopted or adapted by 44 U.S. states and Washington, D.C. (National Science Teaching Association [NSTA], 2022).

Science standards revision and curriculum reform were partially spurred by claims made in *A Nation at Risk* (U.S. National Commission on Excellence in Education, 1983) that the United States' relevance in science and technology was waning. The National Science Education Standards (National Research Council [NRC], 1996) were the iteration of national standards prior to NGSS, and the NRC's charge was to present a vision for scientific literacy across the nation.

In these standards, the phrase "scientific inquiry" was used to describe how teaching and learning science occurred, but without the level of thoroughness and specificity that NGSS undertook. In NGSS "inquiry-based science" became explicitly defined by the SEPs and CCCs. Constructivism as a learning theory undergirds these practices, and student-centered learning became the pedagogy espoused by inquiry-based instruction. For example, the science and engineering practice of "asking questions" is aimed at students with the intent of generating ideas for investigation, mirroring authentic scientific processes. The teacher is positioned as a facilitator who steers students toward a path that remains germane to their learning. The values of constructing and using science knowledge using NGSS's three-dimensional structure is sometimes referred to as "ambitious science teaching" (Davis et al., 2016).

Science Curriculum Implementation

Early NGSS adopters and adapters no longer needed to "unpack" standards, that is, decoding standards to determine content to be learned, skills, and age-appropriate assessment. Curricular flexibility was touted in the *NGSS Lead States* (2013) document. However, this same document containing over 500 pages of standards-related information left additional guidance to be desired (National Academy of Sciences, Engineering, and Medicine [NASEM], 2015). The multidisciplinary approach was clearly a strength in the creation of rigorous science standards. The subsequent curriculum design and implementation process did not immediately continue this same spirit of collaboration (NASEM, 2015). The National Academy of Sciences, Engineering and Medicine (NASEM) met to discuss the “gap between what science teaching and learning could be and the reality of current practices” (NASEM, 2015). NGSS had alluded to general structures for curriculum implementation, but NGSS itself was not curriculum with its own embedded practices. NASEM pointed out that more serious consideration needed to be given to professional learning opportunities, resources, expertise, and implementation strategies.

Davis and colleagues (2016) reviewed literature related to science teachers use of science curriculum materials and what effects science curriculum use has on teachers and students. Their review elaborated on the practical considerations exposed by NASEM. Unsurprisingly, teachers’ prior experiences and beliefs play a central role in curricular decision-making. The shifts proposed by NGSS and the *Framework* were large enough that teachers’ understanding of the changes themselves were lacking (Lowell & McNeill, 2023). A teacher’s understanding of, and agreement with, the intent of an adopted curriculum can shape curriculum use (Davis et al., 2016).

One theme revealed in a review by Davis and colleagues (2016) is that teachers tend to align their curriculum to their current practices, for better or worse. For example, Davis (2006) observed pre-service elementary teachers using science curriculum materials for the purpose of student engagement rather than for the curriculum's intended purpose of building science practices. A separate study followed pre-service elementary teachers that were inconsistently using curriculum materials that assessed students' science inquiry abilities. The belief of the teachers in this study was that inquiry skills themselves did not require assessment in the same vein as content knowledge (Beyer and Davis, 2012).

Other deviations from ambitious science curriculum materials are the result of teachers' conflicting goals. The reduction of the cognitive demand of science tasks such as modeling and writing are commonly observed in the literature. While these curricular alterations are intentional, they represent a compromise that maintains other aspects of the classroom ecosystem such as student participation, lesson momentum, and content coverage.

In light of the literature on curriculum enactment, Janssen and colleagues (2015) provide three practical recommendations that may inform successful science curriculum implementation. One recommendation is that adopted curricula should communicate procedures that have classroom validity. The use of effectiveness studies can help curriculum designers determine the generalizability of their materials. The second recommendation is to make the investment of time and resources worthwhile and realistic to teachers.

Janssen's final recommendation is to promote congruence between curriculum practices and teachers' prevailing practices. The inclination for teachers to slide into habits that are comfortable and have promoted prior success can inhibit the implementation of research-based, reform-oriented practices (Blakely et al., 1987; Ben-Peretz, 1990). Curriculum designers can promote research-based practice by including educative features in curricular materials. Educative materials are included in a curriculum with the intent of supporting teachers' subject matter knowledge (SMK) and pedagogical content knowledge (PCK). Examples of educative materials include text or video-based components that define terms or learning objectives, descriptions and rationales for new teacher practices, and models of sound instruction (Davis & Krajcik, 2005). Marco-Bujosa and colleagues (2017) found that educative materials in one middle school science curriculum supported active reflection and purposeful lesson adaptations that didn't undermine the curriculum's intent to support students' science reasoning and argumentation skills. However, the authors acknowledged that the mere inclusion of educative features did not guarantee their use or teachers' understanding of their underlying intent.

Instructional Models in Science Education

There are various models of inquiry-based learning, but many can trace their lineage to the 5E Instructional Model. This model was developed by the Biological Sciences Curriculum Study (BSCS) in the 1980s and became a cornerstone for science teacher education programs and also provided some utility in other content areas. The five "Es" are *Engage, Explore, Explain, Elaborate, and Evaluate* (Bybee et al., 2006). The purpose of the *engage* phase is to garner student interest and motivation by introducing a phenomenon at the beginning of an instructional unit. The *explore* phase facilitates student learning through labs and activities. The *explain* phase often involves a collaborative effort between teachers and students to analyze and articulate patterns which can be formally described or defined. The *elaborate* phase extends student learning through related labs and activities, and the *evaluate* phase aims to assess students' understanding of learned concepts.

Joswick and Hulings (2024) reviewed existing research on the 5E instructional model to ascertain its effectiveness. It was effective at increasing conceptual knowledge in science, especially in biology. Results about procedural understanding of science and mathematical concepts were less clear because this distinction was not made in many of the reviewed studies. The 5E model also increases attitudes and interest in science. In comparison to “traditional” models of instruction, 5E was more effective at increasing student understanding, but it was not significantly different in outcomes compared to other inquiry-based teaching methods such as cooperative learning structures. The review was unable to find studies that clearly detailed what embedded strategies were included during 5E instruction, which made isolating specific, impactful teaching methods impossible. It is plausible that a combination of inquiry-based teaching methods such as cooperative learning structures embedded within 5E lessons is what makes them effective.

Instructional shifts proposed by the *Framework* and a genuine concern that the 5E instructional model may not fully address the nature of science led to new instructional models in science. Reiser and colleagues (2021) made strides in addressing the shifts from the student perspective. They designed an instructional approach called *storylines* that supports coherence, encouraging students to build on and revise their knowledge in a way that makes the numerous standards of NGSS appear more connected.

The storyline approach involves four teaching routines that manage student investigations (Reiser et al., 2021). The first routine is the anchoring phenomenon routine, which provides the initial context for investigation and elicits student-driven questions for investigation. The second routine is navigation, which orients students toward future investigations by bridging the gap between what has been learned and what is left unanswered. The third “putting the pieces together” routine leverages NGSS’s science and engineering practices to help students develop a deeper understanding of new concepts. Finally, the problematizing routine exposes unanswered questions and flaws in existing explanations to drive further inquiry.

BSCS, the organization responsible for the 5E Instructional Model, has recently taken cues from contemporary research in science education and has revised their 5E model to create the Anchored Inquiry Learning (AIL) model. However, the storylines model has attracted the bulk of interest from curriculum designers and researchers since its inception due to its early and explicit ties to the *Framework*.

Professional Development for Teachers

The shift in instructional practices called upon by the *Framework* and NGSS will continue to require high quality professional development (PD). Two meta-reviews have tentatively identified six characteristics of professional development that improve student achievement (Cordingley et al., 2015; Dunst et al., 2015). The six characteristics are whether PD is sustained, is a collaborative effort, whether it is voluntary/involves participant buy-in, is subject-specific, involves outside expertise, and is practice-based. In the wider literature, these six features of PD are considered as consensus to a point that they inform policy in several countries. For example, 2015's Every Student Succeeds Act requires PD to be sustained, collaborative, and practice based if a program's goal is to receive funding from the U.S. government (Combs & Silverman, 2016).

Sims and Fletcher-Wood (2021) are skeptical of these conclusions, with their meta-analysis pointing to the inappropriate inclusion of many studies within previous meta-reviews. They specifically claim that the "collaboration" and "subject-specific" characteristics of PD currently lack evidential warrant. They also claim that sustained PD may not be what many policymakers expect it to be. Sims and Fletcher-Wood's review reveals that the duration of PD has no relationship on student achievement, at least in English language arts and mathematics. Just as renowned psychologist Hermann Ebbinghaus demonstrated in 1885, distributed practice and the spacing effect may be what undergirds a recommendation for sustained PD (Karpicke & Roediger, 2008). In this case quality is more important than quantity. Even participant buy-in was viewed with scrutiny. While not an undesirable aspect of PD, buy-in was seen as correlated, but not causal to effective PD.

The remaining characteristics of effective PD are that a program is sustained (but not necessarily lengthy), involves outside expertise, and is practice based. The "external expertise" recommendation means that teachers from different schools would be able to bring fresh ideas to the table and a willingness to challenge the status quo. The "practice-based" recommendation is characterized by actively applying ideas learned from a professional development program. Practice could happen in the form of discussion, lesson design or reflection, or teaching mock lessons.

Science Teaching Professional Development

The topic of science teaching professional development is ripe with research opportunities as curriculum developers produce novel NGSS-aligned materials, design professional development workshops, and measure feasibility or effectiveness of their materials. Lowell and McNeill (2023) studied curriculum-based professional development involving OpenSciEd curricular materials. OpenSciEd is a non-profit organization that builds reform-oriented open educational resources for the middle school level and uses the previously mentioned storylines model of instruction (Reiser et al., 2021).

Lowell and McNeill (2023) wanted to know how teacher beliefs and self-efficacy changed over the course of six OpenSciEd PD sessions. They used a PD structure that repeated four principles: encouraging teachers to take a student perspective, analyzing model teaching, analyzing contrasting cases of instruction, and engaging in curriculum enactment and reflection. Their results reiterated the importance of repeated practice in effective PD programs, citing the repeated structure of their sessions as a relevant variable rather than the amount of time used in each session. The program was successful in causing teachers to significantly shift from traditional science teaching beliefs toward reform-oriented beliefs after a single PD session consisting of four days of PD. Storyline implementation self-efficacy significantly increased after three PD sessions, then plateaued. This study shows that beliefs and rationale are important aspects of the teacher's relationship with curriculum as shown in Remillard's (2005) curriculum enactment framework in Figure 1. However, the gap between the change in participants' teaching beliefs and change in self-efficacy beliefs better exemplifies the importance of practice-based professional development. The teacher participants had time during and between PD sessions to rehearse and enact teaching routines, increasing their self-efficacy beliefs over time.

Penuel and colleagues (2023) studied teacher participants in another OpenSciEd PD program to take a closer look at their teaching beliefs and self-efficacy beliefs as they pertained to the teacher routines embedded within the storyline model of instruction. Their results reinforce the importance of two of the effective PD characteristics: sustained, repetitive structures and practice-based programs. Participants' teacher beliefs, knowledge of curricular structures, and science teaching efficacy improved after a three-year period with repeated PD sessions. Notably, teachers' self-efficacy in using storyline teacher routines such as the anchoring phenomenon routine increased, as did self-efficacy in using NGSS SEPs.

These studies provide examples of professional development that improved beliefs and self-efficacy, but demonstrated that there is plenty of room for improvement in other areas of curricular enactment. While some teachers' beliefs and self-efficacy improved more and faster with prior exposure to reform-oriented materials and PD, it was clear that simply repeating the same workshop model had limits to its effectiveness (Lowell & McNeill, 2023). Targeted instructional coaching may be necessary to bolster effectiveness in the numerous teacher routines used in the storylines model of instruction (Gibbons & Cobb, 2017). For example, the use of "talk moves" and promoting discussion structures remains a challenge for teachers new to curricular materials that intend to guide teachers toward using ambitious science practices (Lowell & McNeill, 2023).

Penuel and colleagues (2023) noted that the sheer number of teacher supports and artifacts created as part of storyline routines can be difficult to manage. In the vein of student discussion, there is also difficulty in managing the ability for students to understand that science requires both divergent and convergent coherence seeking. Divergent thinking is necessary to generate questions and critique flaws in current models, but convergent thinking is what develops consensus and theory.

Cherbow and McNeill (2022) also describe an ongoing tension between building students' epistemic agency and pseudoagency. Epistemic agency is the ability of a group of students to shape the knowledge-building of the classroom community. Pseudoagency is an instance that undermines agency where it is clear the teacher predetermined the course of an investigation. This occurs when teachers use pre-made lesson materials. Finally, Cherbow and McNeill address the role of scripting in storyline materials. On one hand, scripting takes on the role of an educative support, but on the other hand, sends a counterproductive message that teachers dare not deviate from the materials in lieu of following student needs and interests. Communicating all these potential pitfalls, even with repeated PD sessions, would be a complex achievement.

Elementary Science Teacher Preparation

Teacher preparation programs broadly prepare their students by offering courses in psychology, multicultural education, special education, and educational technology. Preparation for secondary teachers involves a heavy course load in a specific discipline such as science where teacher candidates learn the content and teaching methods unique to that field's philosophy. However, elementary education programs must take more varied approaches to prepare preservice teachers for their multidisciplinary role. Elementary education majors take content and methods courses in art, physical education, science, social studies, math, and literacy.

In the United States, legislation supporting an environment for high stakes testing has placed an emphasis on tested subjects such as literacy and math. While science is often an assessed subject across the country, science education is not always prioritized in elementary schools or in teacher preparation programs. In a report by The National Survey of Science and Mathematics Education (NSSME), elementary teachers demonstrated lower preparedness in science than their secondary peers and compared to their own preparedness in math (Banilower et al., 2018). As a result, elementary teachers that already see themselves as "literacy people" rather than "science people" may not develop a strong identity as science teachers (Kane & Varelas, 2016).

The National Science Teaching Association (NSTA) recommends that elementary teachers take coursework in life science, Earth science, and physical science to obtain the knowledge needed to teach the full breadth of contemporary science standards (NSTA, 2012). The NSSME report reveals a possible cause for elementary teachers' lower science preparedness, finding that seven percent of preservice elementary teachers take no science courses, while twenty-three percent only take courses in one of the three areas (Banilower et al., 2018).

Elementary Science Teacher Preparation and Subject Matter Knowledge

The NSTA's content preparation standards may seem arbitrary at first, but teachers must develop subject matter knowledge (SMK), especially elementary and ELL teachers who do not have the opportunity to engage deeply in content-specific courses during teacher preparation. Elementary teachers who do not develop SMK in science may hold similar misconceptions to their own students (Tilgner, 1990). However, simply taking science courses may not be sufficient. Learning through formal science courses that use ineffective teaching practices can cause pre-service teachers to develop negative attitudes and beliefs toward science teaching (Mulholand & Wallace, 1996; Rice & Roychoudhury, 2003), so it is important that both secondary and postsecondary science educators use research-based practices that authentically model scientific inquiry.

When teachers have higher subject matter knowledge, they hold and communicate fewer misconceptions about science ideas, reflect more deeply on pedagogy, and are able to determine appropriate scaffolding for students (Pando & Aguirre-Muñoz, 2021). SMK is foundational to pedagogical content knowledge (Abell, 2007), and these gains in teacher knowledge benefit student learning (Sadler et al., 2013). In the literature, SMK has become a strong predictor of the quality of preservice science teachers' lessons (Sullivan-Watts et al., 2013). Findings about teacher SMK parallel the findings about student language development in that learning science vocabulary is more impactful to SMK than learning general academic vocabulary (Ardasheva et al., 2019).

Elementary Science Teacher Preparation and Pedagogical Content Knowledge

Davis and Haverly's (2022) review of elementary science teacher preparation reveals that science methods courses can change preservice teachers' beliefs, attitudes, knowledge, and practices, giving hope that a troubling lack of elementary science teacher preparedness can be ameliorated. One goal of methods courses is to build pedagogical content knowledge (PCK), which is the knowledge of teaching (Shulman, 1987). PCK operationalizes SMK, allowing the teacher to analyze and understand problems of practice, respond to the needs of learners, and to make instructional moves. PCK is a combination of SMK, pedagogical knowledge, and contextual knowledge, or knowledge of the situational environment (Grossman, 1990). PCK offers a useful model of teacher knowledge, but its broad scope makes PCK a construct with fuzzy boundaries (Gess-Newsome & Lederman, 1999). Given this lack of distinction, a variety of pedagogies have been embedded in science methods courses and studied under the umbrella of PCK. Examples of science methods strategies supportive to developing PCK in preservice science teachers include making the nature of science explicit and contextualized (Bell et al., 2011), engaging in structured opportunities to practice science teaching (Wenner & Kittleson, 2018), and focusing on multicultural aspects of science education (Mensah & Jackson, 2018).

Teacher candidates apply what they learn from content and methods courses through experiences in supervised field placements under the tutelage of a mentor teacher at a placement site. Teacher preparation programs typically employ a field supervisor that observes the teacher candidate, ensuring the program's goals are being met by offering feedback and assigning course grades. The expertise of mentor teachers in elementary science may be hard to come by, as the *Framework* and NGSS present models of science teaching and learning that diverge from their experiences as K-12 learners or from their own learning in teacher preparation programs (NASEM, 2015; Reiser, 2013; Stroupe & Hancock II, 2022). Gaps commonly exist in communication and collaboration between teacher preparation programs and mentor science teachers (Zeichner, 2010), exacerbating the transition to reform-oriented science practices.

Teacher preparation programs are not fully to blame for the aforementioned shortcomings in elementary teacher preparedness in science. While there is high variability in teacher credentialing across institutions and states, there is high variability in NGSS implementation at almost every level, including in resources, school-by-school implementation, and professional development opportunities (NASEM, 2018).

Self-Efficacy

Albert Bandura helped lay the foundation for social learning theory in the 1960s, which became social cognitive theory with the inclusion of self-efficacy as a construct. Social learning theory's central feature is reciprocal determinism, which is the dynamic relationship between three factors: the person, their environment, and their behavior (Bandura, 1977). Self-efficacy is an individual's level of confidence in performing a specific behavior and is part of the "person" factor of reciprocal determinism along with personality, experience, beliefs, and attitudes. The "environment" factor includes the social context the person is in, and the "behavior" factor includes any verbal, motor, or social actions.

Bandura (1977) proposed that self-efficacy is influenced by four sources of information: enactive attainments, vicarious experiences, verbal persuasion, and the person's psychological state. Enactive attainments are mastery experiences, which involve incremental successes and gradually decreasing performance aids. Vicarious learning involves learning from competent models. Verbal persuasion serves a motivational role in efficacy beliefs. The final information source is psychological state, and reducing stressful emotive states can improve efficacy beliefs and subsequent performances. Each of the efficacy information sources was shown in Bandura's work to have a significant effect on efficacy in a variety of settings including coping with phobias and abstaining from smoking, but enactive mastery produced the highest increases in efficacy.

Where self-efficacy represents the expectations surrounding a person's ability, outcome expectancy represents the expected consequences of performing a task at a particular level of ability. For example, if a teacher has low self-efficacy in teaching science it may cause the expectation that students will become less interested in science or will not learn science content or skills well.

Teacher Efficacy

Teacher efficacy, or a teacher's confidence in their ability to help students achieve specific educational outcomes, was first coined in a study by Armor and colleagues (1976) who were evaluating the success of reading interventions. Two items in their questionnaire were correlated with student motivation and reading achievement. These items focused on a teacher's ability to overcome external factors beyond their control through effort, which is internally controlled by the teacher.

As expected in Bandura's model of reciprocal determinism, the context or environment that a teacher encounters can influence teacher efficacy. The multifaceted nature of teachers' work made teacher efficacy difficult to measure. Prompts that are too broad or too narrow threaten to limit the generalizability of teacher efficacy measurement tools. Gibson and Dembo (1984) created their own measures of teacher efficacy which revealed two factors that separately captured self-efficacy and outcome expectancy, thus conforming to Bandura's work. *Personal teaching efficacy* captured self-efficacy while *general teaching efficacy* captured outcome expectancy. Personal teaching efficacy was defined by Gibson and Dembo as the "belief that one has the skills and abilities to bring about student learning" (p. 573). Gibson and Dembo described general teaching efficacy as a "belief that any teacher's ability to bring about change is significantly limited by factors external to the teacher, such as the home environment, family background, and parental influences" (p. 574).

Gibson and Dembo's instrument inspired a wave of teacher efficacy research, with researchers modifying the original instrument to investigate teacher efficacy in classroom management, special education, and subject-matter-specific areas (Tschannen-Moran, Hoy, & Hoy, 1998). Bandura (1997) joined the teacher efficacy research arena much later with his Teacher Self-Efficacy Scale, which encompassed measurements of multiple teacher tasks and efficacies such as efficacy in enlisting parent involvement and disciplinary efficacy.

Using reciprocal determinism and social cognitive theory as a theoretical framework, many researchers began to explore self-efficacy's relationship with teacher behaviors and many correlates of teaching efficacy were revealed. For example, Gibson and Dembo (1984) found that teachers with high teacher efficacy expended more effort in the face of adversity, were more likely to help students persist through failure, and were more likely to differentiate instruction. Teacher efficacy is linked to the professional commitment of elementary and middle school teachers (Coladarci, 1992) and a willingness to implement innovative teaching strategies (Allinder, 1994). Furthermore, the positive outcomes of high teacher efficacy are related to improving student achievement

(Moore & Esselman, 1992; Ross, 1992; Watson, 1991) and attitudes toward school, learning, and their teacher (Woolfolk, Rosoff, & Hoy, 1990).

Science Teaching Efficacy and the STEBI

Qualitative work revealed some initial hints as to the connection between teaching efficacy and science teaching. For example, Ramey-Gassert and colleagues (1996) found that highly efficacious elementary science teachers were more willing to take on challenges. Appleton & Kindt (2002) found that new elementary teachers with high science teaching efficacy demonstrated evidence of developing as science teachers. However, self-efficacy beliefs are situational, and as such, teacher efficacy in one subject-specific area may have little effect on science teaching efficacy (Tschannen-Moran, Hoy, & Hoy, 1998).

Riggs and Enochs (1990) developed an instrument based off Gibson and Dembo's (1984) work to measure science teaching efficacy. This tool was called the Science Teaching Efficacy Belief Instrument (STEBI). The STEBI has been further modified to specifically measure science teaching efficacy of in-service teachers by using the STEBI-A, and of pre-service teachers by using the STEBI-B. The STEBI is not the only available tool used to measure science teaching efficacy, but through the past 30 years it has been used over 240 times, demonstrating its prevalence in a variety of international research settings and contexts (Deehan, 2016).

The STEBI consists of two subscale factors, *personal science teaching efficacy* (PSTE) and *science teaching outcome expectancy* (STOE). The PSTE subscale measures teachers' "beliefs about their own capacity to deliver science teaching experiences" (Deehan, 2016, p. 45) to aid students' development of what is now known in NGSS as DCIs and SEPs. Higher PSTE scores are associated with spending more time teaching science (Riggs & Jesunathadas, 1993), a preference to teach science (Lucas, Ginns, Tuli, & Waiters, 1993), and a humanistic orientation toward classroom control (Enochs, Scharmann, & Riggs, 1995). The STOE subscale measures "beliefs about the capacity of science teaching to overcome external factors to aid students' science learning in a general sense" (Deehan, 2016, p. 46). Higher STOE scores are associated with better ratings of science teaching effectiveness and attitude by observers (Enochs et al., 1995). Lower STOE scores are associated with using more text-based approaches, less activity-based approaches, and less cooperative structures for learning science (Riggs, 1995).

STEBI and Elementary Teachers' Enactment of NGSS

Although the STEBI has been used extensively over the past 30 years, its use has mostly occurred prior to the *Framework* and NGSS era. This section will be dedicated to contemporary uses of the STEBI in the elementary setting.

Robertson (2022) conducted a quantitative study to determine if the act of unpacking NGSS standards had an effect on science teaching self-efficacy. In-service elementary teachers' science self-efficacy significantly improved after undergoing formal training in unpacking science standards as indicated by the STEBI-A tool.

Kang, McCarthy, and Donovan (2019) used a modified version of the STEBI-A to measure the effectiveness of a professional development program focusing on enacting NGSS science and engineering practices. Significant gains were made in both STEBI subscales and in a modified scale measuring confidence in NGSS knowledge and skill. However, the sample size of this study was small (n=14).

Menon and Sadler (2016) used the STEBI-B with pre-service elementary teachers in a mixed-methods study to investigate the effectiveness of an elementary-specific content course. PSTE, STOE, and content knowledge all significantly improved. While this study occurred after the introduction of NGSS, it rarely focused on reform-based science teaching practices.

Mentzer, Czerniak, and Brooks (2017) used the STEBI-A with K-12 educators who participated in a three-year science leadership professional development program to implement project-based science practices. Results indicated that outcome expectancies remained high for the three-year period with no significant change. Science teaching efficacy significantly improved, but only after the third year in the program. The authors suggest that project-based science and other reform-oriented science practices may take time to learn no matter how long a professional development program lasts. This finding parallels the findings of Penuel and colleagues (2023), who found that teachers implementing OpenSciEd curriculum at the middle school level experienced higher comfortability with enacting reform-oriented science practices after a three-year period.

Sandholtz and Ringstaff (2014) used a mixed methods approach to evaluate the effectiveness of long-term professional development focused on science content, pedagogy, and teacher collaboration. The science content was delivered in multiple subject areas using scaffolded-guided inquiry to model the use of scientific inquiry. Mathematics and literacy tools were introduced as a means to develop science skills in analysis, modeling, and communication, which happen to be three of NGSS's science and engineering practices. The STEBI results indicated a significant increase in the PSTE subscale, STOE subscale, and overall self-efficacy. Teachers' self-reported use of hands-on science activities and their perception of increased student participation in inquiry practices was corroborated by classroom observations. The positive self-reports and observations were associated with higher science teaching self-efficacy.

Prior to NGSS, engineering had not been prioritized in science standards. Yesilyurt and colleagues (2021) used a mixed-methods approach in which they modified the STEBI-B by replacing the word “science” with “engineering” to focus on engineering teaching efficacy. This modification in measurement is consistent with recommendations from Tschannen-Moran, Hoy, and Hoy (1998), as self-efficacy is highly contextualized. The authors claim NGSS’s science and engineering practices clearly emphasize engineering practices, but not the nature of engineering. The authors taught an engineering unit in an undergraduate science methods course using hands-on, lecture, and literacy strategies. After the intervention, the personal teaching efficacy subscale showed significant improvement with a large effect size. The outcome expectancy subscale showed a significant improvement with a small effect size.

In an attempt to leverage the emphasis on literacy in elementary teacher preparation programs, a quantitative study by Love, Napoli, and Lee (2023) asked whether integrating science themes in a writing methods course for elementary preservice teachers would affect science teaching efficacy. The preservice teachers took the STEBI-B before and after instruction about writing and analyzing poetry with science themes. There was a slight significant difference in STOE, but no significant difference in PSTE. The authors speculated that the students found potential value in STEAM lessons, but were not confident in implementing such instruction.

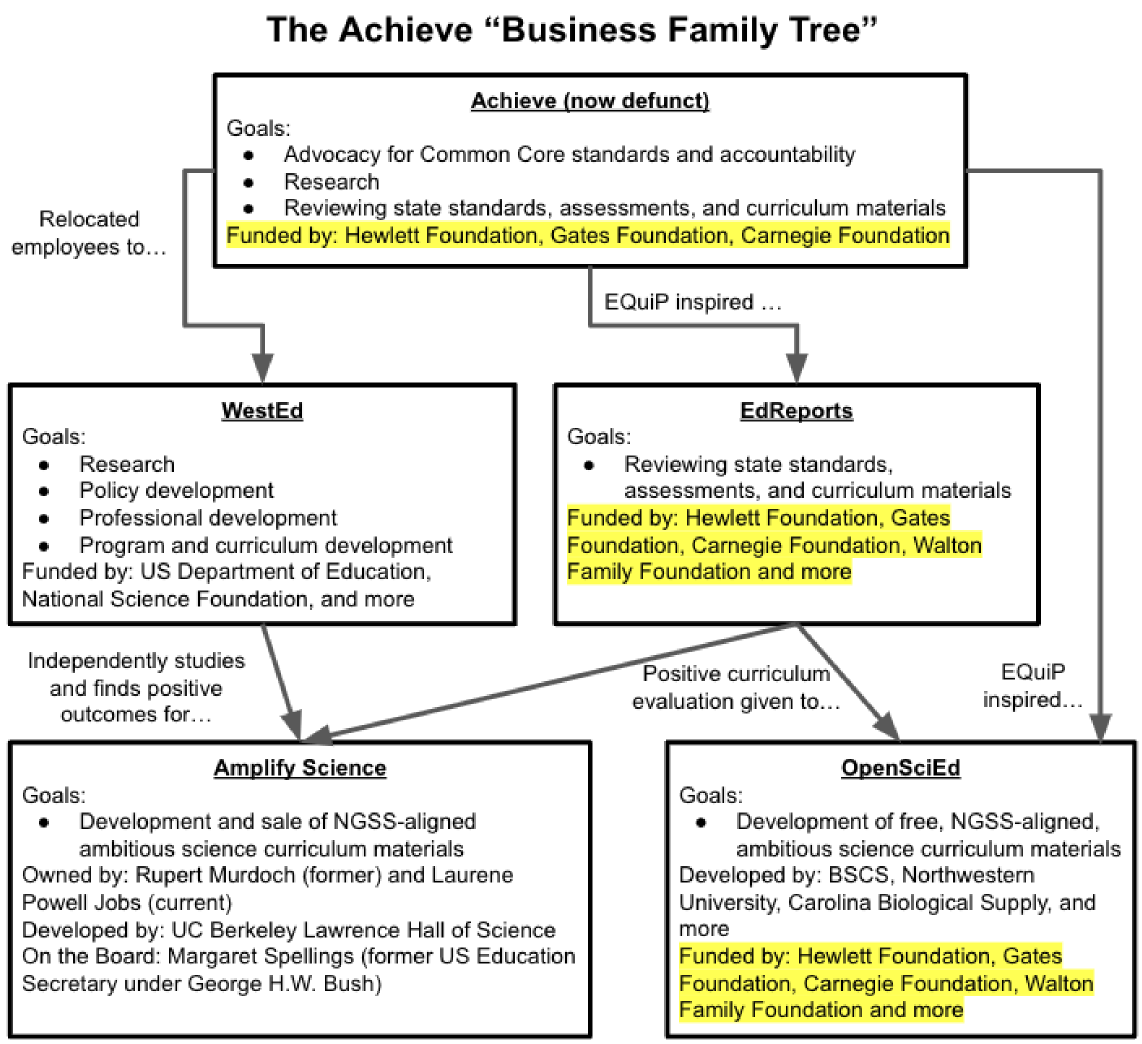
Government-based Incentives for Educational Entrepreneurship

George H. W. Bush’s national educational strategy (not to be confused with federal program or legislation) dubbed “America 2000” laid an inspirational foundation for the Clinton administration to pass an education law titled “Goals 2000” in 1994. Goals 2000 promised federal subsidies to states who vowed to create state standards and testing systems. Federal law prohibits any agency from directing or controlling curriculum and instruction in schools. Goals 2000 circumnavigated these precedents by “bribing” states into complying with the creation of new standards and testing batteries under the threat of withholding funding (Ravitch, 2020). President Bush signed No Child Left Behind (NCLB) in 2002 with broad bipartisan support. It mandated that 100 percent of students should score “proficient” on state tests by 2014.

Common Core, Inc., funded by Bill Gates, jumped at the opportunity to legitimize national literacy and math standards, creating the Common Core State Standards. These standards would be the foundation for their business model. The company worked to create curricular materials and offer workshops with initial financial success. NGSS would later incorporate and reference Common Core math and literacy standards. President Barack Obama and Secretary of Education Arne Duncan announced the Race to the Top Program in 2009, a competitive set of grants that rewarded states that agreed to adopt national K-12 standards among other stipulations. Only 18 states were awarded funding, but this didn’t prevent states from seeking guidance from educational entrepreneurs and their wealthy benefactors.

Amplify Science is the curriculum being used by teachers in this study. Before discussing Amplify Science, it is important to track its “business family tree” to reveal forthcoming conflicts of interest (as shown in Figure 2). Achieve was a company created by governors and business leaders to advocate for Common Core standards and offer consultancies for reviews of standards and assessments which states paid for in part with Race to the Top funds. Achieve is funded by both the William and Flora Hewlett Foundation and Bill and Melinda Gates Foundation (Achieve.org, n.d.).

Figure 2. The Achieve "Business Family Tree"



Achieve helped spark the creation of both EdReports and the previously mentioned OpenSciEd curriculum through their EQuiP Rubrics for Science (Cohen, 2020). EdReports functions like a Consumer Reports for Common Core in that it reviews and rates Common Core aligned curriculum materials. EdReports is supported by many of the same corporate and philanthropic sources as Achieve (EdReports.org, n.d.). Achieve closed its doors in 2020 and relocated most of their employees to WestEd (Cohen, 2020). WestEd conducts many services such as research, policy development, and professional development. WestEd is funded by a variety of sources, including the U.S. Department of Education and the National Science Foundation.

Amplify is a company that creates educational curricula and assessments, and Amplify Science is one of their products. Amplify has ownership ties formerly to Rupert Murdoch and currently to Laurene Powell Jobs. Former U.S. education secretary Margaret Spellings served under George H. W. Bush and is a member of Amplify's board. Spellings was a key proponent of the No Child Left Behind legislation in her role as education secretary.

Amplify Science is the only elementary science curriculum that meets all three "gateways" of the EdReports review process. The first two "gateways" evaluate whether a science curriculum is designed for NGSS and has a coherent design. The final EdReports "gateway" evaluates the usability of curricular materials. Amplify Science and OpenSciEd are the only middle-level curricular materials that meet EdReports's criteria. Amplify Science has had two external efficacy studies (Harris et al., 2022; Harris et al., 2023) conducted by WestEd, which found positive student outcomes.

Amplify Science and OpenSciEd's accolades from EdReports and WestEd are diminished in the light of their connection to Achieve. Most of the Achieve "business family tree" are wealthy philanthropists vying for government funded grants through advocacy of Common Core and its contemporaries. The connections show some troubling entanglements, mainly to the same group of billionaires who stand to make a dollar by creating the standards, legitimizing them through their own rating system, and creating the only "approved" materials (Ravitch, 2020).

Federal, state, and local education systems are often underfunded, and turning away generosity may seem counterintuitive when it comes to the important undertaking of educating the next generation of scientifically literate individuals. However, careful scrutiny of philanthropic motives is not unwarranted given the tax advantages and potential prosperity associated with educational entrepreneurship. The private philanthropy that has powerfully influenced public educational policies that established accountability systems may not similarly be held accountable for negative student outcomes. Understanding the genesis of commercial science curriculum may provide perspective when analyzing beneficial and detrimental educational outcomes.

Amplify Science Efficacy and Literature

Though there is a paucity of academic research surrounding the efficacy of Amplify Science, the company has conducted numerous internal studies. The data for these studies are typically shown in flier or brochure format with the source methods and measures left in relative obscurity. In one document (Amplify Science, 2019c), there are red flags such as graphs that show growth in student achievement, but not indicating statistical significance. In another document (Amplify Science, 2019a), pre and posttest comparisons are made, but the results only show effect sizes without stating statistical significance. One internally conducted efficacy study in Washington state showed positive growth that was not statistically significant (Amplify Science, 2019b).

Amplify Science was used in a study (McNeill et al., 2018) regarding fidelity to implementation of reform-oriented science curriculum. The authors wanted to know whether embedded educative features that supported teacher learning about structural and dialogic aspects of scientific argumentation would affect teachers' enactments of argumentation instruction. The authors found that as long as teachers were faithful to particular goals of the unit's lessons that focused on argumentation, teachers did not need to strictly maintain fidelity to the procedures prescribed by the curricular materials.

Teachers found the embedded educative materials were helpful to achieving fidelity to the goal of improving scientific argumentation.

Uppendahl (2020) wrote their master's thesis using action research to reflect on Amplify Science's impact on their fourth-grade students. Uppendahl found that the curriculum materials lacked repeated practice in written scientific argumentation, resulting in small gains between Amplify's pre and post assessments. However, student curiosity, engagement, and on-task behaviors increased throughout the unit.

An independent WestEd study (Harris et al., 2022) used a randomized controlled trial and found that student achievement on a three-dimensional science assessment significantly improved for students using Amplify Science compared to control schools. A majority of teachers in the Amplify Science group reported that students were more engaged in science. The teachers also reported that the curricular materials supported their science teaching and that their science teaching had changed as a result of using the curriculum.

A second WestEd study from Harris and colleagues (2023) used a randomized controlled trail to examine the impact of the Amplify Science program compared to regular classroom science instruction in first grade classrooms. The research team developed two NGSS-focused assessments since none were available. The Amplify Science condition performed significantly better ($p<.001$) than the control condition in the science learning assessment with an effect size of 0.24. The Amplify Science condition also performed significantly better ($p<.01$) than the control condition in the science-vocabulary-in-use assessment with an effect size of 0.46. The Amplify Science condition did not perform with any significant difference compared to the control condition in standardized science or reading assessments. The authors of this study framed the nonsignificant reading differences as a positive finding, stating that the findings should give teachers confidence that students' reading scores would not be negatively impacted due to literacy-focused science curriculum materials.

CHAPTER III: METHODOLOGY

Context, Recruitment, and Participants

Context

In 2020, the school district in the current study began a curriculum improvement process to determine viable science curriculum options that would align with Nebraska's science standards implementation plan. The curriculum specialist first conducted a needs assessment review resulting in a subsequent "formal program study." The school district initially planned to design their own science curriculum at the elementary level, supplementing it with open-source educational materials. Teachers were offered professional development to understand the pedagogical shifts posed by NGSS to prepare for the possibility of designing and enacting in-house curriculum.

Plans changed as a result of the COVID-19 pandemic, specifically due to an infusion of resources from federal legislation. The Coronavirus Aid, Relief, and Economic Security (CARES) Act, the Coronavirus Response and Relief Supplemental Appropriations Act (CRRSA), and the American Rescue Plan (ARP) Act established various Elementary and Secondary School Emergency Relief (ESSER) funds. Individual states made further stipulations for these funds. ESSER funds could be used to support, select, or purchase high-quality instructional materials. Spring and summer of 2021 were used to review science curricula that might be suitable for purchase.

The state department of education overseeing the school district in this study created an "instructional materials collaborative" that used EdReports to assess the quality existing curriculum. This was significant because the only curriculum that reached the threshold of “partially meeting” expectations for alignment using the EdReports rubric was Amplify Science. Ten curricula were considered, but ultimately one science curriculum was piloted in parallel with Amplify Science.

In anticipation for curriculum selection and implementation, all elementary teachers in the district participated in 1.75 hours of required training to introduce phenomena-based teaching during the summer of 2021. This training was not curriculum-specific. Twenty-four teachers from nine schools piloted Amplify Science during the 2021-2022 school year while thirty-one teachers from twelve schools piloted PhD Science. During the summer of 2022, members of the two pilot groups met to make recommendations for a 2022-2023 school year implementation. Amplify Science was unanimously recommended by piloting teachers for adoption. All K-5 teachers received 3.5 hours of training specific to implementing Amplify Science curriculum during the same summer. Teachers were supplied with physical materials, lesson guides, slideshows, assessments, and guidance for using the assessments.

In terms of continuing supports, an Amplify-employed regional curriculum coach comes to the school district two times every quarter to conduct cycles of observation and feedback. Observation and feedback sessions are by teacher-initiated invitation only, and according to the district science curriculum specialist, are not widely leveraged. Peer-led “office hours” for elementary teachers are available once per quarter. These “office hours” are voluntary, paid professional development opportunities. A teacher leader from within the district models unit planning and advises participants through discussion.

Recruitment and Participants

Purposive, snowball, and convenience sampling methods were utilized in this study. Participants of this study are third, fourth, and fifth grade elementary science teachers at the participating school district in this study. The participating school district provided a list of science curriculum contacts, who are teacher leaders in each elementary building that serve as liaisons between elementary science teachers and the district science curriculum department. An invitation to complete the initial quantitative survey was sent to all members of the list with a request to pass the survey along to the third-to-fifth grade science teachers in their building. There were approximately 300-400 third-to-fifth grade elementary teachers in the participating school district at the time. Only some of the teachers in that population taught science, depending on staffing and scheduling in their building.

The initial goal was to gather survey data from the first fifty participants that responded to include in the quantitative phase. The goal of fifty participants would have been representative enough to elicit a wide variety of STEBI-A scores for case selection purposes and account for the possibility of study attrition.

In this study, the results of the initial quantitative phase aided the purposeful identification of participants for the qualitative phase. Participants in the qualitative phase were a subset of individuals from the quantitative phase. The emergent nature of this study meant that neither the selection of participants nor the data collection methods for the qualitative phase were predictable at the outset of the study (Creswell & Plano Clark, 2018). However, cases were defined by selecting individuals who represented typical scores or extreme scores. This allowed me to ask questions that illuminated similarities and differences between the cases.

Procedures

Research Design

Mixed methods research involves the collection and analysis of both qualitative and quantitative data for the purpose of answering research questions and hypotheses. Integration is a key feature of mixed methods research in which the two forms of data and their results are mixed or combined. The data collection, analysis, and integrated data and results are organized into specific research designs which coincide with the logic and procedures of the study. Theory and philosophy serve to frame the procedures used within the study (Creswell & Plano Clark, 2018).

Mixed methods research allows researchers to explore information that is not accessible through a single approach alone. Qualitative data can provide a thorough understanding of a few individuals' experiences and can draw out previously unknown phenomena. Quantitative data takes information from a wider population and measures known variables. The associations of those variables can lead to inferences of causality, which can then be generalized to larger populations. The strengths of each type of research overlap in mixed methods research to limit the weaknesses of each, such as qualitative data's lack of generalizability, or quantitative data's inability to explore the perspectives of an individual at depth.

This study employs an explanatory sequential mixed methods study design. An explanatory sequential design begins with the collection and analysis of quantitative data.

This phase is followed by the collection and analysis of qualitative data. The intent of explanatory sequential method studies is to use the qualitative phase to explain the initial quantitative results. Figure 3 provides a procedural diagram of the current study. In this study, the quantitative phase may guide purposeful sampling for the qualitative phase.

Figure 3. Study Procedural Diagram

Phase	**Procedure**	**Product**
Quantitative Data Collection	• Demographic survey and STEBI-A	• Numeric Data
Quantitative Data Analysis	• Data screening (univariate)	• Descriptive statistics (mean, standard deviation)
Case Selection; Interview Protocol Development	• Purposefully selecting 1-3 participants from each group based on maximal variation principle • Developing interview questions	• Cases (N=3-5) • Interview protocol
QUALITATIVE Data Collection	• Individual in-depth interviews with participants • Documents	• Text data (interview transcripts, documents, artifact description) • Image data (photographs)
QUALITATIVE Data Analysis	• Coding and thematic analysis • Within-case and across-case theme development • Cross-thematic analysis • Qualitative software	• Visual model of multiple case analysis • Codes and themes • Similar and different theme categories • Cross-thematic matrix
Integration of the Quantitative and Qualitative Results	• Interpretation and explanation of the quantitative and qualitative results	• Discussion • Implications • Future research

Quantitative Data Collection

An initial quantitative survey was conducted to gather identifying information from the participants, which was later anonymized. Individual and school demographics were collected. Distribution of the survey was limited to emailing the district's 41 "science curriculum contacts", one at each elementary building, who were instructed to further distribute the survey to third-to-fifth grade science teachers. Questions eliciting a short description of prior training and experience with the newly implemented science curriculum were asked to provide further context. This survey can be found in Appendix B.

Following the demographics portion of the survey, participants also took the Science Teaching Efficacy Belief Instrument A (STEBI-A), which can be found in Appendix C. Riggs and Enochs (1990) developed this valid and reliable instrument to measure the science teaching efficacy of in-service science teachers. Twenty-five individual items used the Likert scale format in which respondents reported their agreement with statements by selecting one of five responses ranging from "strongly disagree" to "strongly agree."

The STEBI-A measures two constructs, each using their own subscale measure. One subscale consisting of thirteen items is the Personal Science Teaching Efficacy beliefs (PSTE), which measures the respondent's beliefs about their ability to deliver science teaching experiences. An example of a prompt from this subscale is "I am continually finding better ways to teach science." The other subscale consisting of twelve items is Science Teaching Outcome Expectancies (STOE), which measures the respondent's beliefs about the capacity of science teaching to overcome external factors germane to science learning. An example of a prompt from this subscale is "When the science grades of students improve, it is often due to their teacher having found a more effective teaching approach."

Upon item analysis for reliability, the Personal Science Teaching Efficacy subscale achieved an alpha of 0.92 and the Science Teaching Outcome Expectancy subscale produced and alpha of 0.77. Pearson r's for all criteria were assessed and were significantly and positively correlated with at least one scale (Riggs & Enochs, 1990).

Quantitative Data Analysis

Demographic data were used descriptively and information is displayed as tables in the results chapter. Total scores were summed and means were calculated from the entire STEBI-A instrument as well as each subscale of the STEBI-A. These means were rank-ordered and used to compare individuals to the group as a whole. The Likert format of the STEBI-A produces categorical, nonparametric data, meaning that relationships revealed by parametric tests such as correlations are ill-advised even if values are translated from categorical to numerical values.

The quantitative data from the STEBI-A aided the purposeful identification of participants for the qualitative phase. Participants in the qualitative phase were a subset of individuals from the quantitative phase. Cases were defined by selecting individuals who represented typical scores or extreme scores. Individuals scoring high on both subscales, low on both subscales, or high on one subscale and low on another subscale were to be randomly selected from individuals categorized as delineated above. An additional two individuals scoring at the median of both subscales were also sought. Table 2 provides an example of possible cases that were expected to emerge.

Table 2. Joint Display to Describe Purposive Sampling Based on Quantitative Results in an Explanatory Sequential Design.

Participant	PSTE Mean	STOE Mean	Case
Anna	1	1	Case 1: Low PSTE, Low STOE
Ben	2	2	
Carl	1	4	Case 2: Low PSTE, High STOE
Denise	2	5	
Edgar	4	1	Case 3: High PSTE, Low STOE
Frannie	5	2	
Greg	4	5	Case 4: High PSTE, High STOE
Harriet	5	5	
Ingrid	3	3	Case 5: Moderate PSTE and STOE
Jack	3	3	

Note. PSTE = personal science teaching efficacy; STOE = science teaching outcome expectancy. Each new color indicates how participants of each case could be purposively selected to participate in the qualitative follow-up phase

Qualitative Data Collection

A semi-structured interview was conducted using a university-licensed Zoom account to collect video/audio data. The interview used a protocol guideline (Appendix D) that included factual demographic information first to establish a comfortable, neutral rapport. Open-ended questions followed to provide the opportunity for the participants' perspectives to be heard as much as possible. Care was taken to avoid leading questions (Merriam & Tisdell, 2016). Some probes were pre-planned to allow the participants to elaborate on their responses from the quantitative phase, but other probes occurred naturally as unexpected or interesting comments emerged. All participants were asked the same three opening questions about Amplify Science, their professional development leading to implementation, and about justifying their quantitative ratings for science enjoyment.

Merging of the quantitative and qualitative databases was facilitated through the use of parallel questions during the qualitative phase of data collection. Parallel questions use the same concepts in each data collection phase so the separate databases can be readily compared (Creswell & Plano Clark, 2018). No two participants were asked the complete set of protocol questions. This decision was informed by instances when participants answered with “strongly agree” or “strongly disagree” in ways that either supported or ran counter to their STEBI-A subscale totals. The parallel questions of one participant whose scores were near the median were specifically chosen so they could elaborate on the many selections they made in the “uncertain” category of the STEBI-A instrument. However, all participants were asked at least four PSTE parallel questions and two STOE parallel questions. Table 3 provides examples of parallel questions from the PSTE subscale of the STEBI-A and the interview protocol. Table 4 provides examples of parallel questions from the STOE subscale of the STEBI-A and the interview protocol.

All participants were asked the same four parallel questions to get a representative view of their science teaching self-efficacy and outcome expectancy. Question 12 of the protocol asks, “Has your curriculum implementation experience changed your understanding of any of the science concepts you teach? If yes, how so?" This question probes the subject matter knowledge (SMK) of teachers in the study, which is strongly correlated to science teaching self-efficacy. Question 19 of the protocol asks, "What skills do you believe an effective science teacher needs?" This question is meant to elicit pedagogical content knowledge (PCK), which is also strongly correlated to science teaching self-efficacy. Question 2 of the protocol elicits the potential change in teachers' PCK by asking, "In what ways have your science teaching practices changed since the beginning of the curriculum implementation process?"

Question 11 of the protocol asks, "How would you compare the science achievement of your students before and after implementing Amplify Science?" This question was meant to gather a broad view of the outcome expectancy that teachers had of their students.

Table 3. Joint Display for Integration of Parallel Data Collection Questions: PSTE

Quantitative Scale (from the Science Teaching Efficacy Belief Instrument - A)	**Qualitative Questions/Probes** (from interview protocol)
Personal Science Teaching Efficacy	
2. “I am continually finding better ways to teach science.”	In what ways have your science teaching practices changed since the beginning of the curriculum implementation process?
3. “Even when I try hard, I don’t teach science as well as I do most subjects.”	How would you compare your science teaching ability to the other subjects you teach? Why?
8. “I generally teach science ineffectively.”	Describe the components of an ideally planned science lesson or unit. **Probe:** How does your own science teaching compare to what you described as the ideal science lesson?
12. “I understand science concepts well enough to be effective in teaching elementary science.”	Has your curriculum implementation experience changed your understanding of any of the science concepts you teach? How so?
19. “I wonder if I have the necessary skills to teach science."	What skills do you believe an effective science teacher needs?
21. “Given a choice, I would not invite the principal to evaluate my science teaching.”	**For low scores:** What kinds of reservations would you have with your principal evaluating your science teaching?
22. “When a student has difficulty understanding a science concept, I am usually at a loss as to how to help the student understand it better.”	What kinds of challenges has the new science curriculum posed to your ability to address students' understanding of science concepts?
23. “When teaching science, I usually welcome student questions.”	How do you respond when students ask questions during science class?
24. “I don't know what to do to turn students on to science.”	What kind of strategies do you use to build students’ interest in science? **Probe**: Is that a strategy that you’ve always used, or was it embedded in Amplify Science materials?

Table 4. Joint Display for Integration of Parallel Data Collection Questions: STOE

Quantitative Scale (from the Science Teaching Efficacy Belief Instrument - A)	**Qualitative Questions/Probes** (from interview protocol)
Science Teaching Outcome Expectancy	
1. "When a student does better than usual in science, it is often because the teacher exerted a little extra effort."	What impact, if any, does the effort you exert while planning and teaching science have on student achievement in science?
4. "When the science grades of students improve, it is most often due to their teacher having found a more effective teaching approach."	
9. "The inadequacy of a student's science background can be overcome by good teaching."	Can you describe a time when a low-achieving student in science performed well due to your science teaching?
10. "The low science achievement of some students cannot generally be blamed on their teachers."	What do you think are the most important reasons that certain students may not be attaining high achievement in your science class compared to their achievement in other subject areas?
11. "When a low achieving child progresses in science, it is usually due to extra attention given by the teacher."	How would you compare the science achievement of your students before and after implementing Amplify Science?
16. "If parents comment that their child is showing more interest in science at school, it is probably due to the performance of the child's teacher."	Describe how teaching with Amplify Science units has (or has not) impacted your students' interest in science.
20. "Effectiveness in science teaching has little influence on the achievement of students with low motivation."	Describe how teaching with Amplify Science units has (or has not) impacted student motivation.

Qualitative Data Analysis

Written reports in this study used pseudonyms to protect the identities of participating districts and participants through the anonymization of data. Transcripts of interviews were read and initial/open coding occurred by hand (Merriam & Tisdell, 2016). Constant comparative analysis was used throughout to uncover themes from all participants in an inductive manner. For example, curriculum implementation documents and STEBI-A results were used as additional data points to substantiate statements made during interviews (Corbin & Strauss, 2015). A priori codes were considered, but purposely avoided. This decision was consistent with recommendations from Tschannen-Moran, Hoy, and Hoy (1998), as self-efficacy is highly contextualized. Code names came primarily from participant quotes. These are called *in vivo codes*. I interpreted and renamed codes as an organizational aid.

As patterns emerged, they were organized by categories or themes in an increasingly deductive manner. For example, upon initial/open coding, 476 codes were made. These codes were categorized into eight categories and re-grouped during a second round of analysis. 115 representative quotes from the four interviews were organized into these eight categories. Similar categories were subsumed under 3 final themes: knowledge of science content, practices, and curricular resources; limited time and competing demands; and student reactions.

Member checking of interview quotes strengthens confidence that information and contexts are accurate and that credibility within the qualitative phase of data analysis is maintained. Member checking also gives the participants the opportunity to contribute new perspectives that were not retrieved during the interview phases (Baxter & Jack, 2008). This process occurred after interviews were fully transcribed and themes were identified to strengthen validity.

Mixed Methods Interpretation: Data Integration

Mixed method studies are not simply separate quantitative and qualitative studies about the same topic. The quality of inferences drawn from separate strands can be enhanced through the process of integrating the data from each strand into a single interpretation. The intent of integration in an explanatory sequential design is to use the results of the qualitative strand to provide an explanation of results from the quantitative phase. The joint display in Table 5 below provides a useful way to organize the qualitative data, providing further insight, nuance, and explanation about the quantitative database (Creswell & Plano Clark, 2018).

Table 5. Joint Display for Representing Integrated Results

	Low Science Teaching Outcome Expectancy	**High Science Teaching Outcome Expectancy**
Low Personal Science Teaching Efficacy Beliefs	Theme A *"Quote(s) from case 1."* Theme B *"Quote(s) from case 1."*	Theme A *"Quote(s) from case 2."* Theme B *"Quote(s) from case 2."*
High Personal Science Teaching Efficacy Beliefs	Theme A *"Quote(s) from case 3."* Theme B *"Quote(s) from case 3."*	Theme A *"Quote(s) from case 4."* Theme B *"Quote(s) from case 4."*

Researcher Position and Bias

Researchers cannot escape the social world they study. Reflecting on their own experiences reveals a connection between the writer and their subject. My reflexivity on my experience as a science teacher is essential in maintaining the integrity of this case study (Hatch, 2002). Researcher reflexivity and positioning provides the reader with a sense that the researcher is credible, and to prevent unconscious biases (Merriam & Tisdell, 2016).

My undergraduate training occurred just before the NGSS era and science education was presented as guided inquiry. I had experiences using the 5E instructional model (Bybee et al., 2006), which served me well in the first years of my teaching experience. I had little guidance, but nearly limitless autonomy bounded only by the state's science standards and my district's suggested pacing guide. I used this flexibility to hone my skills in Socratic dialogue that elicited student understanding and misconceptions. I was able to teach in ways that allowed students to develop their own questions and develop controlled experiments.

As part of an in-service training in 2013 I was trained how to read NGSS standards and introduced to the instructional shifts posed by NGSS. The state I taught in adopted state standards which were a modified version of NGSS in 2017. In an attempt to respond to the arrival of new state standards, our district implemented "in-house" curriculum, which consisted of teacher-designed 5E lesson plans during the 2018-2019 school year.

The “in-house” curriculum was short lived. Opportunities to use commercially-available science curriculum that claimed to be NGSS-aligned were offered to schools as part of the district’s continuous curriculum improvement process. I was part of a pilot group that was provided access to Amplify Science 6-8 curricular materials for 7th grade science from 2019 to 2021.

I had an overall positive experience with Amplify Science, and I developed professionally as a result of its educative materials and through my enactment experiences. I gained a wider awareness of NGSS through Amplify Science’s educative materials that explicitly taught both teachers and students the definition and use of science and engineering practices (SEPs) and crosscutting concepts (CCCs). I noticed that students developed reasoning skills and were able to communicate scientific reasoning through written argumentation more consistently and with higher quality than with other curricular materials I had previously used. Physical materials and assessments were provided by Amplify Science, the digital platform was user-friendly, and the digital simulations were high quality.

While the digital simulations allowed for some investigative potential, I noticed a distinct lack of student-initiated, student-planned, and student-led investigations. This was unsettling given my background in guided inquiry. It was also strange knowing that Amplify Science claimed to be NGSS-aligned, but was only thinly including the NGSS SEP of “planning and carrying out investigations.” Opportunities for eliciting students’ prior knowledge seemed sparse as well, which is important, because prior knowledge is the foundation for continued learning (Rumelhart, 1980).

Following my experience in teaching using Amplify Science, I was invited to take part in my district's science curriculum selection committee. Teachers used the EdReports Review Tool as a way to rate the quality and NGSS alignment of the "in-house" curriculum, Amplify Science, and OpenSciEd. While this process was completed independently by teachers based on their enactment experiences, a formal review of Amplify Science and OpenSciEd was already publicly available on the EdReports website and made known to our curriculum reviewers. The findings of the district committee matched those of the EdReports reviews, with OpenSciEd holding a rating slightly higher than that of Amplify Science. I have been teaching using OpenSciEd, another NGSS-aligned curriculum, from 2021 to 2024.

This case study presents a small risk of becoming a "*backyard study*," that is, a study of one's own organization. I am an employee of the school district from which the participant sample will be taken. Studying in one's own organization can lead to complicated politics, power imbalances, participant risk, and overall negative influence. However, I am neither an elementary teacher nor a supervisor of elementary teachers, which limits ties to the participants and host schools that would impact participant recruitment and data collection. Multiple strategies of validation will be used in this study to make sure the accounts within are both accurate and to provide insightful learnings (Creswell & Poth, 2018).

Limitations

In general, case study is criticized for its lack of representativeness and generalizability. There are very few proposed participants in the qualitative case study phase of this study, and it may be easy for readers to misunderstand or underestimate the value of unique cases and their rich descriptions. The sample in this study was not randomly selected. This study used a convenience sample, limiting participation only to those who chose to opt in. The part of the population who chose not to participate may have had drastically different experiences than the participants who opted into this study. This presents the possibility of non-response bias. Furthermore, I am relying on my sample to be honest, cooperative individuals. I have chosen a quantitative tool with high validity and reliability, but response biases such as courtesy and social desirability biases are possible in surveys and interviews.

Another criticism of case study is that it is a target for subjective bias. I have carefully considered my interactions with NGSS-aligned curriculum in my positioning statement. I aim to take as measured and objective of a stance on these topics as possible, but realize that my experiences and values may shape the way I view the reality of the participants of this study.

IRB and ethical considerations

The study design was submitted to the University of Nebraska's Internal Review Board (IRB) for approval in January of 2024. After temporary approval was granted, an IRB submission was made to the participating school district. In February of 2024, the participating school district approved the IRB submission and the research project gained certification of exemption. Prospective participants received a study invitation via email and provided their informed consent through a Qualtrics survey. All participation in the study was voluntary and took place outside of the participants' contracted working hours.

Participants were offered an opportunity to receive a $10 virtual Visa eGift card via raffle for completing the initial quantitative survey. Incentives were managed by a university Virtual Incentives program. The odds of winning were at least 1 in 10. I took the participants' code key numbers and entered them into a choice randomizer online. The first 5 (or more) coded numbers selected were cross referenced with the code key in order to notify winners via email after the completion of the interview phase of the study.

Participants' survey responses were stored on a password protected device only available to me. Data access was only available to me and my faculty advisor via a shared folder on my university-licensed OneDrive account. Each survey respondent was assigned a numerical code. The participants selected for interview were only linked by a code key which was secured and not shared. The code key was destroyed upon completion of the dissertation defense. Interview responses were transcribed, the audio/video recording was destroyed, and transcripts were stored on the same password protected device.

Written reports in this study used pseudonyms to protect the identities of participating districts and participants through the anonymization of data. Research results are reported for the quantitative phase in aggregate form and qualitative results (i.e., interview quotes) are reported using an appropriate pseudonym. Individual-level data could be shared in future manuscripts being submitted for publication or presentation, but such data would be de-identified.

Participant burden occurs when extensive amounts of effort and time are required by a study. The participants in the qualitative phase of this study will be a subset of the quantitative phase, meaning participants will take a survey and participate in an interview. The burden is moderate in this study and care will need to be taken to make data collection opportunities more flexible, convenient, or distributed.

CHAPTER IV: RESULTS OF THE STUDY

In this chapter, I describe the data collected in this study. I start by describing a shift in methodology in response to low participant response rates. I then describe the demographics and the science teaching background of the teachers in the study. I also describe the results of the STEBI-A section of the initial quantitative survey. The qualitative results are organized by case, theme, and supporting quotes.

Shifting Methodology

The study was conducted with third to fifth grade science teachers in the same urban school district. This non-probability sample consisted of 7 educators. The sample size fell short of the initial goal of 50 survey responses. Rather than being provided the contact information for all third-to-fifth grade science teachers in the district, I was only provided the contact information for 41 "science curriculum contacts."

The science curriculum contact is a science teacher in each elementary building that is responsible for attending quarterly meetings to learn about science curriculum information, news and updates, and professional development opportunities. The curriculum contact generally participates in abbreviated professional development sessions during these meetings. The curriculum contact then relays the meeting minutes to the rest of their building.

I was invited by the district's science curriculum specialist to personally present the study's informed consent at a quarterly science curriculum contact meeting and to field study-related questions. I sent the formal study invitation and informed consent documents to the curriculum contacts and trusted that they would faithfully distribute the information to the target participants in their building. After a week receiving no survey responses, I sent a reminder message to curriculum contacts with clarifications about distributing the study. A third and final reminder message was sent to curriculum contacts a week after that, just prior to the district's spring break. At this point, one survey response had been received over three weeks.

I changed my recruitment strategy recognizing that either the current distribution model was ineffective or that teachers were potentially overwhelmed with their primary responsibilities. I reached out to the district's director of elementary education, the supervisor of assessment and evaluation, and the science curriculum specialist for other avenues of survey distribution such as acquiring a listserv for the target population or reaching out to principals in-person or via email. These requests were denied, and I was met with the advice to wait patiently for responses or to continue reaching out to elementary teachers.

Through "snowball sampling," I attempted to recruit individual elementary teachers through acquaintances. Nine acquaintances referred the contact information of twelve elementary teachers so I could send personalized recruitment messages. Within two weeks after beginning the snowball sampling process, I had received the seven total quantitative survey responses in this study. Only four of the seven survey respondents agreed to a qualitative interview. While the planned explanatory sequential mixed methods design was still used, a much heavier emphasis was placed on the qualitative phase resulting in a study design reminiscent of a qualitative multiple case study design.

Quantitative Results

Demographics

The study was conducted with third to fifth grade science teachers in the same urban school district. This non-probability sample consisted of 7 educators. An overwhelming majority of the participants in this study (85.7%) identified as White/Caucasian females. Teachers from each grade level from third to fifth grade responded. All participants have been teaching in a K-12 setting for more than ten years with a range of K-12 science teaching experience from one year to 28 years. With respect to educational level, the majority (71.4%) had a master's degree, followed by bachelor's degrees (28.6%). Demographic results are summarized in Table 6.

Table 6. Demographics of Participants

Variable	n=7	%
Race/Ethnicity		
White/Caucasian	6	85.7
I prefer not to answer	1	14.3
Gender Identity		
Female	6	85.7
I prefer not to answer	1	14.3
Teaching Grade Level		
3rd Grade	1	14.3
4th Grade	3	50.0
5th Grade	3	33.3
Years Teaching in K-12		
1-5 years	0	0
6-10 years	2	28.6
11-15 years	3	42.8
16-30 years	1	14.3
31 or more years	1	14.3
Years Teaching Science in K-12		
1-5 years	2	28.6
6-10 years	2	28.6
11-15 years	1	14.3
16 or more years	2	28.6
Highest Degree Received		
Bachelor	2	28.6
Master	5	71.4

All of the teachers in this study are from the same midwestern urban/suburban school district. The school district in this study has 41 elementary schools. Tables 7 and 8 summarize the district's demographics for the 2022-2023 school year. Table 7 shows the district's size with respect to the number of teachers and students. Table 8 illustrates the demographics of the district's students by race and ethnicity as shown by data provided by the state's Department of Education.

Table 7. District Demographics

Number of Teachers	**Number of Students**	**Free/ Reduced Lunch (%)**	**English Language Learners (%)**	**High Ability Learners (%)**	**Special Education (%)**	**Graduation Rate (%)**
3045	41850	47	7	18	17	82

Table 8. District Student Demographics by Race and Ethnicity

American Indian or Alaska Native	**Asian**	**Black or African American**	**Hispanic or Latino**	**Native Hawaiian or Other Pacific Islander**	**White**	**Two or More Races**
256	2003	3057	6615	39	26018	3862

Professional Development

Participants in the quantitative survey were asked about their prior professional development experiences regarding Amplify Science curricular materials, including who lead the development sessions, length of the training, topics covered, and skills learned. Five of the participants mentioned joint training sessions from the school district and Amplify Science representatives in the summer prior to implementing the curriculum. One participant stated that their training came from their building science liaison. Another participant described their involvement in joint efforts between the school district and Amplify Science to "tweak" the curriculum. This same participant helped facilitate training on anchoring phenomenon routines to district elementary teachers.

None of the teachers reported who led the development sessions by name, nor did they describe the length of the training or skills learned.

STEBI-A Results

The study explored elementary teachers' science self-efficacy beliefs and outcome expectancies (PSTE and STOE) during implementation of an NGSS-aligned curriculum. Descriptive statistics for self-efficacy subscales and total scores are shown in Table 9.

Table 9. Descriptive Statistics from STEBI-A

Variable	Mean	SD	Min	Max
PSTE	48.71	0.73	43	52
STOE	38.57	0.87	34	47
STEBI-A	87.26	0.84	79	94

Note: Maximum possible scores: PSTE = 65, STOE = 60, STEBI-A = 125

Case Selection

Cases were selected by virtue of the ranking participants' total scores from the STEBI-A subscales. Originally, selection for the qualitative phase would have ideally involved individuals scoring high on both subscales, low on both subscales, or high on one subscale and low on another subscale. Although seven individuals participated during the quantitative phase with some of them fitting well into each pre-defined case designation, only four of the seven responded to requests for qualitative interviews. The remaining individuals' STEBI-A scores were ranked and new case designations were made based on a continuum of low to high PSTE and STOE subscale scores. Participant names have been pseudonymized. Rank-ordered case designations can be found in Table 10.

Table 10. Rank-Ordered Case Designations

Participant (Pseudonym)	STEBI-A Total	PSTE Total	STOE Total	Case Designation
Amanda	94	52	42	Case 1: High PSTE, High STOE
Sarah	88	49	39	Case 2: Moderate PSTE, Moderate STOE
Rachel	82	47	35	Case 3: Moderate PSTE, Low STOE
Emily	79	43	36	Case 4: Low PSTE, Low STOE

All four teachers' STEBI-A total scores and PSTE subscale totals were higher than the expected median of the STEBI-A instrument (Total: 75, PSTE: 39). Three of the four teachers' STOE total scores were at or above the expected median (STOE: 36). Knowing that the subset of participants chosen for each case had average or higher STEBI-A scores compared to the instrument's expected medians, their STEBI-A scores were ranked relative to each other. For example, Sarah's case designation of "moderate" was based on the fact that her STEBI-A total and subscale totals were nearly identical to the study's population mean for the STEBI-A total and subscale scores.

All participants in the quantitative survey were asked demographic questions. The individual breakdown of these demographics by case can be seen in Table 11.

Table 11. Individual Demographics by Case

Participant (Pseudonym)	Race/ Ethnicity	Gender Identity	Grade Level	Years Teaching in K-12	Years Teaching in K-12 Science	Highest Degree Received
Amanda	White/ Caucasian	Female	3rd Grade	10	10	Master's degree
Sarah	White/ Caucasian	Female	5th Grade	14	8	Bachelor's degree
Rachel	White/ Caucasian	Female	5th Grade	13	1	Master's degree
Emily	Preferred not to answer	Preferred not to answer	4th Grade	13	13	Master's degree

Teachers from each of the three grade levels sampled in this study were represented. Sarah is the only teacher whose highest degree was a Bachelor's degree of the four cases. All of the teachers had ten or more years of general K-12 teaching experience, but Rachel was unique in that this was their first year of science teaching.

All participants in the quantitative survey were also asked questions about their science teaching background. The individual breakdown of these data are shown in Table 12.

Table 12. Science Teaching Background

Participant (Pseudonym)	Years Teaching in K-12 Science	"Do you enjoy science?"	Average class size for science classes	Daily minutes spent teaching science	Amplify Science units taught
Amanda	10	5	21-25	35 minutes/day	Balancing Forces; Environments and Survival; Weather and Climate
Sarah	8	5	16-20	55 minutes/day	Modeling Matter; The Earth System; Ecosystem Restoration
Rachel	1	2	26-30	55 minutes/day	Modeling Matter; The Earth System; Ecosystem Restoration
Emily	13	2	16-20	45 minutes/day	Energy Conversions; Waves, Energy, and Information; Earth's Features

Note. "Do you enjoy science?" was on a scale of 1 to 5, with "5" meaning "I love science" and "1" meaning "Not at all."

Finally, all participants in the quantitative survey were asked to provide the name of their school. Table 13 illustrates the demographics of each school's students by race and ethnicity as they are shown in data provided by the state's Department of Education for the 2022-2023 school year. The state that this school district belongs to conducts one science content assessment in all of the K-5 range, occurring during students' fifth grade year. The percentage of students who reached proficiency in that assessment is listed for each school.

Title I is a federal education program that supports low-income students through additional funding. Schools can voluntarily apply for Title I assistance if 40% or more of their students are eligible for free or reduced-price lunches, which is a proxy determinant of poverty. It is notable that there was a strong positive association between free and reduced lunch rates and English language learners ($R = 0.987$, $R^2 = 0.974$) and a strong negative association between free and reduced lunch rates and high ability learners ($R = -0.966$, $R^2 = 0.933$) and state science assessment proficiency ($R = -0.983$, $R^2 = 0.966$).

Table 13. School Demographics by Case

Participant (Pseudonym)	Number of Students	Number of Teachers	Free and Reduced Lunch Rates (%)	English Language Learners (%)	High Ability Learners (%)	Title I	5th Grade State Science Assessment Proficiency (%)
Amanda	264	23	21	*	16	No	85
Sarah	384	40	**	25	8	Yes	53
Rachel	762	48	17	6	15	No	92
Emily	370	40	**	29	5	Yes	60

Note: * = No data from 2022-2023 school year, ** = 81% in 2021-2022 school year (most recent data)

Qualitative Results

After the conclusion of the initial quantitative phase, qualitative data in the form of semi-structured interview transcripts were collected and analyzed to better understand the quantitative findings. As a result of the limited quantitative data, the qualitative phase was heavily emphasized. The research question guiding this portion of the study was: How do elementary teachers describe their science teaching self-efficacy and outcome expectancy during their curriculum implementation experiences?

Case 1: Amanda (High PSTE, High STOE)

Demographics and Background

Amanda is a third-grade teacher who has been teaching elementary school for 10 years. She holds a master's degree and came into teaching as a second career. Her total STEBI-A score was highest among all of the participants in the quantitative phase of the study and tied for highest score in the PSTE subscale, which measures science teaching self-efficacy. Her STOE subscale score, which measures outcome expectancy, was the second highest of all participants in the quantitative phase. She had the highest STEBI-A total score and subscale scores of all four cases.

Theme 1: Knowledge of Science Content, Practices, and Curricular Resources

Professional Development

A series of events led to professional development and leadership opportunities that were unique to Amanda. At a previous elementary school within the district, Amanda was the science curriculum contact. During her time at this school, she had experienced two science curricular changes. Her regular contact with the district's science curriculum specialist led to an opportunity to join the district's science curriculum selection committee. As a committee member, she took part in early meetings where teachers learned about NGSS, 3-dimensional learning, and phenomena-based lessons because they were originally planning on writing their own science units. Teachers had the opportunity to make a list of "non-negotiables" for science curriculum before taking part in one of two curriculum pilot groups. Amanda was part of the Amplify Science group consisting of nine schools and 24 teachers. She was afforded Amplify-specific training unique to this pilot group.

After the curriculum selection committee used rubrics and written recommendations to rate the two candidate curricula, Amplify Science emerged victorious by unanimous decision. The curriculum was selected, but Amanda's opportunities for professional growth did not end there. She took part in modifying the curriculum's scope to fit the budgeted time elementary schools had to teach science each year. She described the impact this experience had as the full curriculum implementation was underway: "I was helping write and tweak things, so I think I had that advantage of kind of helping me to going into it, but …I feel like I'm constantly learning."

Amanda and a peer were asked to present professional development to all of the district's third grade teachers so they could implement an abridged, four-unit version of the Amplify Science curriculum. She led six sessions, each an hour and 45 minute long, in which she modelled a lesson and familiarized teachers with the curricular materials. Amanda reflected on her own professional development experiences stating, "I feel like I've gotten a lot of great opportunities when I've worked with the curriculum to really become familiar with it and become more comfortable so that I enjoy teaching it even more."

Enjoyment of Science

Amanda rated her enjoyment of science as a 5 out of 5, equating to the phrase "I love science." This wasn't always the case for Amanda.

> I think honestly, I would have said before I started teaching, I probably would have put that closer to a 2, and just because I think when I first started teaching I was not as comfortable with science. And so when I got put as a science

> curriculum contact for my building I was a bit terrified because I didn't feel like that was probably my strongest area of teaching…

Amanda attributed her initial science enjoyment and efficacy to some of her earlier educational experiences in science, stating that "...when I went to school, it was more just memorizing…instead of… figuring things out and how things work on your own." Amanda describes her low science teaching self-efficacy at this stage in leading to her career in science teaching stating "...I just didn't feel confident in my knowledge of the area to be able to pass it on to students. I would say just more discomfort and feeling like I didn't know as much to begin with."

"Gradually" is how Amanda described the change in her science teaching practices. She feels that her science teaching practices have

> progressed and gotten better each year… but I feel like now is most definitely, the most of where students are really figuring things out on their own…for a while I was saying 'This is the concept. Let me show you how it works,' instead of letting them look at something or do something and then figuring out why it worked that way.

Major turning points in building her science teaching self-efficacy came from additional science learning and teaching as well as feedback she received from students: "…the more that I learn about [science] and the more that I've done it and seen how engaged students are, then it really has become one of my favorite things to teach."

NGSS Knowledge and Strategies

As a result of Amanda's work in the curriculum piloting and selection process, she was mostly fluent in NGSS terminology. The rationales she provided about her instructional practices demonstrated that she understood the depth of 3-dimensional science instruction. In subsequent sections, you'll see specific examples of how Amanda's understanding of anchoring phenomenon routines, the storyline model of instruction, and curricular coherence benefit her students.

While Amanda says she uses generally sound instructional strategies such as intentional peer and group work throughout all of the courses she teaches, she also mentioned the specific integration of science and engineering practices (SEPs). For example, she said that the curriculum "requires them to do a lot more modeling." Modeling is used to bridge the gap between the concrete and the abstract.

A major aspect of Amanda's view of science teaching that aligns with NGSS tenets is that she values students' equitable sensemaking. Amanda mentions variations of the phrase "figure out" 22 times, and she uses the phrase "hands-on" to describe instruction five times. Amanda puts an emphasis on using "hands-on" learning to aid students' ability to "figure out" how the natural world works rather than being lectured.

> …kids are capable of figuring those things out…I feel like science takes away a lot of the barriers that some kids have when in math or reading or writing, in part because of their engagement, and in part just because they can talk about things, and figure things out with their hands, and with modeling. I think that that helps a tremendous amount.

Curricular Supports

Amanda's unique experience in the curriculum selection process led to a deep understanding of the rationale for the curriculum's materials. Her expertise didn't seem to necessitate support in the form of coaching or further professional development, nor did she mention district or outside resources she could access if she sought assistance. She also talked more about beliefs and practices than the tangible aspects of the curricular materials. Amanda did describe her use of the embedded videos of lab demonstrations and student investigations, but this was in the context of the occasional instances where limited time led to the necessity of skipping student-led investigations.

Practice Leading to Comfort

With respect to the current curriculum implementation, Amanda reported that she gradually "felt way more comfortable this year with those units that I didn't know as much about." She demonstrated a growth mindset, saying "I feel like there's always things that [students] do or say that I'm making connections to... I feel like more like a student along with them a little bit more than the other subjects."

Theme 2: Limited Time and Competing Demands

Although Amanda brought up the district's miniscule time allotment for science instruction, she did not describe this as an insurmountable obstacle. Third grade teachers are allotted 35 minutes of science instruction per day for 15 days a quarter. She described the occasional instance of running out of time for "hands-on activities." She used embedded video demonstrations as a substitution for investigations in order to stay on track with the unit plan.

Amanda also mentioned not being able to immediately field all student questions. Students' science journals were used to log unanswered questions. Sometimes the questions would get answered as a result of future lessons or get resurfaced at a future date. Amanda communicated that the curriculum materials and her instructional strategies helped her manage short time, but stated that "I just wish we had more time for science. I feel like it is not near enough time focused on that subject."

Theme 3: Student Reactions

Student Interest and Engagement

In regards to Amanda's curricular design experience, she spoke of the role that anchoring phenomena provide in driving student motivation to ask scientific questions. The phrase "anchor" is fitting with Amanda's experience of curriculum implementation. Much like an actual anchor provides a vessel with stability from wind and water currents, Amanda sees the anchoring phenomenon of each unit as a piece of the curricular materials that focuses students' questioning ability. She says the way the units are written provides students with a starting point of "something to figure out. I think the curriculum does a really nice job of making them wonder about something."

> …anchoring phenomenon is really an interesting concept for them. I think they really buy into it and are interested and engaged in it and I feel like it does a good job of guiding them to have the right questions to try to answer to meet the objectives. I think that it's really effective all the way around.

Student Achievement

Amanda framed student achievement in terms of conceptual understanding rather than as grades or test scores. She attributes a change in students' science achievement to her recent curriculum implementation saying "I feel like they have a better conceptual understanding probably than what they had." In Amanda's career, she has taught through three different science curricular changes. She was able to make a comparison of students' conceptual understanding, stating "I do feel like it's progressed and gotten better as I've gone in my career, but I feel like now is most definitely, the most of where students are really figuring things out on their own."

She speculated why students' conceptual understanding has improved, saying Amplify Science "gives them a better structure to build on as they continue into older grades and getting into some deeper ideas." Amanda didn't provide examples that directly substantiated this claim. However, she elaborated on this inference by explaining the cross-curricular benefits she's seeing as students learn throughout the year in her own classroom. For example, she explained that students were making connections in a light and sound unit in their language arts class. She informed me that the district's language arts curriculum is Amplify CKLA. I had the following exchange with Amanda to understand if the cross-curricular connections were intentionally planned.

> **WS**: "With them (Amplify Science/Amplify CKLA) being part of the same company, do you see any similarities with how the English information is presented or structured?"

Amanda: "I do not actually. I kind of thought that I would, but, I do not feel like, like, no. I feel like they probably could have done a better job with it, but no, yeah, they do not."

Despite the lack of explicit attempts by the curriculum materials to mutually reinforce each other, Amanda's students found a way to make it happen.

In summary of Amanda's observations about her students' science learning, she made a comment about the link between her students' engagement and science efficacy stating "I love teaching science and the kids... in large part because that's something that students are always really interested in and can feel successful at."

Case 2: Sarah (Moderate PSTE, Moderate STOE)

Demographics and Background

Sarah is a fifth-grade teacher who has been teaching elementary school for 14 years and has been teaching science for 8 years. She holds a bachelor's degree and is a science curriculum contact for her Title I designated school. Her total STEBI-A score and PSTE subscale was ranked fourth out of all seven participants in the quantitative phase of the study, which measures science teaching self-efficacy. Her STOE subscale score was the third highest of all participants in the quantitative phase. She had the second highest STEBI-A total score and subscale scores of all four cases.

Theme 1: Knowledge of Science Content, Practices, and Curricular Resources

Professional Development

Sarah attended the district's professional development session during the summer prior to curriculum implementation. She also recalled that sessions were organized by grade level and that a facilitator modeled a lesson. She described the remaining time as "time to plan."

Sarah also added that she attended a separate professional development session where she learned about her grade level's district-developed science unit. One district-developed science unit was written at each grade level to address missing science standards due to time constraints posed by the school district and modifications made to the existing Amplify Science curriculum. One of the components of the training was examining the various curricular materials, including the teacher guide, student journals, videos, websites, and simulations. She characterized the district's efforts for unit development as "[trying] to make it as similar as they could" to Amplify Science units, giving examples of student activities such as analyzing graphical data, gathering evidence from videos, and creating models.

She was mostly positive about the district-developed science units except for in one facet, "Unfortunately they don't have the readers that I love so much." The "readers" are short science articles embedded within the Amplify Science curriculum. The digital version of the readers had audio support for multilingual learners and students with learning disabilities.

Enjoyment of Science

Sarah rated her enjoyment of science as a 5 out of 5, equating to the phrase "I love science." Sarah says that, "Students love the experiment side of science, and I love that side, too." She hosts a science club and puts on the school's science fair, each of which take a considerable amount of time and effort to coordinate. Many of her responses recount the joy she sees that her students have in science classes and extracurricular activities, which will be elaborated on shortly in the “student interest and engagement” section.

NGSS Knowledge and Strategies

Sarah shared experiences about her curriculum enactment that hints at a passing knowledge of NGSS instructional shifts. Without uttering the phrase “anchoring phenomenon” Sarah briefly pointed out her approval of Amplify Science’s “unit questions” that guide students through a storyline model of instruction. However, she didn’t elaborate on what she specifically liked about the unit questions.

Sarah was asked about her comfort with fielding unexpected science questions from students. She described how she models humility and vulnerability to her students by admitting that she may not know the answer to every question. She acknowledges student questions by posting them on her marker board. At times, they research the question together on the spot. Other times, Sarah encourages her students to research their questions after they finish an assignment.

The SEP of analyzing and interpreting data allows for natural opportunities to integrate mathematics skills in science, and Sarah noticed that her students’ use of measurement tools such as rulers in science class translated to connections being made in her mathematics class. She made this statement with the proviso that “the connections don't necessarily happen unless they're completely explicit.” Her students were resistant to the idea that math and science should coincide. Her message to students was that

"Math is very much a part of science, so if you like science, make room for some math too..."

Curricular Supports

Sarah described some of the resources available to teachers during curriculum implementation such as the teacher edition (or TE) of the unit guide. The drawback of this resource was that "The TE is very big… Finding the time to read… all that when you're planning all of the subjects can be a challenge." While I did not have access to the teacher unit guides, I did find in the district's resources that one fifth grade unit had a staggering 464 teacher presentation slides. Most slides had notes typical of teacher edition guides, such as "suggested teacher talk," "teacher actions," and "student responses."

Videos were an included curricular support that Sarah appreciated for multiple reasons. Sarah described their value as educative supports stating "If there's something you're struggling with and you're not sure how you're going to teach it… you can watch another teacher teach it." The videos were also used as a substitution when opportunities for student investigations got cut short. The reasons for this substitution will become clearer in the "limited time" section.

The topic of the curriculum's "readers" was touched upon early by Sarah.

> I love that they have different readers that support what we're learning and we do keep those books available for students who want to have, maybe during their personal reading time, are interested in science and want to learn. We let them

access those readers throughout the whole quarter, not just our 2 or 3 weeks with science.

Sarah was making attempts to extend students' opportunity to continue their science learning beyond the 3-week allotment of time. Her enthusiasm about the "readers" was also tied closely to cross-curricular connections in English language arts.

It's really helped students with their non-fiction texts. It helps us to use the table of contents and other text features. Nonfiction texts are usually ones that students struggle with, so it's nice to have an actual book to use when practicing using nonfiction text features. We're not reading the whole book, we're just finding the information that we need and it's teaching that skill as well. Some of them can go back in their books and say, "Oh, in my book it said…" and then they can use the text evidence to support their answer…

Practice Leading to Comfort

Sarah described her and her colleagues' growing comfort and satisfaction with the curriculum implementation process.

Now that we're getting more familiar with it, we love it. Our first Amplify unit was super overwhelming. There is [sic] tons of slides. We had no idea how we were going to fit everything in on time. The more familiar you get with the program I feel like it's very smooth.

Sarah said her rate of growth in comfortability in Amplify Science wasn't nearly as quick as it was when she enacted science curriculum in the past. This was mainly a function of her prior science teaching load, which will be elaborated on in the next section. However, she did note that she is able to notice areas where students were struggling, take note of those areas, and adjust instruction for the following year.

Theme 2: Limited Time and Competing Demands

"I wish we had more time, time, time. There's never gonna [sic] be enough time in the school day for everything that needs to get done."

Reflecting back on the beginning of her teaching career, Sarah shared that she used to teach four sections of science per day. When I asked her about this teaching arrangement, Sarah said she followed the lead of veteran staff members who had made an unsanctioned decision prior to her employment to nearly eliminate instructional time in social studies and health, with that instructional time being divided among and added to other prioritized classes. She reported that the veteran teachers felt compelled to focus on math, reading, and science to improve their school's standardized state testing scores. Students would rotate between teachers who would focus on one of the three "core" subjects.

Sarah found a silver lining in this arrangement, noting that extra time spent on testing preparation seemed to raise state test scores in science. Sarah equated "teaching the same lesson four times a day" to "basically getting four years of experience on a lesson in the same day." Her answer is a comparison to the current curriculum where the next opportunity to teach the same lesson happens during the following school year. Once the veteran teachers had moved on to teach at other schools, Sarah and her new team reached a consensus that they "needed to be focusing on all of the subjects." They reincorporated health and social studies classes again, reducing the instructional time spent on science to its original allotment.

During the current curriculum implementation, Sarah noted that time is at an even shorter supply than it was in the formative years of her teaching career. She described a compromise she makes when time runs low stating "the first thing that gets cut are the experiments." When pressed why student investigations were targeted to save time, she shared concerns with "teaching the expectations… of a successful science experiment" and "fostering those conversations that [students] could be having during the experiments." She felt that students' unacceptable behaviors during science investigations necessitated frequent reteaching, taking away from lab time as a result. She acknowledges that this compromise isn't ideal in stating "...I try not to cut those as much as possible, cause [sic] that's what you remember… watching a video I don't think is as effective."

In regards to momentum, Sarah lamented the negative impacts on student excitement due to the current 15-day allotment for science instruction per unit, stating "We're having great science conversations, and then we go a month or two without talking about science very much at all. Then that excitement builds up again and then it dwindles out."

Theme 3: Student Reactions

Student Interest and Engagement

Whereas many of the participants' interview responses oscillated between teacher-centric and student-centric responses, Sarah talked positively about her students' reactions to learning science in almost every interview prompt. She used variations of the word “excited” seven times in reference to students’ disposition toward learning science, which is more than all the other interviewed participants combined. She used variations of “interest” three times, “fun” twice, and “motivated” twice throughout her interview.

Sarah was noticing increased engagement in her science class. From “the surprise and the excitement on student’s faces when there's a reaction that happens that they're not expecting” to classroom discussions that “get heated” and where “kids get really involved,” Sarah had plenty to say about the fun she was having with her science students.

In regards to SMK and PCK, Sarah says “knowing the material well” has an influence on students’ interest in science and that being able to manage student investigations is important. In her experience:

> I know that lots of teachers shy away from the experiments in the hands-on things just because they're afraid of the behaviors that might come about when they're doing that. That's not necessarily a strength for a lot of elementary teachers.

She provided examples of unexpected management challenges such as deescalating meter stick sword fights and reminding students that lab materials are not edible. She understands that “reteaching just takes away from that experiment time.” Where the punishing outcomes of lost instructional time may deter her colleagues, Sarah feels that she has learned how to prevent these events by “teaching those expectations and holding [students] accountable.”

When initial lab experiences are successful, Sarah fully leverages student interest saying "students are motivated to have better behaviors when they know experiments are coming." She builds up anticipation for labs to the point that positive peer pressure is commonplace, saying she notices students "paying attention, prompting each other like, 'Hey, you gotta [sic] pay attention otherwise we can't experiment…'"

Sarah also appreciated how Amplify Science positions students as scientists at the outset of their units saying, "I love with Amplify Science that they turn the student into the ecologist or into the food scientist…I had no idea that's an actual job." She says "this really helps students get motivated, get interested, start asking questions about careers." Building a science identity has a positive impact on the science achievement of students (Starr et al., 2020), and it appears in this case that the educative supports in the curriculum materials may have made a difference. Sarah says of the curricular resources, "I've got all of the materials I need to guide students through that, so I feel like I'm pretty prepared." Separate from the curriculum, Sarah notices that her science class tends to "rein in" students whose interests may not lie with reading or math.

Student Achievement

Sarah had difficulty comparing the achievement of students before and after implementing Amplify Science, describing the new system of 3-dimensional assessment as "a completely different way of measuring the success of students." The school district has been working on piloting common 3D science assessments at the elementary level. Sarah noted that student participation in the pilot was required to inform the district's assessment revision process, but that recording the results of the assessment in the teacher gradebook was not similarly required. This assessment had the potential to at least partially inform Sarah's view of her students' science achievement had grading been mandatory.

As a fifth-grade teacher, standardized state science assessments were at the forefront of Sarah's mind, as teachers at her grade level are the only one to administer a standardized state science assessment at the elementary level. She recalled, "I know for a while we had done really well on standardized state testing. Our scores came back pretty high and I don't think it's been high again."

Her recollection was accurate. According to the state's Department of Education, in 2018-2019, 49% of students at her school were proficient on the state science assessment. After a two-year hiatus from testing during the COVID-19 pandemic, testing was resumed during the 2021-2022 school year, wherein 73% of students were proficient. At the end of the 2022-2023 school year that coincided with science curriculum implementation, only 53% of students were proficient in the state science assessment.

There are plenty of problems with this metric, many of which will be discussed in depth later. One misleading variable relevant to this metric of student achievement is that the state's science assessment model shifted from entirely multiple choice during the pre-pandemic era to a 3D assessment model in the post-pandemic era. This means there are two years of relevant data on state science achievement. Two data points does not constitute a reliable trend, nor does the comparison of different cohorts of children.

In contrast with Sarah's view of district and state assessment results, Sarah was in tune with her students' classroom grades and with the science skills they've been demonstrating. Sarah noticed gains in students' scientific writing, specifically with their integration of science vocabulary. "Amplify, more than our previous curriculums, give students the words that they need to have those deeper science responses where they're using their vocabulary." She was surprised to hear science vocabulary in use during classroom discussions stating, "…you hear them saying those words in conversations that you typically don't hear fifth graders having, which is great."

When asked to specifically compare the district's previous science curriculum to the current curriculum, Sarah was less enthusiastic about the results she was observing from the current implementation. "I would say we probably don't see as much success, but that's not necessarily due to the curriculum, but our ability to teach science as often as we used to." Recall that Sarah's science teaching schedule near the beginning of her career allocated more time for science instruction at the expense of a reduced or nonexistent health and social studies teaching load.

Sarah provided one final endorsement of Amplify Science with a caveat:

> I do feel like if all grade levels are teaching Amplify Science when they're supposed to be and if they're using it appropriately, correctly, how we're instructed to use it, I do feel like students will have a better understanding of science, but we won't see that for a few more years once they've gone through a few more grades of Amplify. I think with Amplify as our base we could go so far above and beyond what our old curriculum did for us.

She acknowledges that a cohort of students may need to go through the three consecutive years of the Amplify Science program to make meaningful inferences about its efficacy, provided that teachers use it as intended.

Case 3: Rachel (Moderate PSTE, Low STOE)

Demographics and Background

Rachel is a fifth-grade teacher who has been teaching at the elementary level for 13 years, but has only taught science for 1 year. They hold a master's degree and are a science curriculum contact for their school. Rachel was previously a special education teacher before they became a general education teacher. Their total STEBI-A score ranked sixth out of all seven participants in the quantitative phase of the study. Their PSTE subscale score, which measures science teaching self-efficacy, was ranked fifth of the seven participants. Their STOE subscale score was ranked sixth out of all seven participants in the quantitative phase. They had the third highest STEBI-A total score and PSTE subscale score, as well as the lowest STOE score of all four cases.

Theme 1: Knowledge of Science Content, Practices, and Curricular Resources

Professional Development

Rachel had their own unique circumstance leading to the implementation of new science curriculum. Rachel's school opted to temporarily organize their "unit studies" classes by specialist. "Unit studies" includes a quarterly rotation of classes such as science, social studies, health, and in some cases, social-emotional learning. Rachel taught 6 sections of social studies during the year Amplify Science was being implemented while a partner teacher taught the science sections. During the current school year, Rachel switched duties with her teaching partner and taught science for the first time in her career.

The following exchange occurred during Rachel's interview and explains how this teaching structure impacted Rachel's science teaching self-efficacy.

> **WS:** "In your initial survey, you described your prior experience and professional development/training with Amplify by answering "Science Liaison." Can you please clarify that response?"
>
> **Rachel:** "I would say I probably haven't really learned how to teach with Amplify Science. I mean, I went to the classes… summer of 2022 before we put Amplify into place. You know that was an afternoon or something."
>
> **WS:** "You went to the summer training in 2022, but you didn't teach science that whole school year?"
>
> **Rachel:** "No. It had no effect on me because I was going to teach social studies."
>
> **WS:** "So you haven't had any professional development opportunities to get coached?"
>
> **Rachel:** "No."
>
> **WS:** "Are there opportunities that you could find if you wanted to take advantage of something in the district... to get what you need out of [the science curriculum]?"
>
> **Rachel:** "I have no idea to be quite honest. I am teaching sciences here just because felt I couldn't represent our school's needs without having taught the curriculum… [My principal] asked me to be the science liaison for [my school]. I

> didn't think I could be the best science liaison unless I was actually teaching science."
>
> **WS:** "Was there any rhyme or reason to why you were asked to be the science curriculum contact?"
>
> **Rachel:** "No."

Near the end of the interview, Rachel brought up the professional development opportunities for the science curriculum implementation:

> Maybe I just wasn't aware of it. I just don't think that the professional learning piece was there and I know people are going to want more professional learning. With [Amplify] CKLA (the language arts curriculum), I mean, they hammered it down our throat this summer. You could go to so many classes and learning opportunities surrounding that. I just, I didn't feel were provided [for science]... if there are some [PD opportunities], I've never heard about them.

Rachel's perspective was that professional development opportunities were not well-communicated. Rachel described the impact of the timing of their PD experience stating, "It had no effect on me…" Any expected outcomes of the summer professional development may have been severely undercut by the massive delay between learning and applying the information from the session. Recall that there were some supports in place for optional continued learning opportunities throughout the year of curriculum implementation, but Rachel may not have been aware of these because they were only teaching social studies.

Enjoyment of Science

Rachel rated their enjoyment of science from the survey prompt “Do you enjoy science?” as a 2 out of 5, with a 1 on the survey scale representing “Not at all.” Upon further questioning, it was clear that Rachel interpreted the question as a rating of the quality of their teaching experience using Amplify Science. They immediately answered “I truly do love science.” Their mother was also a teacher and had an avid interest in space. Rachel remembers their mother’s model rocket and ownership of a piece of one of NASA’s space shuttles. Rachel clarified their survey response.

> I do love science and I love the thinking and the questions behind it and I would say right now. I don't feel... not the passion, but the... I don't feel like the kids are getting a whole lot out of it like I got doing hands-on…I mean the best time the kids have is when we're doing the hands-on… I feel that that’s lacking.

NGSS Knowledge and Strategies

Rachel’s belief that students learn through “hands-on” activities is firm and rooted in their own experience as a learner, with Rachel stating that “some of my best elementary, middle school memories come from being able to work through a question I had.” Rachel mentions “hands-on” six times in their interview, which is more than any other interviewed participant in this study. However, it was difficult to pin down Rachel’s definition of “hands-on.” At first the definition was vague, with Rachel stating “I do love that manipulation piece of it.” Then, “hands-on” became a strategy that evenly benefitted all content areas, with Rachel stating “Their best ideas come from when they get to use their hands, and that is true, I think, across any curriculum.”

Rachel eventually made it clear that engineering was an important part of their "hands-on" schema.

> ...the kids have the most desire to learn when they can do something with their hands. They built a water filtration system out of cups and tape, and there were questions and ideas coming out of their mouth that I hadn't heard all quarter.

While engineering is involved in all of the science and engineering practices (SEPs), it became clear with another clarification that Rachel was dismayed at the lack of three SEPs: defining problems, planning and carrying out investigations, and designing solutions. The critique of the curricular materials was that "there wasn't a lot of hands-on building activities. There were games and stuff we played, but not a lot of building, testing hypotheses, and things like that."

Interestingly, investigations conducted using simulated models and involving hypothesis testing were not similarly included in Rachel's definition of "hands-on." The variables of student interest and motivation were necessary to cross the threshold into Rachel's definition of "hands-on." This aspect of the "hands-on" definition will be elaborated on in the "student reactions" section.

One theme relevant to PSTE is how science teachers address students' science questions. Rachel had strategies for addressing students' curiosity and embraced it without appearing stressed about limited instructional time.

> I think it's important to know when to take those sidetracked moments… We spend time looking it up ourselves. You know, from different sites to compare

like how are they similar. You know, do they have any differences in that information?

Curricular Supports

Rachel appreciated the “readers” that are embedded into Amplify Science materials. Rachel used the “readers” with their students designated as English Language Learners by teaching them how to move the text into Google Translate. This strategy enabled students to read articles in their preferred language. While this use of the readers wasn’t necessarily an embedded curriculum support, it was a use of Rachel’s pedagogical skills that potentially carried over from their background in special education.

The unit and lesson questions respectively asked at the beginning of Amplify Science and the district-developed curriculum were a support that Rachel didn’t find much value in. Rachel described the unit questions as “surface level” and in need of “bulking up” to “keep [students’] attention.” While Rachel described the questions as “approachable by every student,” they also wanted questions to elicit “deeper level processes.” Rachel’s reservations about the curriculum’s use of unit questions will be elaborated on in the “student reactions” section.

On the other hand, Rachel did find use in what they called “guided questions,” which are educative features that provide teachers more detailed, scripted questions in the “notes” section of the presentation slides. At first, Rachel was not aware of this support, but “second and third quarter I knew they were there and I have been able to… change …which order it's presented, because I have those supports there.”

As for responding to student questions, Rachel felt less supported by the curriculum materials saying "the guide doesn't always go into great detail about some of those questions that [students] have… I don't always get the best answers from the teacher's guide." Rachel used search engines as an impromptu strategy to gathering evidence or scientific explanations on short notice.

Practice Leading to Comfort

A silver lining of Rachel's teaching schedule during their initial curriculum implementation is that they teach multiple sections of science.

> My first group of students get the 'not so great' me and then the second group gets the better me. The second time around, I'm just much more capable of guiding them... I think that's true with everything. You just have to do it and get used to it and comfortable…

Rachel reported focusing on her questioning strategies between class periods stating "My questions are better, I think, because I've really thought about it…I've asked the question before and it didn't sit right with them. So then, how can I rephrase it for the next time or the next day even?" Rachel's commitment to reflection and revision in this example is clear.

Rachel compared their science teaching ability to other content areas such as reading and mathematics stating "Oh, it's very poor. I feel like a lot of the times I'm learning along with the students." Rachel clarified that understanding science content wasn't a barrier for their teaching, but anticipating how their students might interpret directions or information was more of a challenge. "You can read the rubric, you can read the teacher's guide, but until you see the students learning it, you don't know exactly how it's going to come together.” Rachel describes their own questioning as improvisational, stating “I've been modifying things my whole career with math and reading, and it's a new curriculum to me so being able to modify it on the spot, that did not go well.”

Theme 2: Limited Time and Competing Demands

Rachel also had limited time to teach science, but was more cavalier about preparing for science class, stating “I do plan [science questions], it just is lowest on the list of things.” One of the things on Rachel’s hypothetical list was their English language arts curriculum which was being implemented during the same school year that Rachel was also enacting the new science curriculum for the first time. “I'll be real [sic] honest, science is the last thing I think about. Every day with everything else that, you know, CKLA and things like that…”

Theme 3: Student Reactions

Student Interest and Engagement

While on the topic of frustrations with curriculum implementation, Rachel circled back to the perceived lack of “hands-on” activities, stating “The biggest frustration is the days that they just have to sit and listen.” From Rachel’s perspective, the curriculum’s weak support of student-driven investigations made a negative impact on her students’ interest in the district science fair stating, “we had maybe 25 students sign up at [my school], where in the past we had over 100 students sign up. We’ve taken out that ‘test, learn, retest, and learn’ type of thinking.”

Rachel objected to the use of instructional time spent copying down vocabulary terms, saying “that’s not the science I grew up with.” Amplify Science’s simulations were mentioned in the same vein as copying vocabulary. Rachel reports that, occasionally, students “really, really enjoy” the embedded simulations, but that the repetition of lessons involving the simulations eroded student motivation.

Rachel was critical about the curriculum’s “chapter questions.” The chapter questions introduce the anchoring phenomenon for a series of three to five lessons. For example, the first question of the district-developed “Space and Sky Patterns” unit is “What patterns do we see in the stars and constellations of the night sky over time?” The questions are meant to drive student motivation and think about their next investigative steps. In Rachel’s opinion, these questions caused students to answer without putting in the necessary effort and scientific reasoning.

> Once they've answered it, they don't want to think deeper into why. Why is it that the constellations we're seeing are different than the constellations that they're seeing in Australia? … they don't want to continue that conversation…. They don't always put forth the effort that they might when they're given those more higher-level thinking questions.

Rachel wants an answer from students rather than a claim. Claims are testable, and Rachel is noticing that the chapter questions don’t automatically serve their purpose of spurring students into action, which in this case would be testing the veracity of their claims. Rachel’s concern that the questions need “bulked up” do not account for the rationale of anchoring questions. The teacher shouldn’t have to “bulk up” questions.

Instead, the students should be asking questions about data that will need collected, how it will be collected, and what different patterns in the data may mean. In this way, students will be doing the cognitive heavy lifting. Science teachers must guide this process, building a classroom culture for risk-taking and respect for new ideas rather than resorting to an IRE mode of discourse (Michaels & O’Connor, 2012).

Rachel explained the effect of the curriculum’s embedded photos and videos on student motivation, saying it “changes for the positive when they're able to see something to connect to rather than hearing me say something.” Rachel was pointing out the value of leveraging students’ visual register rather than their verbal register when it comes to getting students engaged in a lesson. Students tend to recognize and recall visual information more readily than verbal information (Cohen, Horowitz, & Wolfe, 2009).

Student Achievement

Rachel was teaching science for the first time during the current school year. Some of the questions in the interview protocol were specifically designed to ask teachers to make comparisons before and after their curriculum implementation experience. However, these questions were not pertinent to Rachel’s scenario. One question was specifically chosen to elicit Rachel’s outcome expectancy in the domain of student achievement, which was “What impact, if any, does the effort you exert while planning and teaching science have on student achievement in science?” Rachel’s answer focused on their own teaching, but did not elaborate on its observed effects on students. A similar follow-up question was meant to provide a second opportunity for Rachel to elaborate, which was “What do you think are the most important reasons that certain students may not be attaining high achievement in your science class compared to their achievement in other areas?" Rachel answered this question by referencing students' lack of motivation to deeply consider the chapter questions.

Case 4: Emily (Low PSTE, Low STOE)

Demographics and Background

Emily is a fourth-grade teacher who has been teaching elementary school for 13 years. She holds a master's degree and is a science curriculum contact for her Title I designated school. Her total STEBI-A score and PSTE subscale was ranked lowest out of all seven participants in the quantitative phase of the study, which measures science teaching self-efficacy. Her STOE subscale score was the fifth out of all seven participants in the quantitative phase. She had the lowest STEBI-A total score and PSTE subscale score, as well as the second lowest STOE score of all four cases.

Theme 1: Knowledge of Science Content, Practices, and Curricular Resources

Professional Development

Emily attended the professional development session during the summer prior to curriculum implementation. She recalled an Amplify Science employee running the session. She remembered seeing Amplify's digital environment for students and teachers and also recalled that the facilitator modeled a lesson. She described the remaining time as "time to plan with our coworkers."

Enjoyment of Science

Emily answered the survey question "I enjoy science" as a 2 out of 5, with a score of 1 meaning "Not at all" on the survey scale. She described her educational experience as a contributing factor stating, "…growing up I never liked science. It's really just a personal thing. I do not take that as, like, teaching science… I would rather read or do math, even in high school, and college science was hard for me." Emily's comment about college science being difficult hints at a possible lack of subject matter knowledge in science, which has a relationship with student outcomes (Sadler et al., 2013). However, Emily states, "I never try to instill that on my kids. I never tell them that." She may have an awareness of this bias, but bias and belief may manifest themselves in teaching behaviors in unexpected ways.

NGSS Knowledge and Strategies

When asked to compare her science teaching experiences before and after the implementation of Amplify Science, Emily described a marked difference in perspective.

> I think before Amplify my idea of science was just 'I should just give them a lot of information in whatever form that I can,' and I think through Amplify it's kind of moved me towards 'here's what we want to figure out.'

Emily referred to her pre-implementation practices as "delivering information" and attempting to astonish students with as much "cool stuff" as possible. She says that her role has shifted toward "clarifying" information and putting students into "discovery mode." Without using NGSS terminology, Emily describes the science and engineering practices (SEPs) students are developing, including the SEPs of "asking questions," "obtaining information," and "constructing explanations", as well as the crosscutting concept (CCC) of "cause and effect." She accomplishes "obtaining information" and "asking questions" respectfully through a protocol called "notice and wonder," which positions students to make observations about a phenomenon and ask questions that drive future investigations. Speaking of phenomenon, Emily was the only other interviewed teacher besides Amanda to specifically use the phrase to describe phenomenon-based instruction.

Emily noticed that she has been learning science content as well as SEPs through her experience using both Amplify Science and district-developed science units.

> I don't think if you asked me before Amplify if I knew what amplitude or wavelength was or if I looked at the little chart thing if I could show you the difference and I couldn't. We have another unit that's called 'moonflower.' I think last year before Amplify, if you would have given me a flower and asked me to name the different parts, I would not have been able to.

To clarify, the moonflower unit was developed by the school district to cover standards that were unable to be met due to condensing Amplify Science's curriculum. In the quote, it's clear that the units supported Emily's subject-matter knowledge (SMK), and Emily later elaborates upon the importance of SMK to science teachers.

> I think they would need to have knowledge of the content. Not a deep knowledge, but they need to be able to accurately explain the content. I think a good science teacher would have to have the management of knowing how to make the teacher moves.

Emily uses the phrase "teacher moves" to describe the strategies made available through pedagogical content knowledge (PCK). One "teacher move" that Emily uses is modeling humility and vulnerability to students when she doesn't have the answers to their questions. Emily will say "'I don't know. That's good thinking.' Then I'll kind of open it up to everybody and then let everybody discuss it." She uses research to answer student questions as they arise as well.

Overall, Emily's descriptions of her growth in understanding science concepts as well as her descriptions of sound science teaching practices were surprising given her low scores on the STEBI-A survey. However, self-efficacy and efficacy are different constructs. It is quite possible that Emily does not realize that her understanding of science and science learning are as strong as she articulates, accounting for low science teaching self-efficacy.

Curricular Supports

One feature of the curriculum that Emily appreciates is what she describes as "a really good writing component." Two interconnected science and engineering practices (SEPs) generally take place as students learn how to scientifically write, "engaging in argument from evidence" and "obtaining, evaluating, and communicating evidence." Emily reports that her students are "writing their observations and writing about what they can figure out" which seems to align with the aforementioned SEPs.

Curricular support may not emerge directly from the written or digital curricular materials, but may happen in the form of modeling or coaching. Emily's response on one particular survey prompt made me wonder about the role of administrative feedback within Emily's school. Emily was the only interviewed participant that responded, "Agree" to the prompt "Given a choice, I would not invite the principal to evaluate my science teaching." I asked her about the kinds of reservations she may have with her principal evaluating her science teaching.

> I would rather be observed and get feedback on something that I've had enough experience myself on how to get better. I think right now with just a second year of doing anything I'm still kind of figuring out how to get better and I would like to just have my own reflections before I would invite someone in.

Teacher appraisal can be a difficult process to navigate. The word appraisal itself means "an expert estimate of the value of something." An especially negative appraisal could potentially be detrimental to Emily's science teaching self-efficacy and result in a lack of motivation (Palmisano, 2017). Emily's low PSTE score on the specific STEBI-A question regarding her avoidance of teaching observations seems to be partially corroborated by her interview response.

I asked Emily what kind of self-reflection she takes part in. She leaves herself notes in the lesson's presentation slides that remind her to improve upon specific aspects of a lesson. She also uses her experience as a guide, explaining that her performance during certain lessons the previous year were sometimes memorable enough to warrant adjustments.

Practice Leading to Comfort

Emily described her science teaching ability by comparing it to other subject areas she teaches. While she feels like her "ability has increased," it lags behind other content areas.

> It wouldn't be weaker just because, not because of my personal preference, but just because of the amount of time that I'm doing it. I'm teaching reading two hours every day, I'm teaching math one hour every day, and I've done that for thirteen years… I haven't become much of an expert as I have with math or reading just because the amount of it that I'm doing. Plus, the experience just doesn't add up as fast.

This statement suggests that with more time to practice, that Emily believes she could improve her science teaching ability, and Emily explicitly confirms this stating "the more that I do it, the more I will... understand it." She described her first year of implementation as "learn as a I go." By her second year of using the curricular materials, she was able to anticipate and respond to student misconceptions, stating "Students wouldn't know the difference between amplitude and volume. I feel like I was better this year at being able to [explain] between the two of those."

Theme 2: Limited Time and Competing Demands

Emily focused on the stresses of competing demands at the elementary level. She listed a handful of her responsibilities, explaining "As a fourth-grade teacher, I'm prepping 'Knowledge Block (language arts).' I'm prepping a math lesson. I'm prepping 'Win Time' work. I'm prepping intervention and I'm prepping a second 'Knowledge Block' lesson." Emily paradoxically swings between resignation to the circumstances and the desire to improve her science teaching ability in back-to-back comments.

> It just kind of sucks sometimes to leave something out or it kind of ends up being the last thing on your list… I don't know if I'd call it frustrating, but it's almost

> just the reality of my world....it's really cool stuff and I want to get better, but there's so much going on in a day for a grade level teacher in elementary school.

Science instruction decidedly occupies a small proportion of a fourth-grade teacher's day in this school district, amounting to 45 minutes per day, 15 days per quarter. The amount of budgeted time may convey to teachers the importance of each subject, with more time equating to higher prioritization of variables such as teacher effort, time committed to planning, etc. Emily provided a suggestion that would allow other individuals with more time and expertise to teach science.

> ...many years ago at [my school], my very first year, we had a science specialist. ...all she was doing was prepping the science lessons. ... that just seemed so cool planning-wise and she was such an expert on it. If I were to start my own school I would probably have science with a specialist. And especially with Amplify because it's great materials. There's so much to them. If we could have just one person teaching this that'd be cool.

Theme 3: Student Reactions

Student Interest and Engagement

One of Emily's first impressions of the new science curriculum materials was that "…it just has pretty cool topics. Students are pretty engaged in the topics." One of the student learning outcomes Emily noticed was that "I don't have to try as hard to motivate them because they just naturally are engaged with these topics. They just generally like them more." She has noticed that "specific students who just really are not into reading or math or writing" are drawn to science class. She speculated that those students may not thrive in situations that require algorithmic thinking, such as when they're learning "a grammar skill or how to do fractions" Instead, students are using scientific heuristics, and the learning environment in Emily's science class "has lower stakes in a sense, like, there's not necessarily a right or wrong answer every time."

Emily also noticed a few passing moments where students were making cross-curricular connections. "If we're learning about something in 'Knowledge Block' and it's about something we learned about in science they'll see that connection and they'll just be excited to tell me."

Student Achievement

When asked to compare students' science achievement before and after the implementation of Amplify Science, Emily started with a description of course grades, stating "That ended up being kind of similar, because previously, before Amplify, when I was just giving the information students were still learning it just fine and they were meeting objectives." Emily shifted the conversation toward conceptual understanding to communicate the changes in science achievement.

> It just wasn't as fun or engaging or provocative thinking wise. Now I think they're still receiving very similar grades, just the way that they're getting there makes them think more. Like they're just becoming better thinkers. They were still learning the stuff before just fine, but it was, you know, just kind of sit and listen and 'do you understand' and now it's like … the net positive is they're actually gaining some thinking skills while also still meeting the objective.

These skills were elaborated on in previous sections, but include the domains of asking questions, making observations, arguing using evidence, and in written communication of science ideas.

CHAPTER V: CROSS-CASE COMPARISON

This chapter describes the similarities and differences between each case. This chapter helps answer the mixed methods research question, "In what ways do the interview findings with elementary science teachers help to explain the quantitative results on science teaching self-efficacy and outcome expectancy?" Connections between the quantitative STEBI-A data and the qualitative interview data will be revealed, including instances where the two data sets agreed or disagreed with each other. The shared experience of curriculum implementation led to commonalities in how teachers perceived their own curriculum enactment. However, the STEBI-A results seem to predict trends that suggest variations in self-efficacy and outcome expectancy are related to certain curriculum enactment beliefs and behaviors.

The cases revealed that teachers with higher science teaching self-efficacy and outcome expectancies enjoyed science, prioritized science, and recounted positive student engagement, interest, and motivation more frequently than teachers with lower science teaching self-efficacy and outcome expectancies. Teachers that were involved in a curriculum selection process or that received coaching sessions from a science coordinator reported developing skills and confidence in teaching science. All teachers reported an increased comfort in teaching while using NGSS-aligned curriculum materials over time, but also bemoaned the lack of time to teach science during the school year and referenced their limited experience with professional development. There were a wide range of beliefs about the nature of inquiry in science. Teachers' use of science curriculum support materials was varied as well.

Theme 1: Knowledge of Science Content, Practices, and Curricular Resources

Finding #1: Being involved in a curriculum pilot and selection process was tied to high science teaching self-efficacy

Amanda had the unique experience of participating in her district's curriculum pilot, selection, and implementation processes. She also happened to have the highest STEBI-A total score and PSTE of all quantitative participants as well as the highest STOE of all interviewed participants. Amanda had the distinct advantage of learning about reform-oriented science practices inspired by *A Framework for K-12 Science Education*.

> We went through an entire semester where we did Zoom meetings with Amplify and they really trained us on anchoring phenomenon… then also looking at all of the 3-dimensional science pieces. They gave us a lot of background information on that.

Amanda is one of two teachers who explicitly says the phrase "anchoring phenomenon." She was the only teacher to mention the three-dimensional structure of the *Framework* and to bring up the storyline model of instruction. Amanda met with district science leaders four times to learn about the general structure of the *Framework* and additionally met with Amplify presenters one to two times per month for a semester. These experiences exemplify sustained professional development (PD), which is one of the six characteristics of PD that improves pupil attainment originally put forth by Cordingley et al. (2015) and Dunst et al. (2015) and later partially affirmed by Sims and Fletcher-Wood (2021). Although the similarities of the pilot PD program in this instance and the PD program in Lowell and McNeill (2023) are marginal, they both involved at least three PD sessions resulting in participants leaving with reform-oriented science teaching beliefs.

Amanda left these experiences with a robust understanding of reform-oriented science teaching. Without pretest data, we'll never know if Amanda had high science teaching self-efficacy going into these professional development or curriculum implementation experiences. However, in comparison to the interviewed participants, she seemed to be self-assured to a point where nearly none of her comments framed science, science teaching, science curriculum, or students in a negative light. Obstacles to teaching science such as limited time did not seem to deter her as much as other interviewed participants.

Finding #2: Coaching from a science coordinator provided opportunities for improving science teaching self-efficacy

All of the interviewed participants mentioned the district's science curriculum leadership team as playing a role in influencing their use of science curricular materials. One leader, the district K-12 science curriculum coordinator, offered personalized assistance and mentorship to two of the teachers in the study. The coordinator was able to persuasively reason with Rachel in a way that changed their perspective on the science teaching practices described in the unit materials. "[The district K-12 science curriculum coordinator] and I met on Zoom just one-on-one. Having 30 minutes with her, I was like, 'Oh, that is another way to think about it.'"

Personalized site visits by the coordinator were used to build Sarah's self-efficacy in managing laboratory materials:

> One of our units we do ecosystems and the students build their own terrariums, and we were really thinking we have no idea what we need to do. And [the district K-12 science curriculum coordinator] is always sending us emails saying, ‘please let me know if there's anything you need. I'm happy to come out and plan with you. I'm happy to help you come out and set things up.’ She brought out the lesson guide with us and just walked [us through it]. She had all three fifth grade teachers together and we were able to get all that set up for our next lesson, which really helped us... she was kind of talking us through the lesson and answering any questions we had.

The coordinator has an existing relationship with each of the interviewed participants who all happened to be current or former science curriculum contacts in the district. It’s assumed that building science teaching identity is a goal of the science coordinator’s, based on the following statement from Emily: “It's my first year [as a science curriculum contact]. ...[The district K-12 science curriculum coordinator] really likes the word ‘science champion.’” Building science teaching identity is important because it directly impacts the development of students’ science identity (Kane & Varelas, 2016). Rather than positioning elementary curriculum contacts as conduits for spreading information, the science coordinator attempts to position them as role models, if not leaders. Emily described her role as curriculum contact as leaning more toward the information conduit route, stating “I wouldn't say it’s the expert. It's just more of the communication facilitator or the go-to for questions. Not about content-wise, but just... curriculum-wise.”

Finding #3: Teachers with higher self-efficacy enjoyed science more than those with lower self-efficacy

This finding emerges directly from the STEBI-A results, but was corroborated by the qualitative strand of the study. The teachers with the highest STEBI-A scores had the highest self-rating for enjoying science. The teachers with the lowest STEBI-A scores answered "Do you enjoy science?" with a 2 out of 5, with a "1" on the scale meaning "Not at all." This question was meant to prompt teachers to think about science as a discipline rather than as a class, but oftentimes the question was interpreted as the latter. Since higher PSTE scores are associated with a preference to teach science (Lucas, Ginns, Tuli, & Waiters, 1993), the second interpretation of the question provides some predictive power surrounding the interviewed participants' preferences toward science teaching.

Amanda and Sarah had the highest STEBI-A scores and talked about how their students' engagement and joy, respectfully, influenced their own enjoyment of science. Amanda described the gradual shift in her science teaching self-efficacy from the start of her teaching career.

> I was a bit terrified because I didn't feel like that was probably my strongest area of teaching, but the more that I learned about it and the more that I've done it and seen how engaged students are, then it really has become one of my favorite things to teach.

While Amanda says she began teaching with very little science subject matter knowledge (SMK), the act of teaching science helped her gain confidence in the discipline.

Sarah’s enthusiasm for science is shown by her willingness to run her school’s science club and science fair. Seeing her students enjoy science in and out of the classroom complements her own enjoyment of science.

Of the interviewed participants, Rachel and Emily were on the opposing side of the science teaching self-efficacy spectrum, and their enjoyment of science based on their response to the survey question was the lowest of all participants in both phases of the study. Rachel’s enjoyment in science and science teaching come from early school and family experiences. When Rachel talks fondly of science, they bring up feelings of nostalgia, showing their students their mother’s space program memorabilia and discussing local phenomena to engage students. When Rachel is not enjoying teaching science, it’s due to the incongruence between their prevailing hardline stance on “hands-on” science teaching and the curriculum’s proposed practices. This type of incongruence can inhibit the uptake of research-based, reform-oriented practices (Blakely et al., 1987; Ben-Peretz, 1990). The clash between what Rachel expects out of science curriculum and what they’re experiencing may explain Rachel’s low STOE scores.

Emily was more direct about her response in this question stating that, “college science was hard for me.” SMK is foundational to pedagogical content knowledge (Abell, 2007), but it also contributes to higher levels of science teaching self-efficacy (Velthuis, Fisser, & Pieters, 2014). Emily’s suspected lower SMK may explain her low STEBI-A scores, and therefore, her low enjoyment of science. Lower PSTE scores are associated with a lower preference to teach science (Lucas, Ginns, Tuli, & Waiters, 1993). Emily described her preference for reading and math over science as high school student.

Emily's recommendation that science teaching responsibilities would be better left in the hands of science specialists could be interpreted in a few ways that make corroborating Emily's STEBI-A scores and her interview data less clear. Emily may value science education in such a way that she imagines students being better served by an individual with more time and expertise. This interpretation contradicts with Emily's low PSTE scores. Another interpretation is that she may not feel efficacious in teaching science, enough that passing science teaching responsibilities to another individual would be a relief. This interpretation would align with Emily's low PSTE scores.

Finding #4: Elementary teachers believed science learning should be "hands-on," but with varying rationales about the nature of inquiry

All of the teachers had an implicit level of understanding of different aspects representative of the shifts posed by NGSS. Every teacher held similar beliefs that some form of "hands-on" instruction involving "figuring out" or "discovery" were essential to learning. However, the rationale for these beliefs were varied. The frequency of the use of these terms in interview transcripts is shown in Table 14.

Table 14. Frequency of Inquiry Terms

	Amanda	Sarah	Rachel	Emily
"Hands-on"	5	3	6	0
"Figuring out"	22	0	0	3
"Discovery"	0	0	0	3

One commonality is that Amanda, Sarah, and Rachel all talked about the value of students learning from mistakes. Science and engineering are both incremental and iterative processes that imply that knowledge is never complete and that solutions can generally be improved upon. Although the teachers in this study may be more focused on building persistence and a growth mindset, they're also easing students into the Popperian principle of falsification.

This commonality in "learning from mistakes" helps frame what teachers do next. It is well-established that Amanda understands the structure of the shifts posed by the *Framework* and NGSS. She understands that role of a science teacher is closer to that of a facilitator than a lecturer. She is "trying to help them figure out questions to ask" instead of telling them exactly what to think. While Emily was on the opposing side of the rankings for science teaching self-efficacy, her description of "figuring out" was similar to Amanda's, and it was clear that Emily knew how to play what she called a "clarifying" role.

Contrast the frequency of inquiry-related terms used by Amanda and Emily to those used by Sarah and Rachel, who only talk about science learning as "hands-on." Sarah clearly values "hands-on" because that's "where the learning comes in." There is no rationale provided, and hardly any context is available as to why her assertion must be true. Rachel on the other hand, conflates "hands-on" with "engineering" and sees it as a necessary component to science instruction due to its ability to motivate students. However, Rachel rigidly applies this standard of science learning, and fails to see the utility of models and simulations in facilitating investigations. I asked Rachel about the potential value of the simulations in their capacity to probe and investigate system models.

WS: “The simulation in a way could be an investigation. Do you see any possibility of leveraging those to do investigations?"

Rachel: “It was just the same thing every day. It was 'take away the decomposers. What happens to the plants? Okay, take away the plants. What happens to the bunny rabbits? Take away the bunny rabbits. What happens to the wolves?’ When it was day after day the same thing, which they were getting just tired of, it's hard to keep them motivated for 15 days if you're doing the same thing. ‘Take away the sunlight. What happens to the plants?’”

WS: "Is it a matter of that it just was an unsurprising result? It seems like the pattern is really strong, but not surprising."

Rachel: “Yeah.”

WS: “In a way, it’s not the same. They are testing different things.”

Rachel: “Yeah, they are, but they’re also like, ‘We know that.’”

Rachel’s perception was that students already knew what to expect, thus prematurely stemming their motivation to investigate. However, Rachel was not seeing the simulation’s value in testing hypotheses, something that they were adamant about in describing “hands-on” learning. Each series of questions Rachel mentions in the previous exchange is a new opportunity for students to develop a hypothesis, gather data, analyze evidence, and ask new questions. Should that cycle of investigation culminate in a pattern, then students would be able to, at minimum, make a strong scientific claim. Students could also apply findings of one model to similar phenomena. Rachel only classified this type of systematic thinking as “hands-on” if students’ hands literally touched a physical object.

The teachers that only describe science learning in terms of “hands-on” may hold misconceptions that result in its misuse. Windschitl (2002) found that “hands-on” science activities may simply be used for the sake of giving students something to do. For example, counting the number of drops of water that can fit on the surface of a penny may have scientific value in demonstrating the properties of water. However, students may walk away with the impression that they simply needed to memorize the number of drops of water that could fit on a penny. When teachers don’t have a sense of the learning goal, they may fail to direct students’ attention to the important aspects of a “hands-on” lesson.

Finding #5: The perspectives of continuous supports for science curriculum use were variable

Most of the teachers reported that the Amplify Science curricular materials supported their science teaching consistent with the findings by Harris et al. (2022). Sarah and Amanda both used embedded lab investigation videos as a substitution for student-driven investigations when time was running short in their 15-day science unit. Sarah found additional value in the same educative supports when there was time to plan, saying that the videos helped her understand the purpose of specific lab investigations and that the videos had clear explanations of content.

The embedded educative supports also featured model teachers that explained pedagogical strategies, as described by Davis and Krajcik (2005). Sarah endorsed this educative support, stating, “if there’s something you're struggling with and you're not sure how you're going to teach it, it has the resources where you can go in and you can watch another teacher teach it.” Hodgson and Wilkie (2022) found that observing teacher models can result in shifts in practice. They provided evidence that decomposing pedagogical actions for observers and asking them to confront assumptions about traditional approaches to teaching were necessary to invoke these changes.

Sarah and Rachel used the curriculum's embedded "readers," but valued them for different reasons. Sarah used them as a model of non-fiction text to enhance her students' understanding of English language arts concepts, and they became part of her classroom library as an option for personal reading. Rachel used the "readers" in concert with translation services to support her English language learners.

Sarah and Rachel also mentioned the teacher's guide, which contained lesson plans and educative features. Sarah uses the teacher's guide, but reported that it was so expansive that she doubted teachers would have the time to fully utilize its features. Rachel had both praise and criticism for aspects of the teacher's guide. One criticism was the perception that the teacher's guide did not adequately explain the scientific background content that teachers would need to know. Rachel spent time learning content through web searches when the teacher edition did not adequately explain science concepts. This can be a time-consuming process, but more importantly, the quality and type of information yielded from a web search may lead to misconceptions or teaching above or below what is considered age-appropriate in the standards.

While Rachel criticized the unit questions for their vagueness and perceived lack of engaging qualities, they did find value in the "guided questions." Cherbow and McNeill (2022) found that these scripted resources can be supportive of teacher's talk moves and questioning skills. Rachel describes using the "guided questions" with their "higher level thinkers" to "probe them a little more."

The structure of written scientific argumentation can differ from other types of writing prompts. Emily was the only participant to report on the writing component of the curriculum, which she says helped her students in "writing their observations and writing about what they can figure out."

Finding #6: Practice contributed to increased comfort in using science curricular materials

Subject matter knowledge (SMK) and teacher self-efficacy influence one another (Bleicher & Lindgren, 2005). SMK is foundational to pedagogical content knowledge (PCK) (Abell, 2007). Much of teachers' growth in SMK is due to the reorganization of knowledge rather than the addition of new knowledge, and this can happen as a result of the act of teaching (Arzi & White, 2008). Furthermore, a major source of growth in PCK is teaching experience (van Driel, Verloop, & de Vos, 1998). Nixon and colleagues (2019) found that science teaching experience was related to the development of SMK in elementary teachers and that teaching experience served as a form of self-directed learning. Smith and colleagues (2022) found that teaching experience and science teaching self-efficacy were predictive of participants' science SMK.

Given the importance of teaching experience on SMK, PCK, and teacher self-efficacy, it was an encouraging finding that every teacher in the qualitative phase reported that continual practice led to increased comfort in teaching science. This study did not have a pretest-posttest design to measure changes in science teaching self-efficacy, and "comfort" and "confidence" are not the same as self-efficacy. However, teachers' reports of increased "comfort" suggests that their science teaching self-efficacy may have improved throughout the course of their science curriculum implementation experience.

Amanda reported feeling “discomfort” at the beginning of her teaching career, saying that she didn’t feel like she had the confidence, knowledge, or understanding of science to teach the subject. She said her constant learning and comfort have led to joy in teaching science. Sarah reported a similar outcome, stating “now that we're getting more familiar with it, we love it. I feel like it’s very smooth.”

Rachel had much more to say about the merits of practice and gaining science teaching experience. They felt “much more comfortable with the curriculum” and “much more capable of guiding” students. Rachel said the teacher’s guide cannot fully prepare teachers for unique experiences, stating “until you see the students learning it, you don't know exactly how it's going to come together.” They noted that their questioning strategies and ability to improvise improved after each lesson.

Emily said “my ability has increased because of my experience.” In the same vein as Rachel, Emily said she was able to anticipate and address student misconceptions as she had the chance to teach a lesson a second time.

Theme 2: Limited Time and Competing Demands

Finding #7: There was not enough time to teach science over the course of a year

From the beginning of the science curriculum implementation process, it was clear to Amanda that their school district would be limited on the instructional time allotted to science.

> ...unfortunately the district has about a third of the amount of time for each unit that Amplify provides as far as the amount of curriculum…So we really had to go

through the Amplify curriculum and kind of pare it down and see what…was important enough that we wanted to make sure we included and what pieces we could… weed out for each unit.

The school district in this study budgeted 15 days per quarter of science instructional for grades 3-5 for a total of 60 days of science instruction per school year. This is one-third of the 180 total instructional days for the school year. Third, fourth, and fifth grade are respectively allotted 35, 45, and 55 minutes of science instruction per day, which also amounts to 35, 45, and 55 hours of science instruction per year. These amounts of instruction are eclipsed by the 360 hours per year allotted to English language arts (ELA) instruction and 180 hours per year allotted to mathematics instruction. Teachers were clearly aware of these time constraints, with Sarah stating, "With 55 minutes, 15 days a quarter you just can't get that deep into the topics that you want to."

Difficult decisions had to be made during the curriculum selection and adaptation process. Amanda articulately described the purposeful exclusion of specific lessons by stating, "…we tried to balance…the amount of hands-on activity students were doing, the amount of reading and writing, because most of their lessons had all of those components in each lesson or pretty close." Amplify Science follows Lawrence Hall of Science's "Do, Talk, Read, Write" approach, and the components Amanda describes seems to match the components of Amplify Science's approach.

The "balancing" act was not merely arbitrary pruning the unit to meet a certain quota and proportion of activity types:

Amplify does a lot of building and repeating things so that students can really get that concept, so if there was the same idea repeated multiple days we would take

> out, you know, one or two of those days, so they were at least hitting on it, but unfortunately not going back to it for multiple days. We really just looked at the unit as a whole and the overarching idea and concept and then went back and looked at each day and to see where we could still end up with the same end goal…

According to Amanda, the curriculum adapting committee did its best to remain faithful to the three dimensions of the standards in its compromise, but perhaps at the expense of what the *Framework* calls "coherence." Coherence is when "components across the levels cohere or work together in a harmonious or logical way" (NRC, 2012). This may mean that "vertical alignment" exists, or that learning builds between grade levels. It may also mean that internal consistency is present between lessons and units. A hierarchy of science topics and skills would be learned through logically sequenced and smoothly connected lessons, and Amplify Science calls these coherent units "Progress Builds." Depending on which lessons or activities were removed during the adaptation process, the coherence of the unit may have been disrupted. The cumulative effects of deletions may weaken the coherence of a unit. If lessons are left out that demonstrate causal links such as how components of a system interact, this may cause confusion for learners and force teachers to fill the gaps with learning that should have been present in the first place. If adaptations are undertaken haphazardly, entire learning standards could go missing.

The curriculum adapting committee seems to have had an awareness of these pitfalls. Standards for each grade level were indeed missing, and the district designed one science unit for each grade level to cover those standards. When Amanda was asked how these in-house units compared to the Amplify Science units, they responded:

> … we wrote it with an anchoring phenomenon and we really tried to model it based on Amplify…we went back even last summer and did some editing and had… key questions that we were trying to answer…and we made our… earned word vocabulary and we really tried to… make it as similar as we could so it felt pretty cohesive...we have a storyline that we tried to follow for sure.

Once the units made their way into classrooms, more modifications were bound to happen. Three teachers described the concessions they had to make to conserve instructional time. Sarah stated, “I will tell you if we're running behind schedule, the experiments are the first thing to go.” Amanda described what happens next stating that “Occasionally we would have teacher demonstrations instead of the students doing those hands-on activities just depending on how much time we thought it would take.” If teacher-led demonstrations were too time-prohibitive “…Amplify has some videos… of experiments and then we're like, okay, we don't even have time for us to teach.”

These compromises were surprising given that all of the interviewed participants carried firm beliefs that “hands-on” instruction, “figuring out,” or “discovering” is how most learning occurs in science. This belief is not far from what is described in the *Framework* (NRC, 2012). Students at the elementary level learn procedural skills such as deciding what to measure and how to measure it. Investigations allow students to learn more about the nature of science. An age-appropriate learning outcome is that laboratory exercises don’t have to occur in a laboratory to be legitimate science. Students develop the ability to define patterns, learn how to communicate them, and that those patterns may suggest causal relationships.

A justification to specifically forego investigations was not entirely clear other than “rushing to get through the lessons.” Most teachers showed some sort of regret at the concessions. Emily described how she was resigned to the situation by stating “It just kind of sucks sometimes like to leave out something…I don't know if I'd call it frustrating but it's almost just like the reality of my world.”

Finding #8: The professional development experience for implementation was limited in its effectiveness due to time constraints

As part of their teacher employment contract, the school district in this study requires three and a half hours of professional learning from the district for teachers to increase their knowledge of academic content and teaching skills. At the elementary level, there is generally some level of teacher choice involved in how this time is used. However, the science curriculum department was given preeminent standing due to their implementation of science curriculum, and all elementary teachers were required to attend the training. Any future professional development sessions in science were not compulsory, and the following summer’s required professional development would focus around the implementation of English language arts curriculum. This meant that the science curriculum department had a single 3.5-hour chance to convey their most important messages regarding curriculum implementation. The limited time allotted to this required professional learning meant that major compromises would have to be made to get the implementation off the ground. Amanda described the main components of the grade-specific training she led.

> We went through each of the units and gave them an overview and we did a small, miniature version of an anchoring phenomenon. We introduced them to one of the units so that they would get the idea of how that process works.

The Sims and Fletcher-Wood (2021) narrowed down the characteristics of effective PD from six (Cordingley et al., 2015; Dunst et al., 2015) to only three characteristics. These characteristics include whether a program involves outside expertise, is practice based, and is sustained. The training met the “external expertise” threshold, meaning teachers from different buildings are able to openly discuss and exchange new ideas. Every elementary teacher in the district was in attendance and had the opportunity to bring fresh ideas to the table. Amanda helped lead break-out sessions, and said “we would give them time to process and talk to the people at their table about what that might look like in their classroom.”

The “practice-based” recommendation is characterized by actively applying ideas learned in a professional development program. Practice could happen in the form of discussion, lesson design or reflection, or teaching mock lessons. While some discussion structures were a component of the training, the opportunity to teach a mock lesson was not. Amanda was asked if teachers were allotted time to practice teaching a lesson to each other during the professional development session, they responded, “They didn't unfortunately… they didn't have the opportunity to actually really dig into a lesson and plan it, no.” Teachers who took part in the professional development session as learners described a similar experience stating, “I wouldn't say … practice would be the best way to describe it... they did have one lesson and just kind of ‘How do you plan this out? What materials would you need?’”

The final characteristic of effective PD is that it is sustained. Rachel negatively summarized their experience with science professional development throughout implementation by stating, “I just don't think that the professional learning piece was there and I know people are going to want more professional learning.” Rachel’s testimony supports not only the demand for sustained PD, but the necessity of it.

The science curriculum department in this district had to take a pragmatic stance in following the district’s professional learning model. If this district were to take the lead of the literature on science professional development, they would need a much more sustained PD model. According to Lowell and McNeill (2023), at least one PD session consisting of four days would be necessary just to shift elementary teachers from traditional science teaching beliefs toward reform-oriented beliefs, and it would take at least two additional four-day PD sessions to increase storyline implementation self-efficacy provided that teachers would have time to analyze model teachers in action and go through cycles of rehearsal and enactment. Even if all the previous recommendations are followed, Penuel and colleagues (2023) found that it may take up to three years to increase science teaching self-efficacy.

Finding #9: Science is prioritized less than other school subjects, but even more so by teachers with low PSTE and low STOE

All of the interviewed teachers identified a looming pattern that science is undervalued as a content area. They implied that top-down pressure from the state, their school district, and from building leadership were responsible for some of the perception of science’s diminished status in the educational hierarchy. While there were vague statements such as when Amanda said “I feel like [there] is not near enough time focused on [science]," there were more explicit assertions about state and local policies and structures that influence how science can be taught. Sarah's criticisms were levelled at systems of accountability through state assessments.

> I just don't feel like there's enough importance placed on science in elementary school… there's definitely more of a priority placed on math and reading. That's what we're tested on. Fifth grade is the only elementary school grade level that's tested on science.

In the state this study was conducted, students take annual English Language Arts (ELA) and Mathematics assessments starting at the end of the third grade until the end of their eighth grade. At the end of students' fifth grade year, students take a science assessment encompassing the content, or disciplinary core ideas (DCIs), of that grade level. However, the skills acquired over the preceding three years are used to navigate the other two dimensions of the 3-dimensional assessment. The aggregated outcomes of these assessments are publicly reported and have implications in school funding. Sarah made a connection between test scores and how teachers may prioritize their choices in professional development.

> [ELA and Mathematics test scores] are what principals are looking for a lot of the time. When teachers are looking for … professional development, typically they will choose the one that is going to get them… the best ranking … within the district, because there's always a comparison between all the schools.

Teachers may voluntarily gravitate toward professional development in ELA or Mathematics because annual state assessments in ELA and Mathematics give the illusion of an immediate causal link between the acumen of individual teachers and the test scores of their cohort of students. While fifth grade teachers are responsible for administering the state science assessment, it becomes unreasonable to hold them solely responsible for the cumulative skills that students should have acquired throughout third, fourth, and fifth grade science courses.

It is reasonable to assume that the collective long-term efforts of three consecutive science teachers contributes to students' test scores. However, the three-year model of state science assessment provides little more than contentious conclusions. Does netting high science assessment scores communicate a laudable act of teamwork, or did students simply have an outstanding fifth grade teacher?

Furthermore, what should this school district think about state science scores once students have had a chance to experience three years in a row of a programmatic approach such as Amplify Science? Will low state science assessment scores at a building mean that at least three teachers are to blame? Surely the district's seven-year contract for $3,000,000 worth of Amplify Science materials should carry a price tag that guarantees excellence, that is unless teachers aren't provided with enough time to implement it with fidelity. And what if Amplify Science turned out to be potentially inefficacious? How many school leaders would shoulder the blame of that kind of purchase? Would they jettison Amplify Science in lieu of purchasing materials for another NGSS-aligned science curriculum? There are too many variables at play to make any sort of reasonable conclusions from this testing model, let alone to make policy or classroom decisions based on that data. Perhaps elementary teachers have similar intuitions and adjust their professional development choices accordingly.

The school district in this study also purchased and implemented new ELA curriculum, Amplify CKLA, following the implementation of Amplify Science. Rachel contrasted the communication they received about district-led science and ELA professional development opportunities by saying "…with CKLA…they hammered it down our throat this summer. You could go to so many classes and learning opportunities surrounding that… I didn't feel were provided [for science]."

Educational policy wasn't the only thing deemphasizing the importance of science education for teachers in this study. School scheduling had an impact on science planning and instruction as well. While describing the limited time they had to teach science, Sarah added "…typically if things happen during the day, it's during our "unit studies," so it's the first thing to get cut, whether it's science or social studies or health." The "things" they were describing were guest speakers and other planned interruptions. A conscious effort by school leaders has to be made to mitigate the loss of instructional time in ELA and Mathematics at the expense of instructional time in science and other content areas.

Teachers with the lowest PSTE and STOE subscale scores discussed how a confluence of factors precluded them from prioritizing science planning and instruction. An emphasis on ELA and Mathematics planning was clear.

> I'm teaching reading 2 hours every day and I'm teaching math 1 hour every day and I've done that for 13 years…I'm prepping "Knowledge Block [CKLA]." I'm prepping a math lesson. I'm prepping "Win Time" work. I'm prepping intervention and I'm prepping a second "Knowledge Block" lesson.

The emphasis on ELA was clearest with Rachel who stated "I'll be real [sic] honest, science is the last thing I think about. Every day with everything else… you know, CKLA and things like that…" Given the circumstances, prioritizing ELA and Mathematics makes sense. Once the quarterly 15 days of science were finished, a new unit study course provided teachers with a fresh start, but ELA and Mathematics classes happened every day for multiple hours all year long, requiring constant attention.

Theme 3: Student Reactions

Finding #10: Teachers with higher STOE subscale totals recounted positive student engagement, interest, motivation, and achievement more frequently than teachers with lower STOE subscale totals

All of the interviewed teachers elaborated on their students' engagement with Amplify Science curricular materials, and many of those comments focused on the positive impacts their teaching had on students. However, it was clear that the teachers with higher STOE subscale totals recounted these positive descriptions of student engagement, interest, motivation, and achievement more frequently than teachers with lower STOE subscale totals.

As qualitative analysis was conducted, participants' responses were labeled based on the interview protocol question initially posed to them. Responses to these questions were labelled with one of eight major code-categories prior to the formation of themes. One of those categories was student responses in regards to engagement, interest, motivation, and achievement. A quick count of the frequency of individuals' quotes in this category revealed the pattern seen in Table 15. Participants are listed in descending order by STOE subscale score.

Table 15. Frequency of Quotes in the "Student Responses" Category

	Amanda	Sarah	Emily	Rachel
STOE	42	39	36	35
Positive Outcomes	7	6	3	3
Negative Outcomes	0	3	0	4
Total	7	9	3	7
Positive Outcomes as a Percentage of Total	100%	66.7%	100%	42.9%

This frequency table provides a rudimentary mixed-methods-style summary of the qualitative data. Not all participants were asked the same amount or type of questions, but the amounts and proportions of positive outcomes tell an interesting story about the participants' outcome expectancies.

Amanda exhibited a nearly textbook example of the causal relationship between science teaching self-efficacy and expected outcomes with the statement, "Science is probably my favorite thing to teach. I love teaching science and the kids... in large part because that's something that students are always really interested in and can feel successful at." From Bandura's (1997) point of view, Amanda feels efficacious in science teaching and expects student interest. Amanda also reported, "...the more that I've done it and seen how engaged students are, then it [science] really has become one of my favorite things to teach." If you take Kirsch's (1982) view, it's the outcome expectancy that has influenced Amanda's self-efficacy.

Sarah's only negative student outcomes come in the form of managing student behaviors during investigations and in regards to her school's drop in state science test scores. On the whole, Sarah expects excitement and achievement from her science students. Interestingly, both she and Amanda make a similar comment about their hope that continuous use of the curriculum from grades three to five will lead to better student understanding of science. Emily rarely talks about students, but when she does, she expects engagement and motivation. These teachers' positive outcomes in engagement are consistent with previous findings that student engagement increased while using Amplify Science materials (Uppendahl, 2020; Harris et al., 2022).

Rachel had the lowest STOE subscale score of the interviewed teachers, and most of their comments lean toward describing their students' eroded interest and decreased motivation due to the perceived lack of quality of the embedded unit questions. Rachel expected students to be unmotivated and to engage in learned helplessness behaviors when the anchoring phenomenon didn't carry its weight in terms of eliciting student engagement. Rachel noticed an uptick in engagement while student-driven questions were investigated or as students got to watch videos of phenomenon to prepare for a lesson.

Finding #11: Students' science achievement was unclear in terms of grades and test scores, but skills and conceptual understanding improved

Grades and Test Scores

Three of the four interviewed participants have taught science with at least one other science curriculum prior to implementing Amplify Science. These three participants were able to provide a comparison of what student achievement in science was like before and after implementing Amplify Science. Very few statements surrounding students' science achievement focused on grades as a metric. However, when grades or scores were brought up, it was generally in the context of the shifting landscape of science assessment.

The shifts posed by NGSS have turned the focus of assessment toward measuring students “3-dimensional” (3D) thinking. In 3D assessment, authentic and unfamiliar phenomenon are used to elicit not only the content (DCI) learned by the student, but also leverage crosscutting concepts (CCCs) such as patterns or scale, and science and engineering skills (SEPs) such as analyzing data. 3D assessments would allow learners to express their knowledge in a variety of forms. Sarah summarized 3D assessment as “a completely different way of measuring the success of students.” However, the continual development of 3D assessments by the state department of education and school district in this study make direct comparisons of student achievement difficult at best.

Rachel bemoaned the lack of teacher-friendly assessment.

> I think that would be helpful too, you know, with a reading test, you're like "The main idea was..." [I want] some more hard and fast pieces of assessment. Are they thinking deeper about those questions? I know the assessment piece is coming. It takes time.

The “hard and fast” assessment may have been a reference to traditional, forced-choice tests where only one best answer prevails. As Beyer and Davis (2012) pointed out, science teachers using ambitious science materials may believe inquiry skills themselves do not require assessment in the same vein as content knowledge. Traditional testing is becoming deemphasized in lieu of 3D assessments, but the shift may be a hard sell when it comes to practical considerations such as grading and providing timely feedback, as Rachel has alluded to. Rachel extended some grace about the implementation of 3D assessment stating, “It takes time.” It is important to note that it’s been over a decade since the *Framework* arrived, yet 3D assessment is still lagging behind. The lag is due to a variety of constraints such as the complexity of integrating all three dimensions in assessment and constructing culturally sensitive assessment (Furtak, 2017).

Students’ Conceptual Understanding

Almost all of the anecdotes provided by teachers in the study focused on their observations of growing students’ skills and conceptual understanding in comparison to pre-implementation science curricula. Emily summarized this point.

> They were still learning the stuff before just fine, but it was, you know, just kind of sit and listen and "do you understand" and now … the net positive is they're actually gaining some thinking skills while also still meeting the objective.

In regard to the “thinking skills”, teachers described NGSS science and engineering practices (SEPs) without using NGSS-laden terminology. The SEP “developing and using models” was one of the “thinking skills” that teachers noticed from their students. Amanda summarized this point by stating, “I feel like they have a better conceptual understanding probably than what they had. I feel like it requires them to do a lot more modeling and, you know, trying out models, creating models.”

In the 3rd to 5th grade level band, modeling is used as a way to describe and represent concepts to be learned. Modeling at the elementary level lays the foundation for the middle school and high school grade bands, where modeling is used for predicting and testing ideas (NRC, 2012). However, Amplify Science materials expected students to work slightly above what the *Framework* describes as grade-level by testing ideas during the 5th grade "Ecosystem Restoration" unit. Rachel described the simulation designed to model, predict, and test the effects of population changes on a simple ecosystem.

> …take away the decomposers. What happens to the plants? Okay, take away the plants. What happens to the bunny rabbits? Take away the bunny rabbits. What happens to the wolves?

Teachers were also surprised that their lessons had inspired newfound creativity in their students as they used the SEP "constructing explanations and designing solutions." The latter portion of this SEP aims at engineering design skills such as solving specific building challenges. Creating structures, noticing points of failure, and redesigning structures is age-appropriate for the 3rd to 5th grade level band (NRC, 2012). Amanda noticed that students were able to succeed in a design challenge in unexpected ways during the 3rd grade "Balancing Forces" unit.

> I have students every year that figure out... going back to the magnetic train, one of the experiments that we do every year, I have kids that, the last 2 years, that have figured out how to put it together different than what the curriculum… shows them in the end… if they haven't figured it out.

Emily described Amplify Science's "strong written component" as supporting her students. This may be true of her students, but it is not consistent with the findings of Uppendahl (2020) who wrote that Amplify Science's curriculum materials lacked repeated practice in written scientific argumentation at the elementary level.

CHAPTER VI: CONCLUSION AND IMPLICATIONS

The chapter includes recommendations for school districts that plan on implementing NGSS-aligned science curriculum. The chapter concludes with implications for future study in the areas of science curriculum implementation as well as science teaching self-efficacy and outcome expectancy.

Recommendations Based on Findings

Janssen and colleagues (2015) provided three practical recommendations regarding successful science curriculum implementation. One recommendation relevant to this study was making the investment of time and resources worthwhile and realistic to teachers. Another recommendation was to promote congruence between curriculum practices and teachers' prevailing practices. These recommendations go hand-in-hand with those of Sims and Fletcher-Wood (2021) regarding sustained professional development programs.

The school district in this study was constrained by its professional development model wherein 3.5 hours of required PD was the only guaranteed point of contact between teachers and PD developers. Spreading the 3.5 hours into multiple smaller sessions may not fully satisfy the recommendation that PD be "sustained." Terms of teachers' contracts could be renegotiated when curriculum implementation is approaching. However, outcome-based educational standards in any given content area in the U.S. tend to undergo semi-regular upheaval. This may differentially impact elementary teachers who are responsible for teaching multiple content areas.

A more likely recommendation related to finding 1 is to outsource continuing professional development duties through a shared leadership model, which is broadly defined by Simpson (2021) as utilizing classroom teachers in other school leadership capacities. Simpson's literature review on the topic of teacher leadership highlights its benefits, including those on communities and schools as well as on individuals such as students, parents, and the teacher leaders themselves. An urban school district such as the one in this study has the challenge of working under a large hierarchal structure that depends on developing successive groups of individuals, including principals, coaches, and teachers. The district in this study took the step of not only involving teachers in the curriculum selection process, but also heavily weighed their input in the process. This level of inclusion communicates to teachers that they are credible and trusted. Teachers involved in the curriculum selection process may be fitting candidates for supporting continuing professional development opportunities within their school and across their district.

In a shared leadership model, the responsibility for developing and maintaining the vision and quality of professional development for science teaching would still fall on science curriculum specialists. However, they would work with principals to entrust the responsibility of leading professional development to designated teacher leaders at each building. This type of model may build leadership capacity in teachers while also allowing professional development in science to be sustained, albeit informally, throughout a school year. Outcomes of such an initiative would likely be variable. Special care would need to be taken to prevent teacher leaders from simply propagating familiar practices, but instead demonstrating familiar strategies that are compatible with reform-oriented science practices (Davis et al., 2016). However, a teacher leadership model would at least offer an extension on learning that is more significant than the current PD model of the district in this study.

There are a number of antecedents that must be in place to establish a culture of teacher leadership, and school administrators are key players in this regard. Administrators must work collaboratively with teacher leaders to clearly delineate administrator and teacher leader roles (York-Barr & Duke, 2004). Resentment of teacher leaders by peers is a possible pitfall (Katzenmeyer & Moller, 2009). Knowing this, administrators must carefully select collegial teacher experts. The roles of teacher leaders should be framed around the goals and vision of the school as a whole. A culture of trust and collaboration is possible once teachers know what to expect as well as how and whether their input will be used (Helterbran, 2010).

Setting aside time and incentives to organize teacher leadership is a necessary challenge of the teacher leadership model. Prospective teacher leaders need assurance that leadership roles won't hinder their professional or personal responsibilities, and providing planning time or release time can ease the burden of such commitments. Stipends for teacher leadership work, professional development, or conference attendance are one strategy to incentivize teacher leadership. However, incentives need not be monetary and could come in the form of recognition or the choice of a preferred teaching schedule (Helterbran, 2010). Partnerships with other K-12 school districts, local organizations, and higher education institutions offer opportunities for leadership development as well.

A point more proximate to this study and self-efficacy in particular involves the development of teacher leader identity. Teachers may not readily see themselves as leaders (Helterbran, 2010). Additional training may be necessary to develop the dispositions and skills a teacher leader needs, such as a willingness to lead (Berg & Zoellic, 2019), strong interpersonal skills (Meirink et al., 2020), and a willingness to pursue continual growth and development (York-Barr & Duke, 2004). Administrators may not initially facilitate teacher leadership efforts. In this case, Hinnant-Crawford (2016) recommends that teacher leaders persist in working *with* administrators rather than attempting to work *around* them to achieve their goals.

Domain-specific science knowledge and confidence hold considerable value in developing a professional vision for leadership in science education (Criswell et al., 2018). Teacher leaders with a strong base of existing science content knowledge and pedagogical content knowledge were shown to exhibit strong teacher leadership skills and self-confidence in their leadership abilities after undergoing a teacher leadership development program (Hofstein, Carmeli, & Shore, 2004). Science teachers demonstrated their ability to act as science teacher leaders in programs that, in part, focused on developing subject matter knowledge and pedagogical content knowledge (Criswell et al., 2018; Mentzer et al., 2014). When viewing the findings of these studies alongside the findings of the current study, links between science teaching self-efficacy, outcome expectancy, professional development, SMK, PCK, and teacher leadership demonstrate that a multi-pronged, systematic approach will continue to be necessary in science teacher preparation, professional development, and science curriculum implementation efforts.

Implications for Future Study

Lessons learned from the timing and outcomes of this study inform a number of measures that may be taken in future investigations regarding science teaching self-efficacy and science curriculum implementation. A repeated-measures design that uses the STEBI-A and follows teachers before and throughout curriculum implementation would provide more detailed insight into the connection between curriculum implementation, science teaching self-efficacy, and outcome expectancy. Menon and Sadler (2016) used a similar study methodology using the STEBI-B with preservice teachers. Following their methodology using the STEBI-A with in-service teachers would add to the literature on this topic.

The National Research Council (2004) has called for more rigorous studies involving measuring fidelity of curriculum implementation. The modifications made to the curriculum in this study present an interesting opportunity for comparing the outcomes between partial, adapted, and fully implemented curricula. One unclear outcome in this study was objective measures of student science achievement, which were neither sought nor well-articulated by study participants. Student achievement data would help this district understand whether the compromises made to the curriculum had downstream effects on teachers' classroom enactment, such as in their ability to achieve fidelity to the goal of the full curriculum (McNeill et al., 2018).

References

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

Achieve.org. (n.d.). Contributors. Retrieved November 5, 2023, from https://www.achieve.org/contributors

Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education, 17*, 86-95.

Amplify Science. (2019a). Efficacy: Amplify Science field trial effectiveness research report. https://amplify.com/pdf/uploads/2019/12/AS_field-trial-effectiveness-research-report.pdf

Amplify Science. (2019b). Efficacy: Amplify Science WA grade 5 efficacy research report. https://amplify.com/pdf/uploads/2019/12/AS_WA-grade-5-efficacy-research-report.pdf

Amplify Science. (2019c) Efficacy: Students achieve more with Amplify Science. https://amplify.com/pdf/uploads/2023/02/AS_Flyer_Q1-Case-Study-Grand-Island_021323.pdf

Appleton, K., & Kindt, I. (2002). Beginning elementary teachers' development as teachers of *science. Journal of Science Teacher Education, 13*, 43–61.

Ardasheva, Y., Newcomer, S.N., Firestone, J.B., & Lamb, R.L. (2019) Contributions of language-specific and metacognitive skills to science reading comprehension of middle school English learners. *Bilingual Research Journal, 42*(2), 150-163, DOI: 10.1080/15235882.2019.1597774

Armor, D., Conroy-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., & Zellman, G. (1976). *Analysis of the school preferred reading programs in selected Los Angeles minority schools* (Rep. No. R-2007-LAUSD). Santa Monica, CA: RAND. (ERIC Document Reproduction Service No. 130 243)

Arzi, H. J., & White, R. T. (2008). Change in teachers' knowledge of subject matter: A 17-year longitudinal study. *Science Education (Salem, Mass.), 92*(2), 221–251. https://doi.org/10.1002/sce.20239

Bandura, A. (1977). *Social Learning Theory.* United Kingdom: Prentice Hall.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *The American Psychologist, 37*(2), 122–147. https://doi.org/10.1037/0003-066X.37.2.122

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: Freeman.

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Horizon Research, Inc.

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report, 13,* 544-559.

Bell, R., Matkins, J., & Gansneder, B. (2011). Impacts of contextual and explicit instruction on preservice elementary teachers' understandings of the nature of science. *Journal of Research in Science Teaching, 48*(4), 414–436.

Ben-Peretz, M. (1990). *The teacher–curriculum encounter: Freeing teachers from the tyranny of texts.* Albany: State university of New York Press.

Berg, J., & Zoellick, B. (2019). Teacher leadership: Toward a new conceptual framework. *Journal of Professional Capital and Community, 4*(1), 2–14. https://doi.org/10.1108/JPCC-06-2018-0017

Beyer, C., & Davis, E. A. (2012). Learning to critique and adapt science curriculum materials: Examining the development of preservice elementary teachers' pedagogical content knowledge. *Science Education, 96,* 130–157.

Blakely, C. C., Mayer, J. P., Gottschalk, R. G., Schmitt, N., Davidson, W. S., Roitman, D. B. et al. (1987). The fidelity-adaptation debate: Implications for the implementation of public sector social programs. *American Journal of Community Psychology,15,* 253–268.

Bleicher, R. E., & Lindgren, J. (2005). Success in Science Learning and Preservice Science Teaching Self-Efficacy. *Journal of Science Teacher Education, 16*(3), 205–225. https://doi.org/10.1007/s10972-005-4861-1

Bobbitt, F. (1918) *The Curriculum*. Houghton Mifflin Company, Boston.

Bolin, F. (1987). The teacher as curriculum decision maker. in F. Bolin & J. Falk (Eds.), *Teacher renewal: Professional issues, personal choices* (pp. 92–108). New York, NY: Teachers College Press.

Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). *The BSCS 5E Instructional model: Origins and effectiveness.* (pp. 88–98). Colorado Springs BSCS.

Campbell, T., McKenna, T. J., Fazio, X., Hetherington-Coy, A., & Pierce, P. (2019). Negotiating Coherent Science Teacher Professional Learning Experiences Across a University and Partner School Settings. *Journal of Science Teacher Education, 30*(2), 179–199. https://doi.org/10.1080/1046560X.2018.1547

Cherbow, K., & McNeill, K. L. (2022). Planning for student-driven discussions: A revelatory case of curricular sensemaking for epistemic agency. *The Journal of the Learning Sciences*, *31*(3), 408–457. https://doi.org/10.1080/10508406.2021.2024433

Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences*, *106*(14), 6008-6010.

Cohen, Michael. (2020). Statement from Michael Cohen, President of Achieve. Retrieved November 5, 2023, from https://www.achieve.org/files/Statement%20from%20Michael%20Cohen_FINAL.pdf

Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *Journal of Experimental Education, 60*, 323-337.

Combs, E., & Silverman, S. (2016). *Bridging the gap: Paving the pathway from current practice to exemplary professional learning.* Frontline Research & Learning Institute. https://www.frontlineeducation. com/uploads/2018/01/ESSA_Bridging_the_Gap.pdf

Corbin, J., & Strauss, A. (2015). *Basics of qualitative research* (4th ed.). Thousand Oaks, CA: Sage. Ch. 3: Practical considerations.

Corcoran, K. J., & Rutledge, M. W. (1989). Efficacy expectation changes as a function of hypothetical incentives in smokers. *Psychology of Addictive Behaviors, 3*, 22-29.

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., Saunders, L., & Coe, R. (2015). *Developing great teaching: Lessons from the international reviews into effective professional development.* Teacher Development Trust. https://tdtrust.org/wp-content/uploads/2015/10/DGT-Full-report.pdf

Council, J. R., Ahern, D. K., Follick, M. J., & Kline, C. L. (1988). Expectancies and functional impairment in chronic low back pain. *Pain, 33,* 323-331.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: choosing among five approaches*. Los Angeles, CA: Sage.

Criswell, B.A., Rushton, G. T., Nachtigall, D., Staggs, S., Alemdar, M. & Cappelli, C.J. (2018). Strengthening the vision: Examining the understanding of a framework for teacher leadership development by experienced science teachers. *Science Teacher Education, 102(*6), 1265–1287. https://doi.org/10.1002/sce.21472

Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? *Journal of Teacher Education, 53*(4), 286–302. https://doi.org/10.1177/0022487102053004002

Davis, E. A. (2006). Preservice elementary teachers' critique of instructional materials for science. *Science Education, 90*, 348–375.

Davis, E. A. & Haverly, C. (2022). Well-started beginners: Preparing elementary teachers for rigorous, consequential, just, and equitable science teaching. In Jones, M. G., & Luft, J. (Eds.), *Handbook of research on science teacher education* (pp. 83-96). Routledge.

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, *34*(3), 3–14.

Davis, E. A., Janssen, F. J. J. M., & van Driel, J. H. (2016) Teachers and science curriculum materials: where we are and where we need to go. *Studies in Science Education, 52*(2), 127-160, DOI: 10.1080/03057267.2016.1161701

Deehan, J. (2016). *The Science Teaching Efficacy Belief Instruments (STEBI A and B): A Comprehensive Review of Methods and Findings from 25 Years of Science Education Research* (1st ed. 2017 edition.). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-42465-1

Dunst, C. J., Bruder, M. B., & Hamby, D. W. (2015). Metasynthesis of in-service professional development research: Features associated with positive educator and student outcomes. *Educational Research and Reviews, 10*(12), 1731–1744. https://doi.org/10.5897/ERR2015.2306

EdReports.org. (n.d.). About Us. Retrieved November 5, 2023, from https://www.edreports.org/about

Enochs, L. G., Scharmann, L. C, & Riggs, I. M. (1995). The relationship of pupil control to preservice elementary science teacher self-efficacy and outcome expectancy. *Science Education*, *79*(1), 63-75.

Furtak, E. M. (2017) Confronting Dilemmas Posed by Three-dimensional Classroom Assessment: Introduction to a Virtual Issue of Science Education. *Science Education, 101*(5), 854–67. https://doi.org/10.1002/sce.21283.

Gess-Newsome, J., & Lederman, N. G. (Eds.). (1999). *Examining pedagogical content knowledge: The construct and its implications for science education* (1999th ed.). Springer.

Gibbons, L. K., & Cobb, P. (2017). Focusing on teacher learning opportunities to identify potentially productive coaching activities. *Journal of Teacher Education, 68*(4), 411–425. https://doi.org/10.1177/0022487117702579

Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*(4), 569-582.

Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education.* New York: Teachers College Press.

Hamilton, D. (2014). *Towards a Theory of Schooling*. Routledge.

Harris, C. J., Feng, M., Murphy, R., & Rutstein, D. W. (2022). *Curriculum materials designed for the Next Generation Science Standards show promise: Initial results from a randomized controlled trial in middle schools.* WestEd. https://amplify.com/pdf/uploads/2022/01/WestEd-NGSS-Curriculum-Study-Brief_FINAL-ADA_01_06_2022.pdf

Harris, C. J., Feng, M., Murphy, R., & Rutstein, D. W. (2023). *Supporting science learning and literacy development together: Initial results from a curriculum study in 1st grade classrooms.* WestEd.

Hatch, J.A. (2002). *Doing qualitative research in educational settings.* Albany: State University of New York Press.

Hauk, S., & Kaser, J. (2020). A Search to Capture and Report on Feasibility of Implementation. *The American Journal of Evaluation*, *41*(1), 145–155. https://doi.org/10.1177/1098214019878784

Helterbran, V. R. (2010). Teacher leadership: Overcoming 'I am just a teacher' syndrome. *Education, 131*(2), 363–371.

Hinnant-Crawford, B. (2016). Education policy influence efficacy: Teacher beliefs in their ability to change education policy. *International Journal of Teacher Leadership*, *7*(2), 1–27.

Hodgson, L. M., & Wilkie, K. J. (2022). Modelling lessons for more than imitation: investigating teachers' reactions and decompositions of unfamiliar practices. *Journal of Mathematics Teacher Education*, *25*(6), 749–775. https://doi.org/10.1007/s10857-021-09516-1

Hofstein, A., Carmeli, M., & Shore, R. (2004). The professional development of high school chemistry coordinators. *Journal of Science Teacher Education, 15*(1), 3–24.

Janssen, F. J. J. M., Westbroek, H. B., & Doyle, W. (2015). Practicality studies: How to move from what works in principle to what works in practice. *Journal of the Learning Sciences*, *24*, 176–186.

Joswick, C., & Hulings, M. (2024). A Systematic Review of BSCS 5E Instructional Model Evidence. *International Journal of Science and Mathematics Education*, *22*(1), 167–188. https://doi.org/10.1007/s10763-023-10357-y

Kane, J. M., & Varelas, M. (2016). Elementary school teachers constructing teacher-of-science identities. In L. Avraamidou (Ed.), *Studying science teacher identity: New directions in mathematics and science education* (pp. 177–195). Sense Publishing.

Kang, E. J. S., McCarthy, M. J., & Donovan, C. (2019). Elementary Teachers' Enactment of the NGSS Science and Engineering Practices. *Journal of Science Teacher Education, 30*(7), 788–814. https://doi.org/10.1080/1046560X.2019.1630794

Karpicke, J. D., & Roediger, H. L. I. (2008). Critical Importance of Retrieval for Learning. *Science (American Association for the Advancement of Science), 319*(5865), 966–968. https://doi.org/10.1126/science.1152408

Katzenmeyer, M., & Moller, G. (2009). *Awakening the sleeping giant: Helping teachers develop as leaders* (3rd ed.). Thousand Oaks, CA: Corwin Press.

Kirsch, I. (1982). Efficacy expectations or response predictions: The meaning of efficacy ratings as a function of task characteristics. *Journal of Personality and Social Psychology, 42*, 132-136.

Krajcik, J., Mamlok, R., & Hug, B. (2000). Modern content and the enterprise of science: Science education in the twentieth century. In l. Corno (Ed.), *Education across a century: The centennial volume. One-hundredth Yearbook of the National Society for the Study of Education* (pp. 205–238). Chicago, IL: University of Chicago Press.

Kurz, A., Elliott, S. N., Wehby, J. H., & Smithson, J. L. (2010). Alignment of the Intended, Planned, and Enacted Curriculum in General and Special Education and Its Relation to Student Achievement. *The Journal of Special Education*, *44*(3), 131–145. https://doi.org/10.1177/0022466909341196

Lewis, E. B., Rivero, A. M., Lucas, L. L., Musson, A. A., & Helding, B. A. (2021). Setting empirically informed content knowledge policy benchmarks for physical science teaching. *Journal of Research in Science Teaching, 58*(8), 1238–1277. https://doi.org/10.1002/tea.21709

Lontok, K. S., Zhang, H., & Dougherty, M. J. (2015). Assessing the Genetics Content in the Next Generation Science Standards. *PloS One*, *10*(7), e0132742–e0132742. https://doi.org/10.1371/journal.pone.0132742

Love, T. S., Napoli, M., & Lee, D. (2023). Examining pre-service elementary educators' perceptions of teaching science when integrated with poetry. *School Science and Mathematics, 123*(2), 42-53.

Lowell, B. R., & McNeill, K. L. (2023). Changes in teachers' beliefs: A longitudinal study of science teachers engaging in storyline curriculum-based professional development. *Journal of Research in Science Teaching*, *60*(7), 1457–1487. https://doi.org/10.1002/tea.21839

Lucas, K., Ginns, I., Tulip, D., & Waiters, J. (1993, April). *Science teacher efficacy, locus of control and self-concept of Australian pre service elementary school teachers.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.

Marco-Bujosa, L. M., McNeill, K. L., González-Howard, M., & Loper, S. (2017). An exploration of teacher learning from an educative reform-oriented science curriculum: Case studies of teacher curriculum use. *Journal of Research in Science Teaching, 54*(2), 141–168. https://doi.org/10.1002/tea.21340

McNeill, K. L., Marco-Bujosa, L. M., González-Howard, M., & Loper, S. (2018). Teachers' enactments of curriculum: Fidelity to Procedure versus Fidelity to Goal for scientific argumentation. *International Journal of Science Education*, *40*(12), 1455–1475. https://doi.org/10.1080/09500693.2018.1482508

Meirink, J., van der Want, A., Louws, M., Meijer, P., Oolbekkink-Marchand, H., & Schaap, H. (2020). Beginning teachers' opportunities for enacting informal teacher leadership: Perceptions of teachers and school management staff members. *European Journal of Teacher Education, 43*(2), 243–257. https://doi.org/10.1080/02619768.2019.1672654

Menon, D., & Sadler, T. D. (2016). Preservice Elementary Teachers' Science Self-Efficacy Beliefs and Science Content Knowledge. *Journal of Science Teacher Education*, *27*(6), 649–673. https://doi.org/10.1007/s10972-016-9479-y

Mensah, F. M., & Jackson, I. (2018). Whiteness as property in science teacher education. *Teachers College Record, 120*, 1–38.

Mentzer, G. A., Czerniak, C. M., & Struble, J. L. (2014). Utilizing program theory and contribution analysis to evaluate the development of science teacher leaders. *Studies in Educational Evaluation*, *42*, 100–108. https://doi.org/10.1016/j.stueduc.2014.03.003

Mentzer, G. A., Czerniak, C. M., & Brooks, L. (2017). An examination of teacher understanding of project based science as a result of participating in an extended professional development program: Implications for implementation. *School Science and Mathematics*, *117*(1-2), 76-86.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). San Francisco: Jossey-Bass.

Metz, S. (2013). Editor's Corner: Notes from the field editor. *The Science Teacher (National Science Teaching Association), 80*(3), 6.

Michaels, S., & O'Connor, C.. (2012). *Talk science primer*. Cambridge, MA: TERC.

Moore, W., & Esselman, M. (1992, April). *Teacher efficacy, power, school climate and achievement: A desegregating district's experience.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Mulholland, J., & Wallace, J. (1996). Breaking the cycle: Preparing elementary teachers to teach science. *Journal of Elementary Science Education, 8*(1), 17–38.

National Academy of Sciences, Engineering, and Medicine [NASEM]. (2015). *Science teachers' learning: Enhancing opportunities creating supportive contexts.* Washington, DC: The National Academy Press.

National Academy of Sciences, Engineering, and Medicine [NASEM]. (2018). *Design, selection, and implementation of instructional materials for the next generation science standards (NGSS): Proceedings of a workshop.* Washington, DC: The National Academy Press.

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States.* The National Academies Press. http://www.nextgenscience.org

Nixon, R. S., Smith, L. K., & Sudweeks, R. R. (2019). Elementary teachers' science subject matter knowledge across the teacher career cycle. *Journal of Research in Science Teaching, 56*(6), 707–731. https://doi.org/10.1002/tea.21524

No Child Left Behind Act of 2001, PL. 107-110, 20 U.S.C. 6319 (2002).

NRC (National Research Council). (1996). *National science education standards.* Washington, DC: National Academy Press.

NRC (National Research Council). (2004). *On evaluating curricular effectiveness: Judging the quality of K–12 mathematics evaluations.* Washington, DC: Mathematical Science Education Board, Center for Education.

NRC (National Research Council). (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas.* National Academies Press. https://doi.org/10.17226/13165

NSTA (National Science Teaching Association). (2012). *NSTA science content analysis form: Elementary science specialists or middle school science teachers.* Arlington, VA: NSTA.

NSTA (National Science Teaching Association). (2013). *Program: National conference on science education.* National Science Teaching Association.

NSTA (National Science Teaching Association). (2022). *K–12 science standards adoption.* https://ngss.nsta.org/About.aspx

O’Donnell, C. L. (2008). Defining, Conceptualizing, and Measuring Fidelity of Implementation and Its Relationship to Outcomes in K-12 Curriculum Intervention Research. *Review of Educational Research*, *78*(1), 33–84. https://doi.org/10.3102/0034654307313793

OpenSciEd. (n.d.). Funders. Retrieved November 5, 2023, from https://www.openscied.org/about/our-funders/

Palmisano, A. J. (2017) Does evaluation affect teacher self-efficacy: a qualitative investigation into the effects of evaluation on teachers' sense of efficacy. *Graduate Research Theses & Dissertations*. 2646. https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/2646

Pando, M., & Aguirre-Muñoz, Z. (2021). Case-based Instruction in Science Professional Development: Bilingual/ESL Teachers Reflect about Science Subject Matter Knowledge and Pedagogy. *Journal of Science Teacher Education*, *32*(3), 286–305. https://doi.org/10.1080/1046560X.2020.1819515

Penuel, W. R., Allen, A.-R., Deverel-Rico, C., Singleton, C., & Pazera, C. (2023). How Teachers' Knowledge of Curriculum Supports Partnering with Students in Their Science Learning. *Journal of Science Teacher Education, 34*(8), 861-882. https://doi.org/10.1080/1046560X.2023.2167508

Ramey-Gassert, L., Shroyer, M. G., & Staver, J. R. (1996). A qualitative study of factors influencing science teaching self-efficacy of elementary level teachers. *Science Education, 80*, 283–315.

Ravitch, D. (2020). *Slaying Goliath: the passionate resistance to privatization and the fight to save America's public schools*. Vintage Books, a division of Penguin Random House LLC.

Reiser, B. J. (2013, September). *What professional development strategies are needed for successful implementation of the next generation science standards.* Paper presented at the Invitational Research Symposium on Science Assessment. K-12 Center at ETS, Washington, DC.

Reiser, B. J., Novak, M., McGill, T. A. W., & Penuel, W. R. (2021). Storyline units: An instructional model to support coherence from the students' perspective. *Journal* *of Science Teacher Education, 32*(7), 805–829. https://doi.org/10.1080/1046560X.2021.1884784

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246. https://doi.org/10.3102/00346543075002211

Rhodes, R. E., & Blanchard, C. M. (2007). What do confidence items measure in the physical activity domain? *Journal of Applied Social Psychology, 37*, 759-774.

Rice, D. C., & Roychoudhury, A. (2003). Preparing more confident preservice elementary science teachers: One elementary science methods teacher's self-study. *Journal of Science Teacher Education, 14*, 97–126.

Riggs, I. (1995, April). *The characteristics of high and low efficacy elementary teachers.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco.

Riggs, I. M., & Enochs, L. (1990). Toward the development of an elementary teachers' science teaching efficacy belief instrument. *Science Education, 74*, 625–637.

Riggs, L, & Jesunathadas, J. (1993, April). *Preparing elementary teachers for effective science teaching in diverse settings.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.

Robertson, A. G. (2022). *The Impact of Analyzing Science Standards on Elementary Teacher Science Self-Efficacy* (Order No. 29169330). Available from ProQuest Dissertations & Theses A&I. (2679755392). http://libproxy.unl.edu/login?url=https://www.proquest.com/dissertations-theses/impact-analyzing-science-standards-on-elementary/docview/2679755392/se-2

Ross, J. A. (1992). Teacher efficacy and the effect of coaching on student achievement. *Canadian Journal of Education, 17*(1), 51 -65.

Rumelhart, D. E. (1980). Schemata: the building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–49). Hillsdale: Lawrence Erlbaum Associates.

Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal, 50*(5), 1020–1049. https://doi.org/ 10.3102/0002831213477680

Sandholtz, J. H., & Ringstaff, C. (2014). Inspiring Instructional Change in Elementary School Science: The Relationship Between Enhanced Self-efficacy and Teacher Practices. *Journal of Science Teacher Education*, *25*(6), 729–751. https://doi.org/10.1007/s10972-014-9393-0

Shapiro, L. J., & Kraus, R. V. (2023). The NGSS and the Historical Direction of Science Education Reform. *Science Educator, 29*(2), 63–74.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*, 1–22.

Simpson, J.C. (2021). Fostering Teacher Leadership in K-12 Schools: A Review of the Literature. *Performance Improvement Quarterly*, *34*(3), 229–246. https://doi.org/10.1002/piq.21374

Sims, S., & Fletcher-Wood, H. (2021). Identifying the characteristics of effective teacher professional development: a critical review. *School Effectiveness and School Improvement*, *32*(1), 47–63. https://doi.org/10.1080/09243453.2020.1772841

Smith, L. K., Nixon, R. S., Sudweeks, R. R., & Larsen, R. A. (2022). Elementary teacher characteristics, experiences, and science subject matter knowledge: Understanding the relationships through structural equation modeling. *Teaching and Teacher Education*, *113*, 103661-. https://doi.org/10.1016/j.tate.2022.103661

Starr, C. R., Hunter, L., Dunkin, R., Honig, S., Palomino, R., & Leaper, C. (2020). Engaging in science practices in classrooms predicts increases in undergraduates' STEM motivation, identity, and achievement: A short-term longitudinal study. *Journal of Research in Science Teaching*, *57*(7), 1093–1118. https://doi.org/10.1002/tea.21623

Stroupe, D., & Hancock II, J. B. (2022). Examining mentor teachers' critical pedagogical discourses and participation in an era of changing science standards and pedagogies. *Teaching and Teacher Education*, *109*, 103558-. https://doi.org/10.1016/j.tate.2021.103558

Sullivan-Watts, B., Nowicki, B., Shim, M., & Young, B. (2013). Sustaining reform-based science teaching of preservice and inservice elementary school teachers. *Journal of Science Teacher Education*, *24*(5), 879–905.

Tashakkori, A., & Teddlie, C. (2010). SAGE Handbook of Mixed Methods in Social & Behavioral Research. In *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed.). SAGE Publications, Inc. https://doi.org/10.4135/9781506335193

Taylor, M. W. (2013). Replacing the “teacher-proof” curriculum with the “curriculum-proof” teacher: Toward more effective interactions with mathematics textbooks. *Journal of Curriculum Studies*, *45*(3), 295–321. https://doi.org/10.1080/00220272.2012.710253

Tilgner, P. J. (1990). *A qualitative analysis of preservice elementary teachers' conceptions of heat transfer and temperature*. Thesis (Ed.D.)-University of Nebraska-Lincoln, 1990.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher Efficacy: Its Meaning and Measure. *Review of Educational Research*, *68*(2), 202–248. https://doi.org/10.3102/00346543068002202

United States. National Commission on Excellence in Education. (1983). *A nation at risk: the imperative for educational reform*. Washington, D.C., The National Commission on Excellence in Education.

Uppendahl, K. A. (2020). *How the Amplify Science curriculum impacts 4th grade students and the teacher* [Master's thesis, Montana State University].

van Driel, J. H., Verloop, N., & de Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, *35*(6), 673–695. https://doi.org/10.1002/(SICI)1098-2736(199808)35:6<673::AID-TEA5>3.0.CO;2-J

Velthuis, C., Fisser, P., & Pieters, J. (2014). Teacher Training and Pre-service Primary Teachers' Self-Efficacy for Science Teaching. *Journal of Science Teacher Education*, *25*(4), 445–464. https://doi.org/10.1007/s10972-013-9363-y

Watson, S. (1991). *A study of the effects of teacher efficacy on the academic achievement of third-grade students in selected elementary schools in South Carolina.* Unpublished doctoral dissertation, South Carolina State College, Orangebury. (University Microfilms No. UMI 9230552)

Welch, W. W. (1979). Twenty years of science curriculum development: A look back. *Review of Research in Education, 7,* 282–308.

Wenner, J., & Kittleson, J. (2018). Focused video reflections in concert with practice-based structures to support elementary teacher candidates in learning to teach science. *Journal of Science Teacher Education, 29*(8), 741–759.

Wiles, J. (2008). *Leading Curriculum Development.* Corwin Press.

Williams, D. M. (2010). Outcome Expectancy and Self-Efficacy: Theoretical Implications of an Unresolved Contradiction. *Personality and Social Psychology Review*, *14*(4), 417-425. https://doi.org/10.1177/1088868310368802

Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research*, *72*(2), 131–175. https://doi.org/10.3102/00346543072002131

Woolfolk, A. E., Rosoff, B., & Hoy, W. K. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and Teacher Education, 6,* 137-148.

Yaşar, C. G., & Aslan, B. (2021). Curriculum theory: A review study. *International Journal of Curriculum and Instructional Studies, 11*(2), 237-260. DOI: 10.31704/ijocis.2021.012

Yesilyurt, E., Deniz, H., & Kaya, E. (2021). Exploring sources of engineering teaching self-efficacy for pre-service elementary teachers. *International Journal of STEM Education*, *8*(1), 1-15.

York-Barr, J., & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research*, *74*(3), 255–316. https://doi.org/10.3102/00346543074003255

Zeichner, K. (2010). Rethinking the Connections Between Campus Courses and Field Experiences in College- and University-Based Teacher Education. *Journal of Teacher Education*, *61*(1–2), 89–99. https://doi.org/10.1177/0022487109347671

APPENDICES

Appendix A: Consent form

University of Nebraska Lincoln

Recruitment/Invitation Email

Dear 3rd to 5th grade science teacher:

I am conducting a research study on whether the implementation of a phenomena-based science curriculum relates to science teaching self-efficacy and outcome expectations of elementary science teachers. Participation in this study will require approximately 30 minutes of time to complete an initial online survey. If you are selected for a follow-up interview this will require approximately 45 minutes of additional time. Participation will take place online at your convenience for the initial survey, and via Zoom video conferencing after 3:15 p.m. if you are selected for a follow-up interview. All study participation must be completed outside of district contract hours.

If you are interested in becoming a study participant and certify that you are a 3rd to 5th grade science teacher, please click on the following link to the initial Qualtrics survey: LINK

Reasonable steps will be taken to protect the privacy and the confidentiality of your study data; however, in some circumstances we cannot guarantee absolute privacy and/or confidentiality. Research records will be stored on a password-protected device. Records will only be seen by the research team and/or those authorized to view, access, or use the records during and after the study is complete.

If you have questions about this project, you may either contact the primary investigator, Wes Sliger at wsliger2@unl.edu or the faculty advisor for this project, Dr. Lawrence Scharmann, at lscharmann2@unl.edu.

You will be entered in a lottery to receive a $10 Visa eGift card for participating in this study. You will have a 10% or greater chance of winning, and winners will receive an email notification with a link to the eGift card upon completion of the study.

There are no known risks involved in this research.

If you have any questions, please let me know.

Primary Investigator
Wes Sliger
wsliger2@unl.edu
Faculty Advisor
Lawrence Scharmann
lscharmann2@unl.edu

University of Nebraska Lincoln

WEB-BASED INFORMED CONSENT

IRB Project ID #: 23410

Participant Study Title: Elementary science teacher self-efficacy in the context of NGSS-aligned curriculum enactment: an explanatory sequential mixed methods study

The purpose of this research is to use social cognitive theory to assess whether the implementation of a phenomena-based science curriculum relates to science teaching self-efficacy and outcome expectations of elementary science teachers. An explanatory sequential case selection variant mixed methods design will be used that involves collecting quantitative data first and then explaining the quantitative results with in-depth qualitative data. In the first, quantitative phase of the study, Science Teaching Efficacy Belief Instrument A (STEBI-A) data will be collected from 3rd to 5th grade elementary school teachers that teach science at your school district to assess their current science teaching self-efficacy and outcome expectations. The second phase, qualitative semi-structured interviews will be conducted in a multiple case study as a follow-up to help explore the experiences of four 3rd to 5th grade elementary science teachers representing different combinations of science teaching self-efficacy and outcome expectations at your school district. If you are a 3rd to 5th grade elementary science teacher, you may participate in this research.

Participation in this study will require approximately 30 minutes of time to complete an initial online survey. If you are selected for a follow-up interview this will require approximately 45 minutes of time. Participation will take place online at your convenience for the initial survey, and via Zoom video conferencing after 3:15 p.m. if you are selected for a follow-up interview. All study participation must be completed outside of district contract hours.

Society may benefit by increasing knowledge about teacher preparation and training. Science teaching reforms such as the ones described within NGSS and NGSS-aligned curricula are relatively new and complex, and it can be difficult to understand how teachers understand and use NGSS-aligned curriculum materials. This study may help increase general understanding of the learning community of an elementary science teacher.

You will be entered in a lottery to receive a $10 Visa eGift card for participating in this study. You will have a 10% or greater chance of winning, and winners will receive an email notification with a link to the eGift card upon completion of the study.

Reasonable steps will be taken to protect the privacy and the confidentiality of your study data; however, in some circumstances we cannot guarantee absolute privacy and/or confidentiality. Research records will be stored on a password-protected device. Records will only be seen by the research team and/or those authorized to view, access, or use the records during and after the study is complete.

If you have questions about this project, you may either contact the primary investigator, Wes Sliger at wsliger2@unl.edu or the faculty advisor for this project, Dr. Lawrence Scharmann, at lscharmann2@unl.edu.

If you have questions about your rights or complaints about the research, contact the UNL Institutional Review Board (IRB) at (402)472-6965 or irb@unl.edu.

You can decide not to be in this research study, or you can withdraw at any time before, during, or after the research begins for any reason. Deciding not to be in this research study or deciding to withdraw will not affect your relationship with the investigator, the University of Nebraska-Lincoln, or your school district. You will not lose any benefits to which you are entitled.

You are voluntarily making a decision whether or not to participate in this research study. By clicking on the I Agree button below, your consent to participate is implied. You should print or save a copy of this page for your records.

Appendix B: Initial Quantitative Survey

Initial Quantitative Survey

Demographics Section

First Name: ____

Last Name: ____

Email Address: ____

For demographic purposes, which race or ethnicity best describes you?

- American Indian or Alaskan Native
- Asian/Pacific Islander
- Black or African American
- Hispanic
- White/Caucasian
- Multiple ethnicity/Other (Please specify)
- I prefer not to answer

For demographic purposes, what is your gender identity?

- Male
- Female
- I prefer not to answer
- I identify in another way (Self-describe below)

Which grade level(s) are you teaching during the current school year? Select all that apply.

- 3rd Grade
- 4th Grade
- 5th Grade

Counting the current school year, how many years have you been a K-12 educator, including part-time teaching?

- Dropdown Response: "1" to "40 or more"

Counting this school year, how many years have you taught science in a K-12 setting, including part-time teaching?

- Dropdown Response: "1" to "40 or more"

Do you enjoy science? (1 = "Not at all.", 5 = "I love science.")

- 1
- 2
- 3
- 4
- 5

What is the highest degree or level of school you have completed? *If currently enrolled, highest degree received.*

- High school diploma or GED
- Associate's degree
- Bachelor's degree
- Master's degree
- Professional degree
- Doctorate degree

Which school do you teach in?

- Dropdown Response: "A" to "Z"

What is the average class size of your science classes?

- 0-10
- 11-15
- 16-20
- 21-25
- 25-30
- 30 or more

How much time do you spend per day teaching science during a typical science unit?

- 0-4 minutes
- 5-19 minutes
- 20-34 minutes
- 35-50 minutes
- More than 50 minutes

Which Amplify Science units, if any, have you taught? Select all that apply.

- I have not taught using Amplify Science units.
- 3rd Grade: Balancing Forces
- 3rd Grade: Environments and Survival
- 3rd Grade: Weather and Climate
- 4th Grade: Energy Conversions
- 4th Grade: Waves, Energy, and Information
- 4th Grade: Earth's Features
- 5th Grade: Modeling Matter
- 5th Grade: The Earth System
- 5th Grade: Ecosystem Restoration

Describe any prior professional development sessions you've participated in regarding Amplify Science curricular materials. This may include who provided the training, the length of the training, the topics covered or skills learned, etc.

____

____

____

____

Appendix C: Science Teaching Efficacy Belief Instrument A (STEBI-A)

Science Teaching Efficacy Belief Instrument - A (STEBI-A)

Directions: Please indicate the degree to which you agree or disagree with each statement below by selecting the appropriate description under each statement.

1. When a student does better than usual in science, it is often because the teacher exerted a little extra effort.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

2. I am continually finding better ways to teach science.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

3. Even when I try very hard, I do not teach science as well as I do most subjects.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

4. When the science grades of students improve, it is most often due to their teacher having found a more effective teaching approach.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

5. I know the steps necessary to teach science concepts effectively.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

6. I am not very effective in monitoring science experiments.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

7. If students are underachieving in science, it is most likely due to ineffective science teaching.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

8. I generally teach science ineffectively.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

9. The inadequacy of a student's science background can be overcome by good teaching.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

10. The low science achievement of some students cannot generally be blamed on their teachers.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

11. When a low achieving child progresses in science, it is usually due to extra attention given by the teacher.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

12. I understand science concepts well enough to be effective in teaching elementary science.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

13. Increased effort in science teaching produces little change in some students' science achievement.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

14. The teacher is generally responsible for the achievement of students in science.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

15. Students' achievement in science is directly related to their teacher's effectiveness in science teaching.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

16. If parents comment that their child is showing more interest in science at school, it is probably due to the performance of the child's teacher.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

17. I find it difficult to explain to students why science experiments work.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

18. I am typically able to answer students' science questions.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

19. I wonder if I have the necessary skills to teach science.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

20. Effectiveness in science teaching has little influence on the achievement of students with low motivation.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

21. Given a choice, I would not invite the principal to evaluate my science teaching.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

22. When a student has difficulty understanding a science concept, I am usually at a loss as to how to help the student understand it better.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

23. When teaching science, I usually welcome student questions.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

24. I don't know what to do to turn students on to science.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

25. Even teachers with good science teaching abilities cannot help some kids learn science.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

Appendix D: Qualitative Interview Protocol

Qualitative Interview Protocol

Elementary science teacher self-efficacy in the context of NGSS-aligned curriculum enactment: an explanatory sequential mixed methods study, 2024 | IRB#[project ID 23410]

Interview:

Aim: The goal of this interview is to allow the participant to freely share their ideas, interests, language, understanding, and thoughts related to their experiences as an elementary science teacher during a curriculum implementation experience.

Method: Semi-structured interview

Introduction: The principal investigator will review the purpose of the research study with the participant in conversational language, including the information below.

Thank you for following up the survey with an interview. As you know, the purpose of this research is to understand how elementary science teachers are navigating their science curriculum implementation experiences. I have some follow-up questions that have come to me as I've made some observations and I'd love your insight to help me clarify and increase my understanding.

Interview questions: The interview questions are intended to be open-ended and shared in a natural conversational manner. Not all questions or probes need be asked and specific language of the question may be altered for conversational flow. The purpose of this interview is to provide clarification, elaboration, and accuracy to the observations made during the quantitative phase of the study. As such, a semi-structured approach will be used to ask the participant questions that lend insight to observations previously made by the researcher. Each set of questions may be tailored to the participant based on their role.

***NOTE:* Additional questions may be asked to allow the participant to further explain or expand on a previous response.**

The questions for this interview are not fully predetermined, but some examples of the types of open-ended questions that may be asked are provided below:

Prompt (General)
Tell me about Amplify Science. Response:
In your initial survey, you described your prior experience and professional development/training with Amplify Science. (Read or paraphrase response as needed.) Can you please elaborate on your response? Response:
In your initial survey, you rated your enjoyment of science as a ___ out of 5, with 1 being "Not at all" and 5 being "I love science." Can you tell me about what contributed to this rating? Response:

Open-ended interview questions based on specific STEBI-A prompts and scores for the PSTE subscale:

STEBI-A Question #	Prompt (PSTE)
2	In what ways have your science teaching practices changed since the beginning of the curriculum implementation process?
3	How would you compare your science teaching ability to the other subjects you teach? Why?
8	Describe the components of an ideally planned science lesson or unit. Probe: How does your own science teaching compare to what you described as the ideal science lesson?
12	Has your curriculum implementation experience changed your understanding of any of the science concepts you teach? How so?
19	What skills do you believe an effective science teacher needs?
21	**For low scores:** In your survey, you responded that, given the choice, you would not invite your principal to evaluate your science teaching.

	What kinds of reservations would you have with your principal evaluating your science teaching?
22	How do you address situations in which students are not understanding the science concepts that you've taught?
23	How do you respond when student ask questions during science class?
24	What kind of strategies do you use to build students' interest in science? Probe: Is that a strategy that you've always used, or was it embedded in Amplify Science materials?

Open-ended interview questions based on specific STEBI-A prompts and scores for the STOE subscale:

STEBI-A Question #	**Prompt (STOE)**
1	What impact, if any, does the effort you exert while planning and teaching science have on student achievement in science?
9	Can you describe a time when a low-achieving student in science performed well due to your science teaching?
10	What do you think are the most important reasons that certain students may not be attaining high achievement in your science class compared to their achievement in other subject areas?
11	How would you compare the science achievement of your students before and after implementing Amplify Science?
16	Describe how teaching with Amplify Science units has (or has not) impacted your students' interest in science.

20	Describe how teaching with Amplify Science units has (or has not) impacted student motivation.

Additional Probes

1. In your initial survey, you responded that (summarize PSTE/STOE responses: e.g., “I strongly agree that I am continually finding better ways to teach science.”)
    a. Is this still the case?

    b. Can you provide an example or explanation that you believe contributes to this rating?

2. I noticed that you reported one viewpoint during your STEBI-A survey, and stated a different/opposing viewpoint on a similarly phrased question. Can you tell me more about the differences?

3. I noticed that you reported one viewpoint during your STEBI-A survey, and stated a different/opposing viewpoint during our current interview. Can you tell me more about these differences?

4. During the interview, you stated “ _____ “. Do you have anything to add to or clarify this statement?