Title
Engaging High School Science Teachers in Personalized Professional Learning: A Design Development Study

Author
Reinhardt, Thomas Arthur

Publication Date
2018

Peer reviewed|Thesis/dissertation

**Engaging High School Science Teachers in Personalized Professional Learning:
A Design Development Study**

by

Thomas Arthur Reinhardt

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Education

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Heinrich Mintrop, Chair
Professor Michael Austin
Professor Michelle Wilkerson

Summer 2018

Engaging High School Science Teachers in Personalized Professional Learning: A Design Development Study

by

Thomas Arthur Reinhardt

Abstract

Engaging High School Science Teachers in Personalized Professional Learning:

A Design Development Study

by

Thomas Arthur Reinhardt

Doctor of Education

University of California, Berkeley

Professor Heinrich Mintrop, Chair

The current educational climate, featuring standards-based reforms and reliance on summative assessment and managerial control has strained the professional work of teachers by simultaneously demanding compliance with reform effort and control of classrooms while expecting in-depth professional learning to shift their pedagogical strategies to provide deeper cognitive experiences for students. Currently, the managerial and high-stakes accountability forces on teachers are prevailing, leaving teachers in a state of low motivation with limited time and resources to further their pedagogical content knowledge. In addition, the educational system lacks the resources and technical ability to provide high quality professional learning experiences for teachers. This design development study attempts to motivate high school science teachers into a state of learning through classroom experimentation and reflection on student learning.

This research involved seven science teachers at high-poverty, urban high school with a large portion of English learners and student of color. During a series of seven 90-minute workshops, the participating teachers created personalized professional learning plans based on their self-identified teaching characteristics and goals for their own learning. Subsequently, they attempted pedagogical strategies and presented artifacts of students learning to their colleagues. Being highly contextualized research, design development studies provide the great insight into the specific context where implemented but also illuminate deep challenges facing our educational system, such as motivating teachers to learn advanced pedagogical strategies in response to student learning.

Results from participant responses to structured and semi-structured pre/post interview questions combined with analysis of process data collected during the implementation of the workshop series yielded salient trends in the participating teachers' learning. While most participants attempted new strategies and began tracking learning by reviewing student artifacts, the pedagogical strategies attempted focused on general engagement strategies rather than deeper science pedagogical strategies. This suggested a willingness to try new strategies and a possible need for explicit examples of high-quality science teaching strategies during professional

development experiences. Additionally, the teachers deepened their reflective conversation throughout the project and began to focus more on their learning and control of classroom experiences over the perceived deficits of the student population. However, the language used to describe teaching and learning did not advance to the desired technical level, suggesting a need for additional time repeating learning cycles with a learning experience to bolster technical analysis of learning. This design development study reinforces much of the theorized suggestions for high quality professional learning in science education. Questions remain regarding sustainability and the effects of prolonged and engaging teacher learning opportunities.

Dedication

I dedicate this dissertation to my lifelong teachers, Art and Sue Reinhardt. As my parents, your most important gift continues to be a love of learning. Thank you for your endless support, encouragement, and love.

Table of Contents

List of Tables

Acknowledgments

This dissertation would not have been possible without the support of my professors, colleagues, family, and friends. First, I would like to acknowledge and express deep gratitude for Rick Mintrop. His feedback, critical thought, and willingness to engage me in deep discussions about education and life in general lifted my experience by supporting my development as a leader in education. Additionally, my dissertation committee, Michelle Wilkerson and Michael Austin, provided me with insightful feedback and support during this process to expand my thinking about educational challenges. I would also like to appreciate the eighth cohort of the Leadership of Educational Equity Program (LEEP). Chin, Dennis, Jamila, Julie, Karin, and Peter supported my growth in educational leadership in countless ways. You are inspiring examples of educational leaders. Additionally, I acknowledge the professors of the UC-Berkeley Graduate School of Education for exemplifying academic excellence by offering numerous experiences to broaden my exposure to educational research.

A key component of LEEP is understanding the powerful intersection between educational leadership and research. I am thankful for the support of colleagues and leaders who have believed in me and helped me navigate this intersection, especially Liz Baham, Caleb Cheung, Anthony Cody, Phil Cotty, Claudio Vargas, Rebecca Cheung, Liz Martin, the OUSD Science Department, and Leslie Flynn.

Finally, I could not have finished this work without the loving support of my friends and family. I acknowledge my parents, Art and Sue, for their unwavering support; my partner, Alice for holding me up and believing in me, especially during the challenges of my final year of research and writing; my friend, Steven, for always being there to listen and support me; and my Yummy friends who nurture my creativity and teach me how to be a supportive and caring member of a community.

Introduction

The Common Core State Standards (CCSS) and Next Generation Science Standards (NGSS) necessitate an overhaul in pedagogical practices and course content unlike any previous standards-based reform initiative (Santos et. al., 2012). The lofty demands implied in the new standards create an obvious teacher learning need to enable shifts in pedagogical practice. Desimone's (2009) theoretical writing provides a causal chain of teacher learning through professional development leading to changes in classroom practices with distal effects on student achievement. With the integration of practices, such as argumentation supported by evidence within content areas, the demands on teacher learning are daunting, because they are reconciling a shift in the content and expanding the curricular scope to include ways of practicing science. Most often, institutions rely on professional development (PD) as a means to instill instructional changes. Few PD studies are based on anything but willing volunteers (Borko, 2004, Desimone, 2009). There is hope for PD as an effective venue for change when it provides meaningful, engaging, and sustained focus on teaching practice (Desimone, 2009; Darling-Hammond, 2009; Wilson et. al., 2015). Given the need for teachers to learn content and corresponding pedagogical practices for the new standards, careful attention must be given to attitudes held by teachers, regarding learning.

The intense need for teacher learning is further complicated in urban schools where years of No Child Left Behind policies have left teachers focused on high -stakes, summative tests as accountability metrics, which in turn foster knowledge transmission pedagogies for test preparation (Au, 2007; Ball, 2000; Tienken & Zhao, 2013). Additionally, the societal view that schools are not serving students cause many teachers to conclude their students are incapable of learning advanced material and externalize the root of educational woes, such as blaming student skill deficits or demands from administrators (Mintrop, 2004). This mindset must be overcome in order to advance difficult pedagogical practices. Teachers will need to learn to implement carefully designed instructional experiences for students as they transition to cognitively complex lessons involving problem solving, close reading and other advanced literacy skills, computational thinking, and evidence-based argumentation, etc. (NGSS Lead States, 2013; National Governors Association Center for Best Practices, 2010). Given the challenges of providing cognitively demanding curriculum in urban settings, the complexity of required teacher learning transcends a simple need for increased skills and knowledge to a teacher shift in attitude and motivation to be able to feel efficacious with new styles of teaching in historically underserved populations.

Additionally, current evaluation systems often rely on punitive extrinsic motivation through teacher evaluation. As seen in multiple San Francisco Bay Area school districts, teacher success is often measured by simply meeting minimum laboratory science requirements with simplified activities lacking inquiry components or student learning. To be seen as successful, teachers focus on test scores thus reducing the complexity of the curriculum and working towards superficial markers of classroom success, which often ignore deeper learning on behalf of the students (Achinstein & Ogawa, 2006; Ball, 2000; Au, 2007).

Due to motivational factors such as influence, autonomy, and discourse regarding professional skill and preparation, the current teaching force exhibits signs of de-professionalization, characterized by reduced autonomy in decision making and compliance-driven reform affecting teacher learning and behavior, as teachers have reduced control of their curricula and a focus on basic skills development (Van Veen, 2008). Recent curriculum reforms position teachers as mere implementers of technical solutions (Mintrop & Trujillo, 2007; Van Veen, 2008). Curricular solutions for teachers vary from scripted lessons to mandatory methodologies. Most of these solutions reduce the teaching craft to a set of instructions most people could follow. This implies a reduction in the amount of skill and knowledge needed by teachers (Van Veen, 2008). Although responding to student learning needs is paramount to the role of teachers, this de-professionalized view of teachers reduces teacher morale and desire to work within the educational system (Achinstein & Ogawa, 2006; Finnigan & Gross, 2007). As a response to these beliefs regarding teacher learning and purpose, many schools and districts resort to simplified, one-time, professional learning workshops that seek technical solutions when many teachers would prefer deeper learning opportunities (Darling-Hammond, 2009). This state of de-professionalization is apparent in the focal district of this dissertation, hereafter named Bay Vista Unified to protect anonymity, where disjointed PD workshops focus on discrete skills and administrators reinforce compliance by both teachers and students. During teacher-led professional learning community time, the evidence of the described conditions appears as teachers focus on complaints regarding students and administration, rarely discussing educational practices.

Given the prolonged de-professionalization of teaching, teachers are feeling uncertain which leads to a learning stagnation, indicated by maintenance of complacent attitudes with regard to learning and status quo teaching and learning, and avoidance of risk (Thoonen et. al., 2011). Additionally, a clear message has been sent to teachers; the solutions to educational problems are easy and require the implementation of technical solutions, which purport uniform solutions to complex problems with student and teacher learning; teachers are not to be trusted with student learning. Instead of working to improve their teaching practice, they become complacent with instruction reflecting low cognitive demand on students while maintaining rigid control signified by student compliance, often externalizing the causes of low student achievement (Mintrop, 2004). The aforementioned factors have a cumulative negative effect on teachers' willingness to take risks and take on an attitude of professional learning towards teaching. The problem is many teachers have stagnated in their professional learning at a time requiring great growth in instructional practices (National Academies of Sciences, Engineering, and Medicine [NASEM], 2015). How can teachers overcome stagnation and learn to practice new science pedagogical content knowledge outlined in the literature (NASEM, 2015)? Additionally, how can learning be sustained to counteract the long history of implementation and compliance driven reforms, including knowledge level tests for accountability and management-focused classroom evaluations? Additionally, how can issues of de-professionalization of teaching be addressed to improve motivation and self-efficacy issues for teachers? The following thesis presents the pertinent knowledge base, theoretical framework, and intervention design to work directly with teachers as learners, attempting to address the learning stagnation established by societal and educational norms and constraints. This theoretical framework undergirds a design development study, which will attempt to open teachers to learning and improvement.

Chapter One: Problem of Practice

Within this chapter, the foundational aspects of the design development are constructed. The chapter begins with a detailed assessment of the needs of science teachers in a generalized format and more specific conditions of teacher learning at Bay Vista High. Second, the problematic behaviors are analyzed as a diagnostic endeavor preparing the researcher-practitioner to outline the desired state. The chapter closes by explicitly naming the problem of practice addressed by this intervention design.

Needs Assessment

Design Development Studies rely on careful knowledge of how a problem of practice is situated within the specific context where the intervention takes place. Due to the specificity of the design, a thorough analysis of the context, a needs assessment, must occur. During the initial planning phases of this design, several sites were contacted as potential research locations. As the researcher-practitioner, I originally contacted schools within the district I was working, even implementing the first half of the design at one school before being derailed by school leadership and bureaucratic constraints on teachers' time. Due to the difficulty in establishing a research site, the needs assessment is occurring in two key stages: first, surveys and observations from a nearby school district with similar teacher and student demographics to the research site; and second, interviews and observations of the teachers and teacher leaders at the research site. This provides an understanding of the needs of teachers as a whole and the specific needs of teachers participating in the study.

Teacher leader survey. An initial needs assessment was conducted with 25 teacher leaders, each from a different school site in a district similar to the one where I eventually carried out the project. The teacher leaders are uniquely positioned to report on the disposition of teachers, as they are responsible for conducting site-based professional development and professional learning communities. Two group interviews, one for high school and another for middle school teachers, were conducted with the teacher leaders to ascertain the challenges faced by teachers. The results of the interview were categorized and labeled. Two patterns of response emerged. Both groups of teacher leaders identified teacher mindset as a major obstacle. In this category, they included resistance to changing strategies and their current teaching styles. They also cited low student skills, compliance, and motivation and the need to teach more content rather than the skills-based standards of the NGSS as reasons not to shift to a more cognitively demanding teaching style.

The second category described the needs of teachers to shift their mindset. The teacher leaders suggested a lack of exemplary NGSS-aligned strategies to demonstrate possible pedagogical changes as a major obstacle to convincing teachers to shift their mindset. They also alluded to a need for more intense time devoted to professional development and sense-making of the NGSS. This brief assessment of needs in the demographically similar district contributes to the more pointed needs assessment that follows.

Classroom observations. After I conducted many observations and interviews with teachers who were identified as struggling by their principals, the following patterns were observed. Regarding the implementation of NGSS pilot curriculum and assessments, teachers expressed frustration, citing two primary causes. First, teachers say that the students are not prepared for the open-ended learning and critical thinking that NGSS demands. Second, teachers mention the need to reproduce typical lecture-style college education in their own classrooms. Under these auspices, teachers stress the need for fact-rich knowledge transmission over the time spent on inquiry and scientific practices. Additionally, several teachers describe reform fatigue resulting from frequently shifting school-wide goals and implementation of new standards.

During classroom observations, I observed the teachers delivering the following types of instruction. Many were seen providing lectures, worksheets, and step-by-step labs that requiring minimal critical thought or explanation. In a few instances, students seemed to participate in disjointed activities, staying busy the entire time. A final pedagogical strategy, seen in one classroom involved students with almost no guidance working on a research project. Although teachers should utilize various pedagogical strategies, these patterns are interesting to note because of the limited cognitive demand and sense-making opportunities for students. Prior to the initiation of the intervention design, a final needs assessment was conducted at Bay Vista High School.

Pre-intervention conversations with school leaders. Given the quick turnaround between determining a suitable research site and beginning the design implementation, unstructured interviews with Bay Vista district and school leadership were used to characterize the state of teacher learning at Bay Vista High School. The instructional leader at the district level had one primary focus: to increase student academic engagement. Although little specific information regarding Bay Vista High School was revealed, this focus was telling. Initially, it is, on the surface, an incredible worthwhile goal for a district. However, implementation of such a goal falls on the school site administrators who frequently use whole school professional learning workshops with sound strategies and little follow through. Bay Vista High School experienced this as the leadership hired a consultant to provide training in Kagan strategies, a set of generic classroom teaching strategies to foster successful collaboration and discussion. While Kagan strategies have demonstrated success for student engagement, the strategies alone do not produce connections to the classroom subject matter. Continued teacher learning through experimentation is likely necessary for the successful use of Kagan strategies, applied to science education.

During conversations with the Bay Vista High School instructional teacher leader and the science department chair, concerns arose regarding the science department's preparedness to implement NGSS-aligned strategies and continue to learn how to use those strategies to maximize benefit for students. The department meetings and teacher learning interactions were characterized as "unfocused", "rarely discussing deeper science instruction and NGSS implementation", and "hour long complaint sessions about students and administration with the occasional sharing of a good science activity" by the teacher leaders interviewed. These conversations indicated the presence of externalization of blame for instructional quality and limited focus on learning for improvement. The following observations and interviews further illuminated these challenges.

Pre-intervention needs assessment. During the weeks prior to the implementation study, classroom observations were conducted. Audio recordings and observational scripts were collected for 20 to 45 minutes of classroom teaching. The scripts were coded for indicators of pedagogical and lesson implementation style, using the Teacher Typology Matrix (TTM) (Appendix A). Additionally, a short post conference semi-structured interview helped determine how participants reflected on their own teaching decisions and styles. The observation and interview with the Bay Vista teachers served to analyze differences between teachers' self-reported styles and learning needs versus observed behaviors and learning needs proposed in this study.

Classroom observation pre-assessment. Although more frequent and longer observations would provide more accurate assessment of the teachers' pedagogical styles, the 30-minute observations provide a snapshot of the participating teachers' instruction and a foundation for informational interviews and design development activities. Table 1 illustrates each participants' observed levels within the four areas of teaching style utilized as focal points for teacher professional learning trajectories. Each category is characterized in the Teacher Typology Matrix (TTM) (Appendix A).

Following the observations of teachers for the Bay Vista High School needs assessment, several patterns emerged. First, activities taught in most of the classrooms rely on the transfer of knowledge or practice of simple skills. Students were not observed to be using higher order thinking skills. Additionally, very little evidence of NGSS alignment to grade-level appropriate practice and concepts was observed. In two different instances, students were building models that served only to reproduce images from the textbook. In other instances, worksheets and guided notes were used to deliver information. Many of the lesson segments seemed to focus on the facts of science with limited connection to previous learning and students' lives outside of the science classroom. Finally, teachers were performing limited or cursory checks for understanding in the classroom. For example, worksheets were scanned for correct information or whole class discussion, involving minimal student response, were used to determine the level of student learning. To deepen an understanding of these patterns, short structured interviews were conducted with the teachers following each observation.

Table 1: Needs assessment teacher observation summaries

Participant	Teaching Style	Activity Coherence	Relevance to Students	Responsiveness to Students
Claudia	**Emerging** Knowledge reinforcement with models	N/A	**Initial** Modeling only relevant to science facts	**Emerging** Teacher circulating to check answers, unaware of struggling students
Emily	**Initial** Students cut and paste models according to directions	**Initial** Transitions from building a molecule to lecture on different reactions	**Initial** Model and reactions not contextualized	**Initial** Students are completing task without engaging
Kathy	**Initial** Students copying notes from the screen	N/A	**Emerging** Teacher names ways the concept is seen in the real world	**Initial** One whole group check for understanding with a few students calling out answer
Marcus	**Emerging** Students complete laboratory activity with limited inquiry opportunity, primarily reinforces previous learning	N/A	**Emerging** Students are interested in the procedural aspects of the activity	**Emerging** Some students ask and have questions answer, many students simply copy work of other group members
Nancy	**Emerging** Students complete laboratory activity with limited inquiry opportunity, primarily reinforces previous learning	**Developing** Laboratory builds on questions from opening activity	**Emerging** Activity is loosely connected to a real-world phenomenon	**Emerging** Students have the opportunity to respond to questions and student work is checked at the end of class
Sara	**Emerging** Students working in small groups on a worksheet	N/A	**Initial** Students producing answers to scientific questions with no context	**Emerging** Few students respond to the check for understanding, most groups not working together effectively
Tony	**Emerging** Students work in small groups to revise ideas	N/A	**Initial** Questions are about abstract science concepts	**Emerging** All students responding to questions

Note: An entry of "N/A" indicated there was insufficient evidence observed in that category.

Teacher interview pre-assessment. The needs assessment included a semi-structured interview using the following seven questions:

1. Please talk about something that stood out to you from the lesson?

2. How successful were your students?

3. What informal assessment strategies did you use?

4. How well did your lesson execute your planning?

5. Did you adjust your lesson in the moment? If so, how and why?

6. What will you consider for future lessons?

7. How do you think that will affect the experience for students?

During the teacher interviews, several salient patterns emerged to help explain the pedagogical and curricular decisions made for the lessons observed. When discussing the types of reflections used for decisions, teachers expressed a desire to serve the students. However, most of the teachers' decisions were acknowledged as reducing expectations and standards for the students due to the perceived deficiencies in the students' abilities and prior knowledge. Further, teachers named several external factors such as student behavior, skill, and lack of materials as the reason classroom improvements were difficult.

In conclusion, the science teachers of Bay Vista High School exhibit many characteristics of concern for this design development study, including:

- Reduction in cognitive demand to feel success
- Summative student data used more than daily reflection on student artifacts of learning to determine successes with students
- Externalization of problem source to student behavior or lack of resources
- Limited attempts to connect teaching strategies with NGSS and student learning needs

In order to attempt to address these concerns, the Theory of Action presented in the next chapter outlines factors, grounded in research literature, related to the observed patterns.

Problematic Behaviors

Following myriad observations of science teaching, I have gleaned certain patterns in teacher attitude, behavior, and teacher type. Most salient are the treatment of students, especially the curriculum delivered, and the defensiveness exhibited with regards to professional learning and improvement efforts. Within classrooms, teachers emphasize behavior and control and basic skills, citing these as the types of supports students need before more advanced pedagogical techniques can be attempted. Observationally, I concluded several teachers are capable of advancing their pedagogy by identifying a next level of learning but remain in states of repetition of comfortable pedagogy.

While a spectrum exists from rigid classroom control and teacher-centered instruction to constructivist student-centered teaching, the emphasis on classroom management contributes to lessons featuring low pedagogical skill. Most of the teachers focus on direct instruction and knowledge transmission through teacher-centered methods. Some teachers are pioneering their own constructivist approaches and NGSS-aligned science content. However, my observations yielded little evidence of teachers working to shift their practice and engage more students in deeper science learning. Often these teachers implement strategies sporadically, resulting in incoherent lessons. A final group interprets constructivist pedagogies to a misguided end which allows investigative freedom lacking guidance; students are in a totally open inquiry environment lacking many necessary structures to successfully internalize learning from inquiry activities. Given the gap between high school science teacher experience and behavior and the expectations advanced in the NGSS (NASEM, 2015), teachers could benefit from familiarizing themselves with the pedagogical expectations and begin to practice teaching in these ways.

The second major area of concern is classified broadly as teacher fatigue. Teachers use a variety of explanations to justify their resistance to implementing new pedagogical approaches. Most commonly, they cite lack of student skill and preparedness. Secondarily, the perceived need for knowledge-transfer pedagogy and demands of post-secondary education provide an additional area of concentration for the teachers. During informal interaction, teachers express a perception that universities are knowledge transfer oriented and teacher-centered, leading to the claim that students must experience that type of education to be better prepared. In this case, teachers are overloaded with post-secondary expectations making it difficult to shift to more student-centered pedagogies. The final component, contributing to teacher fatigue during professional learning, stems from initiative overload. Teachers mention multiple professional development initiatives and a constant stream of different district foci as tiring. In addition, teachers reflect that time spent grading, compiling data, and working on other school projects limits their capacity to create new lessons, meaningfully reflect on teaching practices, and respond to student needs.

The most obvious observable symptom of this issue is in the implementation stance of new teachers. Most teachers seek the curriculum to be implemented and become frustrated when curriculum materials are sparse, forcing them to design their own curriculum and lesson plans. Additionally, teachers remark that their administrators and the central office administrators prioritize their ideas for school reform without respecting the skills, desires, and working conditions of the teachers. The question arises from comments such as, "what does having 32 students in six sections per day do to your ability to professionally plan, reflect, assess, and learn?"

Given a lack of professionalism in teaching, the defensiveness towards new initiatives, and the focus on compliance, teachers maintain their status quo and avoid taking learning risks in their classrooms. Although many of Bay Vista High School's science staff are fairly knowledgeable in the subject matter they teach and in ways of teaching, they do not always demonstrate the pedagogical content knowledge required to synthesize the content of science with teaching strategies that result in coherent lesson delivery for students (Shulman, 1986). In order to attend to student need and determine the best approaches for science teaching, teachers must develop habits of personal growth and reflection. Teachers need ways of acting as professional learners, determining and testing new pedagogical skills, to counter the implementation narrative. This final point must be considered when designing a viable intervention.

Desired State

To rectify the concerns presented in the literature and my observations of science teachers, especially Bay Vista teachers, carefully designed experiences for teachers should focus on creating a dynamic teacher learning stance, wherein teachers are trying new classroom techniques suited to their next level work, their zone of proximal development. For the purposes of this study, the teachers' zone of proximal development represents the learning they are capable of turning into classroom change while moving towards the developed skills and behaviors of the TTM (Appendix A). In this new disposition, teachers would be aware of the lofty goals of CCSS and NGSS while concentrating on their growth towards these goals. Ideally, participating teachers will self-assess their next level work need and create a classroom experience to match based on peer feedback and learning inputs. To deepen their learning in one specific area, teachers explore strategies to increase student learning in one focal area of their choosing. Subsequently, they could observe small incremental improvements to their teaching through their own learning as professionals. Teachers could appraise their progress by monitoring signs of student engagement and learning.

While the proximal outcome of trying new techniques within their zone of proximal development is desired for teachers, distally, a shift in teacher attitudes and behaviors would be observed. For example, teachers would be attempting and reflecting on new pedagogical practices, including an increase in student-centered teaching. Teachers would understand the need for all students, especially those in marginalized urban schools, to experience hands-on, relevant, and cognitively demanding teacher practices. Even without a radical shift towards constructivist pedagogies, teachers would exhibit the most important characteristics of active learning, such as taking risks, reflecting on successes and failures, modifying teaching to work towards improvement, and seeking out opportunities to increase their understanding of research and pedagogy. For some teachers, this may mean creating small group discussions to make initial models of a discrepant event. For others, the next level work may be to simply engage more students in active listening to a teacher-centered lesson by using interesting demonstrations. Given all teachers could work on improving their strategies, an initial learning trajectory towards high quality science pedagogical practices will be established and roughly followed by participating teachers. Additionally, teachers would begin taking risks and shift their pedagogical approaches. On a daily or even hourly basis, teachers would respond to evidence of deeper learning and adapt pedagogy accordingly.

Chapter 2: Theory of Action

The following chapter consults the knowledge base for two primary reasons. First, the problem is explored through prior studies and theoretical writing from the corpus of academic literature. Within the first section of the chapter, root causes of the observed challenges at Bay Vista High School are examined with an additional look at general practices dominating teacher professional learning. Subsequently, the second portion of this chapter, Theory of Change, applies principles from various fields of professional learning to determine the best course of action to address the key challenges.

Exploration of the Problem

The knowledge base on teacher learning was explored to identify observed behaviors similar to those observed within the Bay Vista High science teaching staff. Although teachers share some classroom materials, the teacher learning at Bay Vista follows typical patterns found in other schools. Those behaviors and learning stances are visible in both classroom teaching and teacher collaboration periods. Within the classroom, some new strategies are attempted, but few meet the expectations of the NGSS. Further, teaching focuses on knowledge transmission to prepare for summative checks for understanding. During collaboration time, the focus of conversation revolves around descriptions of the obstacles to better student learning, such as student prior skills and administrative requirements. Overall, the environmental teaching conditions contribute to a lack of motivation and to the stagnation in learning and trying deeper science pedagogical strategies. Further, the professional learning opportunities afforded to science teachers in Bay Vista and similar districts are insufficient to address the demands of cognitively-advanced professional standards, the NGSS, outlined previously. Three main concepts regarding teachers form the theoretical underpinnings of the identified problem: motivation, prevalent forms of professional learning and complex learning demands to prepare for NGSS. Each is explored in depth to connect to the problem described above.

Motivation. In order to determine the most effective ways to motivate teachers to internalize new standards, such as NGSS, and take an open and dynamic stance toward professional learning, I turn to an examination of the knowledge base on issues affecting general motivation and teacher motivation specifically. The research base has many studies of individual motivation. The following section details theoretical frameworks used in educational contexts to explain the state of teacher learning and motivation, specifically the effects of decades of high stakes accountability and the view of teaching as a marginal profession.

Motivational strategies. Over the past few decades, school reform literature maintained a myopic focus on one primary motivational factor, high stakes accountability. High stakes accountability is seen as an extrinsic motivator to increase student achievement. While the specific assumptions of studied strategies vary slightly, the common reasoning posits motivation through expectancy theory, where teacher behaviors are attached to external rewards (Kelley & Protsik, 1997). Myriad studies, with varying result, use some form of expectancy theory for teachers to test for an increase in student achievement (Au, 2007; Finnigan & Gross, 2007; Kelley & Protsik, 1997; Mintrop, 2004 and 2012). Within the same vein of research, Skrla and colleagues (2001) study the assumption that high stakes accountability highlights inequitable education and reduces deficit thinking, and Booher-Jennings (2005) looks at the effects of educational triage, distribution of resources to students close to achieving the next performance rank, in the context of a Texas accountability study. While extrinsic motivation through incentives proves a common theme, the following section highlights areas of research implicating motivation through high stakes accountability as a counterproductive effort towards educational reform and teacher learning.

Effect of high stakes accountability on motivation. Much of the knowledge base contains evidence of undesirable outcomes from high stakes accountability which are important to highlight as counterproductive to district-wide curriculum initiatives grounded in new standards, such as increased teacher burnout, simplification of pedagogy and curriculum, and decreased collegial collaboration, (Au, 2007; Booher-Jennings, 2005; Finnigan & Gross, 2007; and Mintrop, 2012). These findings are significant when considering the complexities of NGSS implementation, such as the need for innovation and a learning stance, presented in the first section of this paper. While high-stakes initiatives present undesirable effects towards change, Mintrop (2012) noted increases in teachers' satisfaction with increasing integrity. Additionally, Stone and his colleagues (2008) found increased graduation rates and teacher professional learning when principles of self-determination theory were applied.

Linda McNeil (1982, 2013) raises salient concerns regarding the lasting effects of the high stakes accountability regime. She previously unveiled how the desire to control classrooms and students leads teachers towards a defensive curriculum, wherein content is simplified or obscured from the students (McNeil, 1982). More recently, McNeil (2013) connects this argument to the lasting effects of the managerial control stemming from high stakes accountability by recognizing that teachers who are well-versed in their content area often fall back to these defensive teaching stances to control their classrooms. She characterizes this teacher stance as de-skilling, whereby teachers are reducing their own ability to provide the best lessons to their students, despite having many of the necessary skills to provide a cognitively demanding curriculum.

Marginalization of the teaching profession. Tangential to the issues of high stakes accountability directly controlling teachers is the greater issue of teaching as a profession. Notably, this is an issue of control, or lack thereof, of the work of teaching. Ingersoll (2003) compares the work of teachers in various schools to conclude the professional view of teaching relies on individual teachers controlling their classrooms and being committed to individual students, thus limiting the time and energy expended towards investigation of deeper pedagogical strategies. The predominant societal view of teachers and the type of structure teachers work within can lead to the repetition of the desire to control a classroom and its students (Van Veen, 2005). Teacher motivation to learn advanced and cognitively demanding pedagogical strategies is sidelined in favor of classrooms that give that maintain control and care for individual students. If teachers fail to exhibit basic classroom control, the system leaves teachers with only the students' behaviors and skills to blame, since the entire system is designed to maintain this control (McNeil, 2013). Research on teachers' surveys of their definitions of effective science instruction supports the aforementioned focus on control of the classroom environment and students rather than more cognitively demanding pedagogy and the more cognitively demanding learning suggested in the NGSS (NASEM, 2015). This study must consider the motivation to control classrooms while enabling teacher experimentation with new pedagogical approaches that may seem uncontrolled at first.

This design development study is concerned with shifting these deeply entrenched behaviors and attitudes exhibited in many teachers, especially those at Bay Vista High School. Intrinsically, teachers in Bay Vista have a great desire to help their students. Although society couldn't ask for a greater sense of commitment from teachers to their students, societal narratives regarding poor teacher effectiveness obscure this quality. As management and society at large describe the effective work of teachers as maintaining control of their classroom and delivering basic curriculum, teachers have lost motivation to continue professionalize their own learning in deeper and meaningful ways, opting for more generalized engagement strategies in professional development workshops. While shifting societal perceptions of the teaching profession or managerial desire for controlled classrooms and curriculum is well beyond the scope of this study, it may be possible to motivate teacher learning of complex pedagogies through other means. This research design attempts to acknowledge these barriers to teacher learning and investigate factors that provide teachers awareness and space to deepen their own professional learning

Teacher professional learning. Within this section, I illuminate how prevalent professional development programs attempt to address the needs of teachers. Subsequently, I turn to a discussion of how powerful professional learning is not happening for many teaches, especially those teaching in low socio-economic schools.

While consensus is emerging in the literature regarding the purpose of professional development and the characteristics of effective professional development workshops and programs, researchers are developing a nascent understanding of the effects of professional development on student achievement, increasing teacher outcomes, and the challenges of implementing effective professional development in practice (Desimone, 2009; Darling-Hammond et. al., 2009). The underlying theoretical frame for teacher professional learning posits effective professional development as the practice suited to change teachers' skills and knowledge leading to proximal changes in their behavior and classroom instruction, thus distally increasing opportunity and achievement for students (Desimone, 2009; Supovitz & Turner, 2000).

However, the experiences of Bay Vista science teachers and many other teachers around the country exhibit breakdowns in this system of professional development for two primary reasons relevant to this research. First, many of the professional development programs are not effective for changing teachers' attitudes and behaviors, since they repeat the prevalent, knowledge transmission strategies. Second, even though a picture of high quality PD is forming, few schools have the resources to actualize such experiences for their teachers.

PD reinforcing undesirable patterns. Overall concerns regarding teacher PD, such as only half of teachers reporting satisfactory PD experiences (Darling-Hammond et. al., 2009), provide impetus for deeper investigation into the quality of science teacher PD. When examining reports of the state of science teacher professional development and teacher PD in general, concerning themes emerge with respect to the quality of PD and the perpetuation of basic knowledge transmission. First, current patterns in professional learning exhibit a focus on subject matter learning but lack depth of understanding (Darling-Hammond et. al., 2009). Additionally, most PD for science teachers is delivered in one-time or a short series (Darling-Hammond et. al., 2009; Banilower et. al. 2013). Banilower and colleagues (2013) further illustrated the lack of depth by reporting that fewer than one-third of science teachers are averaging more than seven hours per year of PD. This problem is complicated by the desired PD focus of teachers in high poverty school, traditional knowledge-transmission and content learning over inquiry-based and student-centered approaches (Suppovitz and Turner, 2000).

The second salient issue with the quality of teacher PD is related to collaboration and deeper active learning for teachers. Darling-Hammond and colleagues (2009) report fewer than 20% of teachers of U.S. teachers focus their collaboration on teaching of their content area. Within science collegial learning, there is a high rate of reported participation, nearly 70%, in professional collaboration with science teachers in similar context. However, only one-third of the science teachers reporting collaboration have used student artifacts of learning (Banilower et. al., 2013).

These two issues are visible within the Bay Vista teaching staff, as they have time to collaborate and spend little of it focusing on the deeper pedagogical content knowledge required for learning. Prolonged opportunities to focus on student science learning and the pedagogical content knowledge helpful for NGSS-alignment are necessary to enable effective teacher learning.

Absence of promising PD. In general, teachers in the U.S. are not receiving enough professional development directly related to the pedagogical content knowledge needed for their discipline (Darling-Hammond et al., 2009; Banilower, 2013). Darling-Hammond and colleagues (2009) further explain that a disconnect exists between the professional development programs in the U.S., leading to a patchwork of strategies (Wilson & Berne, 1999). With the limited opportunity for professional learning, teachers' learning is distributed to multiple strategies without the time to experiment and make meaning of the strategies as a coherent approach to improve teaching.

Another challenge within the teacher professional learning knowledge base is a reliance on volunteer teacher participation and self-reports and surveys to calculate the effectiveness of PD (Borko, 2004; Hattie, 2008). This illustrates significant deficits in the professional learning research that are addressed in this design development study. More research is needed to describe the opportunities provided to all teachers, not just volunteers, using more descriptive metrics of learning. Current research provides some insight into strategies that may work for many teachers.

Complexities of the NGSS. As seen in the previous section, many professional development programs are not creating deeper learning and pedagogical shifts in teachers. Those that are successful require long term, intense learning opportunities that are not feasible in many school systems. To complicate matters of professional learning, the Next Generation Science Standards represent a weighty divergence from the pedagogy and content found in many science classrooms, including those at Bay Vista. The following outlines some salient areas of difference between the old standards and NGSS, exemplifying the need for teacher learning.

The NGSS are written to merge the disciplinary core ideas of science, science and engineering practices, and the crosscutting concepts found in all science disciplines. Reiser (2013) and NASEM (2015) highlight key shifts in learning demand linked to the three areas of the NGSS: science teaching should explore and explain phenomena rather than transmit science facts; inquiry and scientific practices, often taught as discrete units, must be taught and used as the basis for all lessons; storylines must be built across time for students to investigate science concepts. For the teachers at Bay Vista and elsewhere across the country, these three divergences completely shift the classroom expectations and teaching patterns of the past. A learning design for the teachers should include aspects of these three key shifts represented by the NGSS. Additionally, the NASEM (2015) report outlines significant needs in the skills required to teach science to a widely diverse group of students, and pedagogical content knowledge required for teaching rigorous NGSS-aligned science. Considering the mandate to make rigorous science, aligned to the key shifts outlined above, teachers have incredible learning needs at this time.

National Survey of Science and Mathematics Education, NASEM (2015) highlights a national problem I observed in Bay Vista and other schools. The prevalent form of instruction in high schools remains didactic, text book oriented and relatively void of science practices such as using evidence to support claims (NASEM, 2015). While NASEM (2015) outlines promising professional learning practices, to be discussed in the subsequent section, the group cautions that these are based on an extremely limited knowledge base and go well beyond what is typically observed in schools. What is clear through the research and my personal observations, the learning needs of science teachers, especially during this transition to NGSS, are deep and require concerted efforts by school systems to invest in the learning of science teachers.

The question remains, how can a professional learning experience shift teachers from the repetition of normed behaviors and styles towards a dynamic and open stance toward deeper learning of science content and pedagogy outlined in the NGSS for all students?

Theory of Change

Within this section, I provide a framework for addressing the challenges outlined previously through change drivers rooted in the research of academics from various spheres and the intuition of practicing educational leaders embedded in the urban context of Bay Vista High School.

Developing the motivation and skills needed to work towards gradual improvement in teaching practice is a difficult pursuit. As previously explored, teachers face internal and external obstacles discouraging them from opening to new learning. However, the teacher professional learning knowledge base provides insight into potentially effective promising learning practices. To begin addressing the discrepancies between traditional teaching and cognitively-demanding pedagogies, teachers must be inspired and motivated to learn about and attempt new practices aligned to NGSS. Through professional learning programs, teachers must also find inspiration and increasing efficacy by focusing on the positive aspects of student learning. Finally, increasing their sense of self-efficacy and sustaining motivation are key to creating lasting shifts in behavior and attitude. Through classroom experimentation and sharing of artifacts in collegial learning environments, teachers may begin these important learning shifts.

The subsequent sections will detail a research-based approach to rectifying the aforementioned concerns regarding teachers' persistent attitudes and behaviors. First, I examine the literature on effective professional learning before turning to theories indicating potential to motivate, increase self-efficacy, and work towards incremental improvement in classrooms. I will subsequently provide a proposed intervention while analyzing the feasibility of the proposed project.

Insights from professional learning literature. In the following section, I explore the knowledge base on professional learning. By starting with consensus definitions of high quality professional learning for all teachers and moving to current extensions for science teacher learning, I highlight concepts important to the development of a successful intervention at Bay Vista.

Desimone (2009) builds a persuasive case for a consensus on the characteristics of high quality professional development; focus on content, active learning, coherence, duration, and collective participation. Focusing on subject matter content and how students should engage with that content may be one of the most influential factors of effective professional development (Desimone et al. 2002, Desimone 2009, Penuel et al. 2007, Supovitz & Turner, 2000). Active learning encourages teachers to engage authentically in their own learning through interactive activities and contributes to teacher learning and change in practice (Penuel et al., 2007, van Driel et al., 2012). Additionally, aligning professional development coherently to the goals teachers have for themselves, their students, and state and district initiatives increases the sense that the experience will be valuable (Desimone 2009, Penuel et al. 2007). This theory of change suggests building a stronger sense of teacher professionalism through autonomy in decisions regarding their professional learning and respect for each other as professionals. Teachers' ability to participate meaningfully with a group of other teachers enhanced the effects of the professional learning experience, potentially leading to school-wide changes (Darling-Hammond et al., 2009; Penuel et al. 2007, Wilson & Berne, 1999). This study needs to account for the strength of collaboration while determining ways to foster a sense of autonomy.

The NASEM (2015) report synthesizes current research to contribute two salient ideas to the previously mentioned notions of professional learning by addressing science teachers specifically. Science teachers benefit from analyzing content and pedagogy, specifically artifacts of student learning, using tools developed to facilitate analysis and collaboration (NASEM, 2015). The report adds the benefits of allowing teachers to focus their professional learning on problems they are experiencing (NASEM, 2015). These ideas align with notions of increasing professionalism and motivating teachers by providing deep learning opportunities focused on their personal learning needs.

Additional insight is gained by analyzing the research on school context and collaborative group work. School leaders and professional development designers should focus on the creation of sustained and focused professional development that aims to engage teachers as intellectuals pursuing knowledge and skill in a focused area (Wilson & Berne, 1999). This aspect aims to increase the professional attitudes of teachers by providing them autonomy within the concentration of science teaching. Finally, collaborative structures encouraging collective participation alone will not generate the desired learning (Hattie, 2007). Additional structures, motivational levers, and foci are discussed in the following sections. The intervention studied here will consider these notions of high quality PD when constructing activities for a group of teachers in a regular school collaboration setting, not volunteers for additional learning.

Motivating and sustaining teacher learning. A powerful asset, among the science teachers studied for this intervention, is their desire to help students. Many of them express deep care for their students. The late hours, open door policies, and caring descriptions of students exemplify Bay Vista science teachers' commitment to their students. However, predominant teaching styles and foci of teacher learning indicate a need for pushing more cognitively demanding science pedagogy. In this case, teachers need to understand that attempting more student-centered, contextualized, lessons with a focus on practicing science will be of great benefit to students. The challenge is to motivate teachers to accurately identify and attempt learning in their zone of proximal development, as they work towards deeper instructional practices through experimentation and reflection. Additionally, examples of teacher learning trajectories will illuminate potential areas for growth in a reduced stress and blame environment. During reflections on their chosen next level work, teachers may increase their desire to continue to learn by identifying small successes in their classrooms. Subsequently, incremental changes will be applied to classroom practice and accumulate into deeper pedagogical shifts along a learning trajectory. I then apply goal setting theory to develop motivational activities, based on critical reflection, as an approach to engage teachers in their own learning. These conditions may also serve to increase teachers' self-efficacy (Luthans et. al., 2015).

Notably, self-determination theory, presented by Stone and his colleagues (2009), assumes humans will be motivated by their basic needs: competence in their ability to complete tasks, relatedness to their colleagues, and autonomy to guide portions of their own work. Alongside self-determination theory, Amabile and Kramer (2011) assume that individuals' happiness positively correlates to effective work outcomes. Additionally, Mintrop (2012) contributes a fourth major assumption, self-concept theory based on integrity leads to more coherent approaches to the high stakes accountability environment. More importantly than the assumptions made in the research are the observations which lead to eventual frameworks of motivational theory.

Setting goals. In a revision of their pioneering earlier work on goal setting theory, Locke and Latham (2002) explore various types of goals and conditions needed for success. Previous studies focused solely on performance goals, those with an explicit end result. In their 2002 article, Locke and Latham suggest that learning goals may be most appropriate for increasing motivation and self-efficacy. Luthans and colleagues (2015) support this notion with the idea that reachable goals should be accompanied by self-determination to persevere and learn approaches to complete the goals. NGSS practices are often described by teachers as out of reach of their own classroom abilities. For these teachers, identifying one area of improvement would be useful; goals and attempted pedagogical shifts must lie in an area of attainability while maintaining a cognitive push, the zone of proximal development (Vygotsky, 1980). Instead of seeing their goal as students creating high quality, for example, teachers should set goals for their own learning. For example, a teacher may investigate the student creation of high quality models through various learning sequences or peer review experiences. In this example, the goal is based on teachers' learning the best methods for teaching the practice of modeling. Seijts and Latham (2001) found greater commitment to cognitively demanding goals if a focus was set on learning rather than outcome. Learning goals necessitate teacher problem solving leading to changes in classroom practice. In the context of Bay Vista High, goals should be set by the teacher for their own learning.

Increasing self-efficacy. Given the reliance on self-efficacy in many discussions of motivational improvement and the externalization of problems facing educational improvement, it is important to define self-efficacy for teachers and explore the differences between teachers with and without high levels of self-efficacy. Self-efficacy, in teachers, is commonly held as the belief that one can have a positive effect on student learning (Bandura, 1977; Ashton 1985). Tschannen-Moran and Hoy (1998) specify efficacy as teacher's belief with regard to specific tasks and contexts. To appreciate the importance of self-efficacy, as it pertains to the transition to NGSS, I turn to the research on teacher self-efficacy and the relationship to student outcomes, and teacher behaviors.

A primary goal of increasing teachers' desire to learn is to positively shift their self-efficacy. Among other benefits, this has been shown to increase creativity and problem-solving abilities, crucial attributes for a transition to NGSS (Tierney and Farmer, 2011). Caprara and colleagues (2006) found teachers with higher self-efficacy are better able to handle complex tasks and are more adaptable to student needs and contexts. Luthans and colleagues (2015) identify positive feedback, encouragement, observation of masters, as means to positively boost self-efficacy. Locke and Latham (2002) demonstrate training, role modeling, and expressions of confidence as means to boost self-efficacy. Since self-efficacy is positively correlated to work performance, working to improve self-efficacy is a worthy endeavor to help the studied teachers work to improve their instructional effectiveness through personal professional learning (Luthans et al., 2015).

According to the body of literature, contextual and collaborative factors contribute to the self-efficacy beliefs of teachers. Several themes emerge in this literature. First, the context teachers work within seems to affect the levels of self-efficacy. Caprara and colleagues (2006) also found collaboration with colleagues to lead to increased self-efficacy. Further, increased stress from student behavior and decreased comfort levels are related to lower levels of self-efficacy. By contrast, highly professionalized workplaces respect the autonomy and advanced training of individuals in the system and create environments and structures to support that professional work (Ingersoll, 2003; Van Veen, 2005). Logically, addressing issues of professionalism and feelings of autonomy, while boosting collegial learning, should improve feelings of efficacy for teachers.

Experiencing successes. The need to maintain the motivation to learn and improve by experiencing multiple small success is well documented (Amabile & Kramer, 2011; Locke & Latham, 2002; Luthans et al., 2015). Most importantly, Amabile and Kramer (2011) illustrate the positive correlation between positive emotions, creativity and productivity. They further their case by connecting incremental progress to positive work life conditions. For a teacher who has stagnated in learning, the progress associated with finding small successes could lead to a feedback system causing greater happiness and creativity at work. Creativity is synonymous with learning and experimenting in the classroom. Therefore, recognizing small gains towards goals is vital to the success of teachers working towards learning goals.

Once self-efficacy begins to increase, learning for the sake of instructional improvement must be sustained. To address the continued improvement of teachers' learning for growth in practices contributing to NGSS education, I draw on two theories highlighting the importance of happiness, progress, excellence, engagement and ethical views to form a scaffold of sustenance for teacher learning. Amabile and Kramer's (2011) progress principle suggests adequate and incremental improvement through continued attempts and celebrations of success, outlined above, can create feelings of happiness which contribute to a positive inner work life. This notion suggests the creation of a positive feedback loop, where small achievements generate feelings of happiness, in turn boosting self-efficacy and increasing work quality, which will inevitably lead to further experiences of success.

Gardner and his colleagues (2002) posit the theory of good work creates satisfaction for workers by focusing on the work of purposefully changing and challenging the status quo, committing to excellence, and ethical alignment. Teaching certainly falls into these categories. Those striving to do "good work" in education will find greater happiness, thereby improving their inner work life and further contributing to the desire to improve their practice. The aforementioned strategies will be folded into the subsequent plan to improve the work of science teachers, inspiring them to strive for continuous improvement through careful study of their practice and science pedagogy.

Collegial learning. Professional learning communities (PLCs) have been at the forefront of research on teacher learning for nearly 20 years (Council, 2016; Darling-Hammond and Richardson, 2009; Kruse & Louis, 1993; Vescio et. al., 2008). Purported benefits to teacher learning and student achievement are well documented (Council, 2016; Darling-Hammond and Richardson, 2009; Kruse & Louis, 1993; Vescio et. al., 2008). However, some dissonance appears in the literature between the types of PLCs and conditions necessary for success (Darling-Hammond and Richardson, 2009; Kruse & Louis, 1993). The corpus of PLC research has even grown to include specific activities to be conducted in science teacher PLCs (Lewis et. al., 2014). For the purposes of this study, teacher collaboration is proposed as a means to develop trusting and psychologically safe learning environments wherein teachers explore advanced and cognitively demanding science pedagogy.

Darling-Hammond and Richardson (2009) highlight teachers experiencing trusting environments capable of fostering deeper reflection and growth, potentially leading to whole department or school-wide change. Additionally, they cite Little (2003) and Dunne, Nave & Lewis (2000) to emphasize that change occurs as these groups develop discourse to describe science teaching practices aligned to changes tried in classrooms. Additionally, Vescio and colleagues (2008) describe the need for PLCs to allow teachers to bring their own learning needs to the table. This design focuses on analyzing artifacts of student learning and teaching within a structured small group discussion, while allowing teachers to choose their own focus of learning.

Summary of literature applied to theory of action. Following the literature review from diverse fields, this intervention design was created to leverage principles of motivation and promising professional learning practices to attempt to shift teachers to deeper learning of science pedagogical strategies. By controlling their own plans, analyzing teaching and learning artifacts, observing student successes and discussing all aspects with their small department, the teachers may shift from informal and repetitive routines to a more in-depth analysis of teaching and learning in science classrooms.

Underlying these activities used to shift teachers' learning stance are the concepts of high quality professional learning, goal setting, increasing self-efficacy and experiencing successes. This design development incorporated notions from the literature on professional learning to oppose the de-professionalization most teachers experience to some degree. Further, having the autonomy to choose their own learning goals contributes to professional attitudes. The challenge of deepening learning strategies is addressed by the focus on narrow science teaching strategies with opportunities for collegial sharing and reflection.

Feasibility. The implementation of this improvement framework must match successful professional learning strategies, enact change drivers, and be a practically applicable to the context at Bay Vista, in order to become a feasible model for improvement. Professional learning must be inextricably linked to specific contexts, sustained and tied to classroom practice (Darling-Hammond, 2009). As such, this proposal contextualizes teacher learning and improvement in work towards deepening science teaching with regard to student learning and current issues affecting science pedagogical content knowledge.

The project is enabled by the deep ethically rooted caring stance teachers have towards their students. This attitude will facilitate motivation towards good work and progress. Additionally, some teachers are motivated to improve their practice. These teachers actively engage in professional learning workshops and frequently serve as formal and informal teacher leaders. They represent a pool of experience to be used during observations, exemplar videos and positive peer support sessions. As previously stated, exposure to coaching and positive role models increases self-efficacy, thereby generating the conditions for teacher learning and improvement.

With the pre-established theoretical underpinnings grounding the study, the logistics of implementing an intervention at Bay Vista High School must be considered. Fortunately for this design, the science department has a 90-minute collaboration meeting scheduled every week. The intervention was designed for seven workshops to be held during these meetings. However, the department needs time to discuss department issues. Additionally, meetings times are occasionally filled with schoolwide meetings. Therefore, the seven-workshop series occurred over the course of 12 weeks.

The design requires facilitation and consideration of individual teachers' needs. I served as the facilitator and principal designer. It was fortunate that the current department chair was open and reflective in her leadership. She served as a critical friend and co-designer, which was particularly impactful due to her knowledge of participants' experience, teaching styles, and personalities.

Finally, the department had some experience with cycle of inquiry and sharing practice from previous years. Although the practice did not reach the depth of this study, the design may have benefited from of the routines that were attempted the previous year.

The reliance of the design on teacher participation in collegial environments and focus on artifact-based learning plans presented challenges to feasibility. Specifically, teachers at Bay Vista have heavy preparation loads and limited time outside of the classroom. Teacher attendance and completion of portions of the intervention happening in their own classrooms affected participants' completion of all activities.

Intervention Design

A carefully designed intervention, in a design development study, must align components of the intervention with anticipated learning and the theoretical groundings of the inquiry (Mintrop, 2016). The following intervention design (Table 2) was planned to take approximately three months with main activities occurring two or three times per month during 90-minute workshops. Generally, the beginning workshops covered norms of interaction, goal development, and skill building in order to create a safe and professional group dynamic, identify next level work, and create an understanding of the learning process needed to meet high level NGSS-aligned instructional expectations. Additionally, teachers used goal setting to establish their specific learning and growth areas as professionals. The second half of the intervention centered around self and peer reflection of artifacts and classroom experiences. Table 2 summarizes the seven main workshops of the design.

The activities in this intervention were designed for individual learning in small groups. During the first group activity, each participant reviewed vignettes of five example teachers (Appendix B). These composite sketches of classroom teaching were designed to represent five commonly seen teaching practices, based on the TTM (Appendix A). Participants read the vignettes and identified commonalities between the vignettes and their own practices. Probing questions were used to encourage selecting a combination of attributes. This activity was designed to help participants establish a baseline for the next level learning, learn to self-analyze, and begin to see the learning trajectory towards NGSS learning.

During the first group activity, participants established norms to focus on improvement, professionalism, and technical feedback by practicing lower inference observation with reduced judgmental language. Additionally, participants practiced technical feedback, feedback based on low-inference observations of teaching and learning, while discussing examples of NGSS-aligned activities. To contribute to notions of professionalism, this activity began to provide a groundwork for low-inference, technical feedback, a key component of providing professional feedback to colleagues and self-describing teaching practices. By acknowledging that all forms of teaching are acceptable and can be critiqued through non-threatening language, participants were to build the skills and comfort to participate in later feedback activities. This activity was also intended to reduce the defensiveness often observed in mandatory professional development workshops.

Table 2: Summary of design implementation activities and anticipated outcomes

Workshop	Activities	Anticipated Outcomes/Objectives
1 – Teaching Styles and NGSS Activities	Phenomena-based science investigation Teaching style vignettes	Understand phenomena-based learning as an NGSS-aligned, engaging strategy. Generate ideas for phenomena. Identify characteristics of your teaching style.
2 – Examples of Learning Trajectory	Review learning plans associated with vignettes. Write a learning and classroom experimentation plan.	Identify areas of work and strategies to try that will enable learning to improve teaching.
3 – Establish Next Level Learning Goals	Fishbowl with department leader Incremental improvement tips Peer feedback on improvement plans	Determine a next level learning goal aligned to your teaching. Understand how to use classroom inquiry for incremental improvement. Finish a learning plan with a strategy and data source.
4 and 5 – Evidence Sharing and Feedback	Presentations by colleagues	Learn from classroom experiments of colleagues. Receive feedback to inform next steps.
6 – Building on Learning Goals	Evidence-Strategy connection activity. Design new learning plans.	Understand the link between strategies and evidence. Determine sources of information to best inform teaching practice. Improve learning plans for another round of classroom experimentation.
7 – Evidence Sharing and Feedback	Evidence choice group discussion Peer Review of plan implementation.	Improvement seen between first round of goals and evidence presentation to the second round.

During the second and sixth workshops, participants were to construct personal learning plans (Appendix B). The key learning inputs were to be example learning trajectories outlined in the extended vignettes (Appendix B). Additionally, participants were to glean ideas from each other's professional learning plans. This activity would connect to the change drivers by motivating teachers to work at their personal next level and develop meaningfulness in the work of teaching.

The final type of learning activity, workshops three, four, five, and seven attempts to motivate individuals by having them focus on the changes experimented with in their classrooms, as evident in the artifacts shared for peer feedback. A peer feedback structure was introduced to reduce feelings of judgment by focusing on low inference technical feedback. Volunteers who would share in front of colleagues could opt to present their inquiries first, during workshop four. The remaining participants would be asked to share during workshop five. Video and audio recording would be encouraged due to proven effectiveness and observers' ease of making low inference comments (Sherin & Han, 2004). Student work samples also encouraged a focus on the student learning. During the final workshop, all participants would share in small groups. Participant self-reflection would play an instrumental role during these activities. Not only would participants think about the quality of the teaching and outcomes, but they would also reflect on the teaching moves from one class period to the next. During the workshop activities, careful collection of data by the researcher would make connections between the learning and the change drivers. The following section details data collection procedures and connections.

Chapter Three: Research Design and Methods

This design development study investigated an intervention working with secondary science teachers who have been observed implementing consistent teaching strategies without directed attempts to improve practices. Having created strategies to manage classroom routines, these teachers tend to exhibit the following attitudes and beliefs that the proposed intervention attempted to change:

- low self-efficacy with regard to student learning of complex science processes
- traditional knowledge-based curriculum relying on memorization of science facts over engagement in scientific practices
- different levels of pedagogical content knowledge
- diminished teacher motivation due to de-professionalization
- repetition of status quo teaching techniques with a focus on classroom control and curriculum completion
- resistance and fatigue when facing standards-based reform initiatives

By addressing the aforementioned attitudes and beliefs, the goal of the outlined design intervention was to create a willingness to experiment with new strategies and become motivated to research, through literature and personal experience, the types of pedagogy implicit in the NGSS. This shift in attitude and behavior represented the desired proximal outcome while distal outcomes included increased collaboration and greater student outcome through teaching improvements.

The issues of teacher motivation, self-efficacy, and defensiveness to instructional change vary widely from one teacher to another. Highly variable contexts, a practitioner approach to research intervention based on a theoretical base, an inability to control the natural research environment, and specific details of each teacher's situation and experiences, necessitated a research methodology with a systematized implementation plan and data collection procedure allowing for flexibility to adjust and rework the intervention. Design development studies (DDS) are well suited for implementation flexibility in complex educational contexts where the researcher proposes an intervention to shift behaviors, attitudes and beliefs to a desired state (Van den Akker et al., 2008; Mintrop, 2016).

Whereas experimental designs attempt to control the research environment and contrive laboratory-like conditions, DDS embrace the complexities found in naturalistic settings, thereby allowing a more complete understanding of the research situation. Similarly, DDS do not purport to generate generalizability as is the case for experimental and quasi-experimental methodologies. Further, experimental designs compare control and treatment groups. Design development studies contain treatments or interventions without the need for controlled comparison groups, opting instead for clear descriptions of the implementation process and making a plausible connection between process and outcome of the intervention.

DDS methodology provides a structure in which interventions can be tracked and data can be systematically collected, thus documenting progress towards the desired state (Van den Akker et. al., 2008). By systematically collecting data on the process and outcome, the causal links between the intervention and the outcome may be established, and the knowledge may be applied to similar contexts.

The intervention, in this study, sought a solution to increase Bay Vista High School science teacher motivation and self-efficacy through learning goal setting, identification of next level learning, reducing negative pressures, building skills and reflection on small successes. The literature lacks a synthesis of these foci to positively shift teacher motivation and self-efficacy. Therefore, a DDS was employed to allow for tinkering with the combination of factors contributing to the desired state, teacher feeling more self-efficacious and motivated to work towards the implementation of pedagogical strategies more aligned with NGSS. Given the strength of the knowledge base regarding teacher motivation, skill building and self-efficacy, a design development study, in this case, allowed for intervention rather than mere exploration found in other forms of action research. However, action research provided an influential methodological base for this design. Miles and Huberman (1994) suggest rigorously tracking and reflecting on the story of the change process. Although more structured than action research, much of this design followed the embedded researcher reflecting on an intervention that arises from action research. Both methodologies are well-suited for natural environments with identified problems.

In conclusion, DDS allowed me to systematically track the intervention process and document the proximity of the observed state to the desired state. Design development was ideal for this research problem due to the variety of contexts teachers work in and the relative stability of the challenge of implementing NGSS. DDS allows for an intervention to be attempted in multiple, slightly varied contexts. In additional iterations, the intervention can be tweaked and implemented in response to the research team's process data (Van den Akker et. al., 2008).

Elements of the Research

A defining characteristic of research methodologies is the type and purpose of data. Most methodologies contain data describing the research environment and intentional summative measures designed to determine effect. Design development research is unique in its use of two main types of data. Given the thoughtful intervention involved in design development studies, impact data are one focus of design development research. Unlike the outcome data of experimental research, design development studies compare outcome data to the baseline data before the intervention, rather than comparing the outcome data derived from the treatment group to the data derived from a control group.

The second type of data crucial to design development data is process data. Although pure action research also relies on process data, it does so more heavily than design development studies, since action research is often ongoing, and the research/practitioner is documenting a change process. In contrast, design development studies have an intended, research-based theory of change and intervention plan that need to be documented with process data. Therefore, both process and impact can be planned and standardized, to varying degrees, in advance of the intervention.

Impact data. Measuring the impact data for design development studies must be aligned to the treatment group and desired changes. It can often be compared to baseline data using similar measurements. Impact data should be highly standardized, before the research project begins. Additionally, multiple metrics should be employed to triangulate the data and sufficiently measure impact in each area of desired change. The multiple metrics provide a more robust connection between the intervention and potential outcomes.

For this study, several forms of impact data were collected before and after the implementation of the intervention. First, participants were asked a series of structured interview questions designed to illuminate the effectiveness of change drivers with regard to the underlying problems of the study. The questions (Appendix C) investigated the ability to self-analyze, the ability to set next level learning, sense of confidence and motivation, sense of learning quality and sense of learning safety. Each domain contained a few Likert scale response questions with probing questions to elicit the underlying meaning to participants' responses, check the accuracy of the Likert responses, and assist in the identification of patterns during the analysis. The questions regarding self-analysis aimed to determine changes in the participants' assessment with regard to strengths and areas of growth in their teaching. Setting the next level learning determined the participants' recognition of the importance of properly set learning goals and their sense of the goals being appropriate to their growth needs. The study also attempted to build self-confidence and motivation, as evaluated with questions pertaining to noticeable improvement in one's own teaching practice and student indicators. Learning quality indicators assessed the progress participants feel they have made, while learning safety was an indicator of how well the design created an environment where all participants could take learning risks, primarily by sharing work from their classrooms.

The second form of impact data was a performance-based assessment designed to evaluate several of the same domains as the structured interview questions, namely evaluation of teaching, setting next level learning, and capacity and skill building. The key difference, however, was in the subject of the domains. In the interview, participants reflected on their own changes. The performance assessment investigated teachers' ability to reflect on teaching practice and work towards improvement by focusing on another teacher. Videos of unfamiliar teachers were shown. Participants reflected on the observed teaching style, identify areas of growth and suggest a learning plan for the teacher. Simply seeing change in the interviews, performance data, and surveys is insufficient for DDS without linking the impact data to the process. Table 3.1 outlines the design levers being implemented and the corresponding sources of impact data that tracked the study outcomes in each area.

Table 3.1: Design levers and corresponding metrics of impact data.

DESIGN LEVERS	IMPACT DATA SOURCES
Ability to evaluate teaching	Semi-structured Interview Questions, Performance Assessment
Ability to set next level learning	Semi-structured Interview Questions, Performance Assessment
Sense of confidence, motivation, and efficacy	Semi-structured Interview Questions
Sense of Learning Quality	Semi-structured Interview Questions
Sense of Learning Safety	Semi-structured Interview Questions

Process data. Unique to design development and action research is the use of process data. This data stream allows for the tracking of intervention progress. The purpose is to capture key events that eventually influence the outcome, as the design unfolds. In both action research and design development studies, the researcher, who is also working in the context, attempts to shift outcomes to a desired state. In order to document the key moments creating a change, as measured by the impact data, process data records the implementation process (Mintrop, 2016). While the standardization level is lower than in impact data, process data is useful for several salient reasons. Table 3.2 highlights the types of data that tracked the progress made with regard to specific activities in the intervention.

Table 3.2: Summary of design activities and process data.

Workshop	Anticipated Outcomes/Objectives	Sources of Process Data
1	Understand phenomena-based learning as an NGSS-aligned, engaging strategy. Generate ideas for phenomena. Identify characteristics of your teaching style.	Low inference field notes and audio recordings. Survey including teaching style reflection and phenomena-based instruction.
2	Identify areas of work and strategies to try that will enable learning to improve teaching.	Low inference field notes and audio recordings. Survey to determine influences on individual learning plans
3	Determine a next level learning goal aligned to your teaching. Understand how to use classroom inquiry for incremental improvement. Finish a learning plan with a strategy and data source.	Discussion questions: What did you consider when choosing a learning plan goals and strategy to try? Low inference field notes and audio recordings. Learning plans.
4 and 5	Learn from classroom experiments of colleagues. Receive feedback to inform next steps.	Low inference field notes and audio recordings.
6	Understand the link between strategies and evidence. Determine sources of information to best inform teaching practice. Improve learning plans for another round of classroom experimentation.	Responses to Evidence-Strategy activity. Low inference field notes and audio recordings.
7	Improvement seen between first round of goals and evidence presentation to the second round.	Shared evidence statements. Low inference field notes and audio recordings.

In design development studies and action research, process data is beneficial to the researcher as it provides evidence of the learning process to determine key features of the implementation as it relates to the outcome. Table 3.2 exhibits the major process data collection points as related to the stages of the intervention implementation. Process data must remain fluid and malleable to enable tracking of unexpected and important events during the implementation. For example, this design relied on teacher reflection and conversation during the intervention. If a key experience happened, audio documentation was available. However, rubrics and coding schemes were not predetermined to allow themes to emerge from participants' conversations. Additionally, process data generated many unique data points within an inquisitive stance, thereby reducing biases commonly found in impact data.

To be most effective, Miles and Huberman (1994) suggest limiting process data to areas within the conceptual framework and research question. While process data should be fluid, design development process data should be planned in advance in several key ways, such as chunking into key activities of the intervention. Once chunks have been established, salient behaviors and corresponding indicators should be predicted so the process data focuses on key portions, explicit or invisible, can be categorized and the data streamlined into manageable portions. For example, it was anticipated that personal learning plans will provide insights into participant thinking and development of self-reflection and setting next level work. Therefore, this data point is repeated in the study. By carefully planning the key foci and instruments or process data, the researcher can align the process data to the key outcomes represented in the impact data (Mintrop, 2016).

This design had three several sources of process data. First, teacher learning plans were collected during the workshops. Since two separate plans were being created, some ability to track changes in the types of goals for the plan are key. A second source was audio recording and observations of the teacher discussions to identify patterns in their thinking during the intervention. Following the events, a detailed field notes with low inference observations (Miles & Huberman, 1994) were compiled by the researcher to track key micro events, potentially influential moments within a workshop. The third source entailed the reflection and feedback given to teachers during Workshops Four, Six, and Seven.

Additionally, teachers shared various student-based artifacts from lessons with newly attempted strategies. These artifacts were analyzed over the course of the intervention to identify any possible changes. Teachers also engaged in a self-reflection process to determine their own improvement. Analysis of the self-reflections yielded valuable information as to how teachers are self-evaluating and focusing on the improvements and needs in their classrooms. A third source gauged participants' learning through direct post-intervention questionnaires and discussions, sometimes occurring as warm ups in subsequent sessions. The questions were intended to ascertain the experience of the participants with regard to the specific intended outcomes of each intervention step. In combination, these process data sources served to describe the learning process and how it is linked to observations in the impact data.

Participant Sample Selection

According to Maxwell's (2012) criteria for qualitative research sampling are representativeness, variability, theoretical testing, comparison, and relationships. Additionally, he suggests convenience of implementation and data collection to be factors affecting the sampling used. In DDS, participants are selected due to the fact that they are centrally involved in producing the problem of practice and in potentially benefiting from new ways of thinking and acting. Therefore, purposeful convenience sampling is employed in this case of this study.

As opposed to probability sampling, which selects participants randomly, purposeful sampling selects participants from known groups. This research design employed purposeful sampling by choosing participants intentionally due to the insights potentially gleaned from their participation; the intervention was expected to affect the participants selected for the study. Participants were also selected due to their membership in the institution the researcher has access to and wants to change. In this study, a school site science team was intentionally selected based on a set of criteria, primarily that several members of the team display a closed stance towards learning and improvement. Although this could be viewed as convenience sampling, since principals will need to permit the research to occur, schools with all novice teachers or high functioning professional learning teams were not considered for the study.

Unit of Analysis

The unit of analysis defines the parameters of observation for a research study. By carefully selecting the subjects, groups, or systems under observation, the scope of potential data becomes better defined. In several types of research methodology, design development being one, a unit of treatment is also defined. A unit of "treatment" does not need to be specified for methodologies without the intent to intervene, such as case studies, ethnographies, and other exploratory methods. However, design development studies have intentional treatments, interventions, but without a comparison group.

In this DDS, individual science teachers belonging to the same school site were the unit of analysis. The study used pseudonyms and refrained from including identifying information to preserve as much anonymity as possible for the participants. All seven of the participants have been teaching for at least four years. None of them expressed an intention to retire in the next five years, making them ideal mid-career candidates for the study. Two teachers participated in a couple workshops but opted out of the study group.

As Borko (2004) points out, most professional development literature relies on motivated volunteers. This study purposefully grouped teachers typically volunteering to participate in learning with those who rarely seek out professional learning opportunities. However, the study itself occurred during mandated collaboration time. The teachers were expected to participate in some type of professional growth during the time. While the research collected data on some participants exhibiting high functioning learning dispositions, the focus of analysis will be on those teachers who were identified in the needs assessment as stagnated in learning or having significant areas of growth towards advanced science pedagogical practices.

Level of Design Flexibility

Since this design investigated a common problem with limited documentation in qualitative research, a more structured research design was constructed. Overall, the original plan was highly structured to investigate the effects of the aforementioned domains. However, there was flexibility to respond to individual participants' needs and the needs of the group as a whole. For example, an activity on choosing best types of evidence and a fishbowl discussion with a colleague were added to address concerns of the department chair. Additionally, workshops were monitored through facilitator reflection to identify areas of adjustment.

Data Collection Methods

A necessary component of research designs is a methods section detailing the use of particular types of data. Research designs looking for a high degree of generalizability tend to generate standardized large numeric data sets, primarily quantitative data. In this design development study, the complex behaviors and changing stances and attitudes of teachers were tracked. This required a set of qualitative data capable of providing richer insight into the implementation process and the impact of the design. Methods for collecting qualitative data were structured and determined in advance of the study.

In this design development study, several major types of qualitative data were collected. Highly structured interviews were used at the beginning and end of the intervention as a metric of focal teacher disposition to learning and experimenting with science practice. Interviews allow for a deeper look into the thinking of participants (Patton, 2005). Performance assessment data was compared to determine if teachers altered their ability to evaluate and suggest learning plans for an unfamiliar teacher. Contact summary forms, a version of field notes, were completed following each intervention activity. To increase the thoroughness, audio recordings were taken of the intervention activities. Participants and researchers analyzed video and other classroom artifacts. Multiple sources of data and careful record of impact data were needed to ensure the rigor of the design.

Data Analysis

In this section the four main data sources, by type, are discussed. Each type of data used in this DDS requires different analytic techniques. Although other forms of data will be collected, presented here are the forms necessitating the most thoughtful analysis and with the most important projected connections between process and impact.

Structured interview questions (Appendix C) consisted of two types of questions requiring two different data analysis formats. First, participants provided a Likert scaled response to scripted questions. These were compared directly between baseline and outcome interviews. Second, a set of probing questions was coded and analyzed using a rubric (Appendix D). The second piece of impact data, the performance assessment, was also reviewed and coded using the same rubric. Additionally, process data was analyzed.

Twice during the intervention, participants completed personal learning plans (Appendix C). The plans were analyzed for signs of appropriate next level work. Although initially compared to the teacher typologies and sample learning plans, possible next level learning developed out of the teacher thinking. By analyzing teachers' plans, new themes and possibilities emerged.

Self-reflections and peer feedback data were noted by the facilitators as well as recorded. When noting an interesting matter, the audio recording was consulted to determine the significance of the interaction. Of particular interest were interactions that led to growth moments, moments where a safe yet critical discussion exists, and moments that shut down participation. Additional coding of the process data was considered, if the investigation necessitated it.

Rigorous Research Design

To contribute to the academic knowledge base, all studies must surpass basic requirements specific to the type of research. Although positivist scientific research has strict definitions of rigor, concepts of reliability, validity, and transferability can be useful when creating a DDS. As an action researcher implementing a predetermined design, specific concerns regarding threats to rigor and bias were considered.

Reliability. Reliable research designs all have one salient feature in common, the methods are enumerated clearly enough to provide other researchers the opportunity to reproduce the design. In experimental designs, reliability is ensured with careful planning of treatment and control groups, listing the conditions of the experiment, and providing a detailed description of the treatment and data collection methods. In DDS, reliability is ensured by using metrics and procedures that can be clearly followed and understood by other researchers (Mintrop, 2016). The nature of malleability in the design implementation is a cause for concern, in terms of reliability. As moment-by-moment decisions are made, the design could become unclear. However, careful planning of observation tools and recording changes to the implementation can assuage reliability concerns. As see in the appendices, impact and process data tools were carefully crafted to capture key moments in the implementation of the design. Additionally, when the carefully planned design is altered, the tool and procedures were adjusted and documented accordingly.

Validity. Validity describes the strength of logical connections between the theoretical frameworks, empirical data, such as observations, causal patterns and real-world situations. While many research methods, relying on large quantitative data sets, establish validity with statistical analyses and generalizability to show connections, design development studies have unique validity challenges to address. First, DDS establish construct validity by connecting the theoretical framework, or theory of action, which describes the problem and possible solutions to the implemented intervention (Mintrop, 2016). In the case of this study, the change drivers regarding learning and motivation served as a connection between the need to reduce an implementation stance and defensiveness. Internal validity, the strength of causal relationships between the research design and the outcomes, was complicated in DDS since the implementation of the design remains fluid. However, this was overcome by carefully documenting process data and relating the intervention designs (Mintrop, 2016). In this study, facilitator reflections, personal learning plans, feedback notes, and recordings were used to increase internal validity. Each of those three pieces of process data related to impact data areas: signs of teacher learning comfort and reflection, ability to self-analyze teaching and set next level learning, and skill building respectively. The final connection is potentially the most tenuous as skills were built with the vignettes, personal learning plans and experience of others in the group. Based on a theory of action and based on making a connection between impact and process, design development results in a plausibility argument.

A third factor affecting validity, external validity, is easy in one respect and difficult in another respect to maintain in DDS. Design studies take place in natural settings, thus findings emanate from real life. But given the naturalistic settings of DDS and the complex nature of the relationships involved in educational systems, it is difficult to connect conclusions to other contexts. However, learning may be appreciated by the greater educational research community through careful description of interventions and documentation of process and impact data, while clearly describing the contextual conditions of the intervention site. Additionally, DDS are highly contextualized to specific educational settings, the studies have a high degree of external validity for those specific contexts (Mintrop, 2016). For example, other urban secondary schools may look to this DDS for ideas regarding science teacher learning at school sites.

Transferability. As previously mentioned, external validity is low in DDS, due to the uncontrollable natural settings of the research. Without a large sample size and controlled experiment, it is difficult to generalize results as applicable of the population in general. Further, each educational research project is so unique, it is impossible to control external factors and replicate an implementation. However, a carefully crafted DDS is able to provide an understanding of the organizational context and change, and possibly suggest interventions that may be viable in similar contexts. In order to ensure some level of transferability, the connections between the theory of action, process data, impact data, and real-world context must be strong. These measures of research rigor can be easily disrupted, especially in action and design development research. Given this design is occurring with a science department in an urban secondary school with challenges regarding student population, resources, and capacity, there is promise that other schools may find the research valuable. Eliminating the reliance on expensive professional development additions and including the often-ignored teachers who do not volunteer should help the study transfer to a wider range of contexts, because it is unlikely that a school would not have the minimum requirements and constraints of the school in this study.

Threats to rigor. The validity and reliability of research can be affected by a number of researcher biases. While all research designs are prone to bias, action research and design development face particular challenges due to the desire of the research to positively influence the outcomes of the systems they are both acting within and researching (Mintrop, 2016; Coghlan & Brannick, 2001). Researchers acting as facilitators must carefully guard against two main types of bias. Advocacy bias, the tendency to claim greater effect or causal relationship from an intervention, is a risk to educational leaders conducting research (Mintrop, 2016). In my study, advocacy bias had the potential to develop as I am attempting to shift the learning of science teachers. Professionally, I have a great incentive to prove useful in this endeavor. Additionally, I had to guard against observer-expectancy bias, the devotion to my proposed solutions as answers to the perceived problems. By definition, I am proposed an intervention that I thought would have some effect on teachers' learning stance and improvement of lessons. However, I tended to internal validity regarding the connection between the intervention and impact. Each precaution was essential to safeguard against bias.

The observer-expectancy bias can be reduced by continually referring to the connections of the problem to the theoretical base of the design and structuring process and impact metrics at the beginning of the study (Mintrop, 2016). Advocacy bias requires some additional care, especially when the roles of designer, implementer and evaluator are conflated. If they can be separated, advocacy bias is reduced because the three separate individuals check the work of one another (Mintrop, 2016). In a design such as mine, the carefully crafted metrics and co-design of the intervention helped to reduce the observer-expectancy bias. Advocacy bias remained a major concern in my study, as I will act as primary designer, implementer, and evaluator. A rich corpus of action research rigor informed strategies to reduce this bias. Coghlan and Brannick (2001) suggest careful cycles of reflection during the implementation, followed by a critical analysis where assumptions are put forth and disconfirming data is explicitly sought.

Relationship to research subjects. My relationship with the participants was altered by implementing this design. By creating transparency with the participants about the purpose of the study, to investigate NGSS implementation with actual classroom teachers, stronger rapport and openness to participation was established. Overall, I made many attempts to level barriers between me, the researcher practitioner, and the teacher participants to foster a collaborative partnership.

Intrusion into the teacher group. Many participants have heard of the work I have done previously for NGSS implementation. The intervention allowed them to engage in further extensions of NGSS implementation. One concern was that I was seen as an authority figure. Data may have been influenced by the difference in position. These factors were mitigated with thoughtful transparency regarding the purpose and design of the study. However, I did not do much to reduce the factors that shaped my experience as a teacher. In particular, the degree of privilege I brought to the implementation due to my dominant culture identities. I toiled to approach the work empathetically and technically to reduce the overt effects of my identity. In addition to my cultural background, I brought a set of teaching and learning experiences to intervention with deeply rooted constructivist ideologies.

I asked the group of teachers to agree to participate fully in the meetings and activities. This was incentivized by working with their principals to negotiate for this work to happen during their regular workday.

Participants had some reservations regarding the reporting process. Since they will know the facilitators and participants personally, they could potentially link data to specific participants. Norms of confidentiality were established. Additional measures were taken to anonymize the following data and analysis, including using pseudonyms for district, school and teacher.

Chapter Four: Data Analysis and Findings

In this design development study, a series of seven workshops was implemented attempting to shift teachers' motivation to learn by creating the conditions necessary for incremental improvement and professional inquiry. The following chapter documents the design development process using both impact and process data. Subsequently, connections are drawn between the key findings and trends in the impact data to critical moments observed in the process data. Table 4.1 summarizes the objectives and data sources for each workshop in the intervention.

Table 4.1: Relationship between workshops, impact, and process data.

Workshop	Anticipated Outcomes/Objectives	Impact Design Lever Addresses	Sources of Process Data
1	Understand phenomena-based learning as an NGSS-aligned, engaging strategy. Generate ideas for phenomena. Identify characteristics of your teaching style.	Self-assessing science instruction Setting next level learning	Low inference field notes and audio recordings. Survey including teaching style reflection and phenomena-based instruction.
2	Identify areas of work and strategies to try that will enable learning to improve teaching.	Self-assessing science instruction Setting next level learning Learning safety in groups	Low inference field notes and audio recordings. Exit questions to determine influences on individual learning plans
3	Determine a next level learning goal aligned to your teaching. Understand how to use classroom inquiry for incremental improvement. Finish a learning plan with a strategy and data source.	Sense of learning quality Setting next level learning Learning safety in groups	Discussion questions: What did you consider when choosing a learning plan goals and strategy to try? Low inference field notes and audio recordings. Learning plans.
4 and 5	Learn from classroom experiments of colleagues. Receive feedback to inform next steps.	Sense of confidence, motivation and efficacy	Low inference field notes and audio recordings.
6 and 7	Understand the link between strategies and evidence. Determine sources of information to best inform teaching practice. Improve learning plans for another round of classroom experimentation. Improvement seen between first round of goals and evidence presentation to the second round.	Setting next level learning Sense of learning quality Sense of confidence, motivation and efficacy Learning safety in groups	Responses to Evidence-Strategy activity. Low inference field notes and audio recordings. Shared evidence statements. Low inference field notes and audio recordings.

Within this chapter, the process data and impact data are thoughtfully documented, analyzed and interpreted to identify relationships between the baseline condition of the participants, planned design, actuated design implementation, degree of implementation attainment of intended purpose, and the outcome condition of the participants. First, the process data collected during the seven workshops is analyzed and interpreted to determine discrepancies between the workshop intention and the realities of implementation. Next, this chapter explores the progression of participants in terms of the five domains of the design (evaluating teaching, setting next level learning, sense of motivation, sense of learning quality, and sense of learning safety), culminating each section of impact data with a conclusion identifying deficiencies in the design and achievements linked to the process data.

As seen in Table 4.1, the design workshops begin with learning engagement around deeper science pedagogical practices before moving into personal learning goal writing and inquiry into science teaching practices. Through two iterations of planning and reflection, the design intends to shift participants from their current generalized state of externalizing lack of improvements in classroom teaching due to factors such as student behavior, skill level, lack of time, poor use of collaboration time, to a state in which they take control of their learning and achieve a deeper state of goal setting, technical reflection, and motivation to improve toward NGSS-aligned pedagogical practices within a high functioning collaborative group.

In order to document the impact, semi-structured interviews and a video observation performance assessment were conducted with seven high school science teachers before and after the series of workshops. These sources comprise the baseline and outcome data analyzed for participating teachers' shifts over the five dimensions. Process data, including low inference field notes, audio recordings, teacher responses to questions, teacher learning plans, and student artifacts is analyzed to find critical moments of learning and growth during the workshops.

Process Data

The following section analyzes several forms of process data, primarily audio recordings, field notes, and participant artifacts from each of the seven workshops. Through the analysis, critical incidents are highlighted to exhibit key points in the implementation of the intervention potentially giving insight to the features of the design which had influence on the decisions made during implementation and the outcome. The process data is presented for each workshop sequentially. Table 4.1 summarizes the sources of process data. The subsequent sections describe the decision points for changes and key pieces of evidence found in each workshop by comparing workshop plans and expectations for the workshops to low inference observations of the workshops, identifying discrepancies between the plan and the actual experience.

Workshop one. The opening workshop of this design development study included two primary activities intended to improve the evaluation of science instruction and determine next level learning. During the workshop, low inference notes and audio were recorded as sources of process data. Additionally, a closing survey was administered to gauge participants' self-evaluation of teaching and willingness to try phenomena-based instruction.

Planned activities and anticipated outcomes. First, participants worked with several examples of science phenomena that could be used in a classroom setting. As described in Chapter 2, science teachers learning about high-quality, NGSS-aligned strategies, such as phenomena-based lessons, is seen as critical step in determining the types of lessons needed to academically engage students and move towards effective NGSS-aligned strategies focused on the technical aspects of teaching science. These phenomena-based demonstrations can serve as examples of strategies for next level learning goals and high-quality instructional strategies to support evaluation of teaching quality for the participants. This activity was meant as a hook to engage participants with high quality examples of NGSS-aligned science instruction. During the second activity, participants review five vignettes of teaching styles (Appendix B). The vignettes cover a wide range of styles represented on the Teacher Typology Matrix (TTM) (Appendix A). This activity is intended to provide anchor points from which participants could self-identify their teaching style within the four categories of the TTM: teaching style, activity coherence, relevance to students, and responsiveness to students. The design workshops intentionally avoid explicit instruction and awareness of the TTM in order to diminish expectancy bias of the participants when reporting their teaching styles. Determining current teaching style is seen as a precursor to setting next level learning.

Observations. During the workshop, several critical incidents were captured in audio recordings and field notes. First, I presented three example phenomena. The participants treated the examples as specific ideas for lessons. This was evident as they critiqued the activities and discussed the specific units in which they would use the activities. Despite my attempts to redirect with the question, 'How might activities like this inspire student engagement or lead to student questioning?", participants continued to discuss the different units in which they could use the specific activities. For example, Kathy remarked, "I could use the slinky activity during my unit on energy and waves." This may indicate a need to have specific science activities as examples for the participants. Additionally, the use of the activities as a hook may have been insufficient to describe the type of phenomena-based learning intended.

Second, Marcus and Nancy spent the entire small group discussion time expressing their frustration with students' lack of basic algebra and reading skills before ending with a quick reflection, "How could we do something like this conceptual for students who don't get the math?" The participants spent significantly more time discussing issues of student behavior and skills than ways the teachers could modify instruction to help the students. Finally, a few teachers struggled with technology meant to enable reading of the vignettes and complete the reflection survey. The results of the survey are explored below.

Vignette reflection. All of the participants in this intervention were present and submitted a response to the vignette reflection. Four of the seven participants saw components of their teaching in a combination of multiple vignettes. Table 4.2 contains the connections to teaching style made by each of the participants.

Table 4.2: Self-identification with vignette styles (Appendix B).

Participant	Vignette(s) identified as most like their teaching style
Claudia	Disjointed implementer and Open inquirer
Emily	NGSS-aligned constructivist
Kathy	Disjointed implementer
Marcus	Didactic lecturer and Front loader
Nancy	Didactic lecturer and NGSS-aligned constructivist
Sara	Disjointed implementer and NGSS-aligned constructivist
Tony	Front loader

The "disjointed implementer" and NGSS-aligned "constructivist styles" were most commonly referenced by the participants. The other styles were referenced twice each. Interestingly, the majority of the participants referred to multiple vignettes to represent their teaching style and even added additional information. Marcus stated, "I am most like the more controlled vignettes (Didactic lecturer and Front loader), but I like to lead students to answers with hints." Claudia listed numerous student activities, such as making comic books and white board quizzes, not represented by the vignettes. Participants provided various rationale for their choices that rely on the types of activities, such as, "I like to check in with my students' knowledge like Stevens…"

Finally, all participants mentioned the desire to try some type of phenomena-based teaching in their class by stating, "Yes. [I am likely to try phenomena-based learning]," adding comments such as, "I think it gives the students more motivation as they relate it to real life," and "I would like to plan a whole unit with phenomena. I want to learn how I can plan instructions with phenomena as the base line and etc."

Interpretation. Workshop one was designed to begin to shift the amount and quality of teaching evaluation and establish anchor points for setting next level learning and teaching goals. During the workshop, participants approached the phenomena as activities to implement rather than representations of a high-quality strategy. Further, they cited low student skills and poor behavior and reasons they couldn't attempt more advanced strategies. For these reasons, the activities did not achieve a movement towards quality evaluation of teaching as comments remained centered around basic student ability and compliance.

However, teachers began to open up on two fronts. First, they used the vignettes to begin to describe the teaching strategies and styles they employ in their classrooms, thereby using some language helpful to analysis of classroom teaching. Second, they expressed a willingness to learn about and try phenomena-based teaching, a sign of next level learning goal motivation. Analysis of process data from subsequent workshops should determine if phenomena-based teaching is incorporated as a learning goal and classroom pedagogical experiment.

Overall, workshop one demonstrated limited effectiveness towards the established goal within both the engagement activity and the diagnostic vignettes. Regarding the opening activity, teachers reduced the activities to their exact purpose, rather than expanding and creating additional examples. Activities were taken at face value while perceived student attributes hindered application of the NGSS-aligned principle.

During the diagnostic vignette portion, two additional challenges to the design emerged. First, the descriptive vignettes may have been perceived literally. By focusing on specific examples, participants refused to identify the vignettes with general teaching types, thereby not identifying deeply with a typology of teaching. The teachers looked at superficial features of the teaching and remained focused on specific instructional strategies. Subsequent analysis is needed to investigate the lasting effects of this superficial analysis of teaching on the remainder of the design.

Workshop two. The primary focus of the second workshop in the implementation focused on setting next level learning goals with additional foci on learning safety in groups and self-assessing science teaching. A review of vignettes, writing of professional learning plans and peer feedback were the featured activities. Process data during this workshop comprised field notes, audio recordings, and responses to brief exit questions.

Planned activities and anticipated outcomes. During the second workshop, participants revisited vignettes with the addition of a learning plan for each scenario in the vignettes. With the assumption that participants self-assessed their own teaching styles and strengths in the first workshop, the activity attempted to provide inspiration for the types of next level learning based on the styles represented by the five vignettes. As seen in Appendix B, each vignette provided a possible next level learning goal for professional development, thereby providing an example from which participants could build their learning plans. As demonstrated previously, providing guidance but allowing autonomy in goal setting and demonstrating learning goal language may generate a deeper sense of autonomy and self-determined motivation to achieve changes in learning.

Following the vignette review, participants presented their learning plans to their colleagues. In this activity, participants were exposed to additional ideas from their colleagues and in turn helped their colleagues refine the strategies and evidence used to learn about the success of their classroom experiments, thus tapping into their responsibility to peer learning.

Observations. The participants entered the second workshop with a noticeably diminished energy level. As the facilitator, I sensed a need for a few minutes of general lamentation characterized by discussion of the overwhelming workload they faced. Five of the seven participants began the workshop by informally discussing the pressures of grading and the need to preserve time for their own work; Marcus and Emily were not present for the workshop. At the close of the discussion, participants were prompted to review the vignettes and identify one example from their teaching that made them associate with the vignettes. Nancy asked for more time to read, because she had forgotten which vignette she related to the previous week. After a few additional minutes, each participant shared an example of their teaching related to one of the vignettes. Nancy, for example, related to the vignette of Ms. Lewis stating, “My class is very interactive. They review problems, then make some models, and do an experiment.” Additionally, Tony looking forlorn remarked, “I’m mostly doing lecture,” identifying with the Mr. Stevens vignette. With the exception of Tony, the discussions between teachers were lively as they recounted the work of the week.

The second activity of the workshop featured time for participants to develop work plans, talk to peers, and speak with me regarding professional learning plan development. Participants were encouraged to review the learning plans attached to the vignettes (Appendix B) and create a small scale, repeatable strategy to implement in their classrooms with corresponding evidence documentation. Additionally, the plans were meant to include a focal problem or teaching challenge, learning goal statements for the teachers, and ideas for addressing the challenge. Claudia reacted, “We aren’t going to teach the same thing in two weeks, how can we write a plan,” indicating a possible misunderstanding of the scope of the inquiry or the intervention tapping into PD fatigue. She engaged me in a conversation about the types of problems she was encountering, specifically “dealing with hitchhikers (students copying work of more advanced students).” After describing the problem, she settled on attempting a Kagan grouping strategy to put students of more similar ability together. Kagan strategies focus on student collaboration and discussion strategies meant to encourage participation and achievement for all students. The Bay Vista staff had recently completed a training on many Kagan strategies. She finished with a general complaint that observing group interactions was too much additional work to perform during a class. Several of the participants suggested using student evidence from summative assessments. After 5 minutes of work, three of the teachers began discussing the behavior of the sixth period classes, concluding, “We are tired from the other classes, and they (students) are just finished learning.” This theme of engaging students generically, not in science specific activities, repeated multiple times in the study. I redirected them to complete their learning plans and prepare to share ideas with the whole group. The workshop concluded with a whole group discussion summarized in the subsequent section.

Interpretation. Following the implementation of workshop two and analysis of the process data, I noticed two salient trends which are addressed and interpreted here. First, I discuss the use of the vignettes to create learning plans, analyzing the self-evaluation of teaching strategies to design goals. Second, I connect the expectations and ideas presented by participants to the engrained notions of cycles of inquiry.

During the workshop, participants reviewed and commented on the components of colleagues’ individual learning plans (Appendix B). I analyze the learning plans in more detail in the context of Workshop Three, after participants had a chance to complete their ideas regarding strategy and evidence to be collected. As participants shared their teaching style, current classroom foci, and ideas for next level strategies, a few noticeable patterns emerged. First, participants aligned the vignettes to a snapshot of their teaching that day. The goal of the vignettes was to characterize teaching styles, thereby facilitating the identification of next level learning goals for the teachers. The fluctuating identities obscured the focus on a specific learning goal.

Second, three participants suggested the use of test scores as the source of data, indicating a misunderstanding of appropriate evidence on the learning plan. Student performance data would best be considered teacher performance goals, not learning goals. Additionally, two other participants described "better student engagement" during lecture as a point of focus for their learning. The most student-centered area of learning came from one participant who stated, "I want to increase the student talk and academic language during large group discussions." The focus on summative student performance is expected, due to the typical way cycles of inquiry are conducted in school settings; a common strategy is agreed upon, and teachers compare summative student performance.

The final piece of process data, the exit survey, had an unreasonably low response rate, which prompted me to ask these key questions as warm ups in future lessons. The survey asked, "What did you consider when choosing a problem, learning goal, and strategy to try? What was most influential in your decision?" Interestingly, one respondent named Kagan strategies, those introduced at a recent whole-school professional development session. The other respondent talked about ease of implementation as the most influential characteristic of strategies they want to try. While the first respondent named student engagement as a rationale, these responses and the statements during the workshop point to a general desire to try simple, easy, or comfortable strategies while looking at summative test scores or general "student engagement" as indicators of learning. This does not indicate a move towards defining next level learning which would be typified by more immediate student responses during activities designed to engage them academically.

The participant responses and discussions during workshop two indicated a strong inclination towards compliance driven accountability with summative assessment. The design may have lacked sufficient activities to shift the mindsets of participants to focus internally on science related teaching practices, the area of control for teachers.

Communication with department chair. Between the second and third workshops, the department chair raised questions with me via telephone regarding her personal learning plan and group sentiments regarding the learning in the first two workshops. Given that co-design can be a powerful tool in design development studies, the conversation had implications for the design of workshop three, which was originally slated for the initial round of presentation of learning. The department chair expressed concern that the teachers did not embrace the work. She concluded that the participants would appreciate more examples of specific strategies to try and more individual help. She suggested that I offer up individual consultations to all participants and preview the meeting agendas in advance. Workshop three was modified to address some of these concerns. Additionally, the offer for individual consultation was extended to each participant, none responded, and all meeting agendas were posted at least three days in advance of the meeting.

Workshop three. The third workshop in the series involved increasing learning quality, setting attainable next level learning goals, and developing safety to learn in groups. Three activities formed the basis for learning and change: a fishbowl conversation with the department chair presenting student work, an incremental improvement didactic explaining the importance of collecting timely student evidence and responding to student need, and an opportunity for peer feedback. The following sections contrast the planned activities with the process data collected during the workshop, including: opening discussion, low inference field notes, and professional learning plans.

Planned activities and anticipated outcomes. The third workshop, in this design aimed to increase teachers' sense of their own learning quality by detailing steps in incremental school improvement. This is seen as an important step to understand learning quality. Further, the fishbowl featuring the department chair sharing her learning plan and student work served as an attempt to increase a sense of learning safety by decreasing the lofty expectations and lengthy classroom experiments talked about in previous sessions and focus on smaller scale, achievable outcomes. This activity was added following a discussion with the department chair and my critical friend. The fishbowl emphasized the focus on teacher learning goals, as opposed to focus of many of the participants' plans, student learning outcomes. In the following analysis of process data, responses to opening questions, discussion during the workshop and learning plans were reviewed.

Observations. During the opening discussion, reviewed below, participants all shared their thinking regarding implementing new strategies in the classroom. Following the opening discussion, the department chair shared her work around asking questions to encourage participation in class.

The fishbowl was a conversation between the department chair (Emily) and me, wherein I guided Emily through reflective dialogue regarding her plan and the evidence collected. During the discussion, all of the participants listened intently and interjected some clarifying questions and suggestions. A key learning moment came near the beginning of the discussion when reviewing Emily's learning goal. I reminded Emily that the goal should indicate a learning that she wants to accomplish as a teacher. We discussed changing the wording from, "Students use academic language in response to discussion questions" to "learn the best strategies and types of questions to improve the quality of student conversation." Claudia and Marcus exclaimed points of clarity as they stated they had not previously understood how to write learning goals for themselves. Following that moment, the discussion flowed freely.

After the fishbowl, I reiterated the three key points which highlighted the importance of using the learning plans to make incremental improvements: 1. write learning goals to describe what the teacher should learn from trying strategies in the classroom; 2. try promising classroom strategies that have the potential to improve science learning; and 3. plan to collect evidence that will provide insight into the strategy and learning goal. Participants then worked with partners based on the types of strategies they named on their plans.

During the peer sharing and feedback portion, the three points from above were used as discussion stems to provide feedback to peers. The 20 minutes went quickly and were full of lively discussion sharing ideas and strategies. I spent the time with Sara and Tony who had not yet completed a draft learning plan. During that time, I helped Sara come up with a discussion strategy that will reduce the risk of sharing to whole groups, so students can share their understanding. Tony started by discussing the challenge of students wanting to just name the answer in discussions when he really wants them to describe their understanding of science concepts. I coached both participants to plan strategies that had students explaining their scientific thinking to other students, in a lower risk environment. Following the workshop, I asked participants a reflection question.

When responding to the question, 'What do you consider when choosing a personal learning goal and strategy to try?' participants responded with two main themes: the amount of student engagement and the amount of science content covered. First, student engagement was explained by participants as the amount of time students are engaged in the tasks. To turn this student-focused goal into a teacher learning goal, one could try several strategies to determine which works best to engage students in academic science conversation. This is exemplified by Marcus, "I was going to try a Rally Coach. I was going to try and do a strategy to help students get a little clearer on it. It's a Kagan strategy, because we had an inservice on it. I use it because it gets them to talk through it and think about how to do a Lewis structure." With regard to content covered, participants were focused on the idea that successful teaching is indicated by the amount of material you are able to cover in class, not the depth of the material nor student understanding. The comments made by the participants are consistent with the following analysis of their learning plans.

A review of the participants' learning plans revealed a few salient trends. Four of seven participants named a Kagan strategy as the focus of their teaching experiment. This is especially compelling given that Sara and Tony did not name a strategy in their plans. Kagan strategies were easy to implement and avoided deeper science teaching goals in favor of student participation and compliance-oriented goals. Second, every plan contained a mention of student engagement or participation as the problem. There has been a district-wide and school-wide focus on student academic engagement which could affect this preference. Third, most participants exhibited strong alignment between problem, strategy, and evidence in their plans. For example, one Emily proposed trying three discussion strategies to see which produced the most academic conversation. As evidence she observed various student conversations. However, two participants exhibited poor alignment; Kathy and Marcus described student work artifacts, such as posters and problems on worksheets to analyze the quality of discussion students are having in collaborative groups. These forms of student artifacts did not yield evidence pertaining to student discussion. Given this misalignment between goal and evidence, I adjusted Workshop Six to contain an activity exploring examples of appropriate connections between strategy and evidence.

Interpretation. In summary, Workshop Four contained intervention steps enabling learning and slight shifts to participant routine, such as the fishbowl observation and learning plan process. However, the depth of academic focus within the learning plans was not evident. In some cases, the evidence lacked alignment to the proposed learning.

The use of the fishbowl discussion was meant to raise participant awareness to the ease and utility of short term experimental learning and reflection cycles in the classroom. It served to reduce anxiety about the scope of the inquiry, based on feedback from participants reviewed previously. Yet the activity lacked persuasive influence to push participants toward attempting NGSS-aligned pedagogical strategies. The strategies in the fishbowl were discussion strategies meant to elicit student scientific thinking, but participants seemed to embrace the short cycles alone, not the focus on evidence of student science learning.

While the professional learning plans exhibited internal alignment between goal, strategy, and evidence, deeper science pedagogical practices were noticeably absent in the plans in favor of Kagan engagement strategies. Several participants were concerned with improving student engagement in science learning. Although there is some evidence that strategies were chosen with purposeful connection to science learning, many seemed to choose Kagan strategies due to the explicit learning in a school-wide professional learning workshop and the desire to for student compliant engagement. In the analysis of the following two workshops, actually evidence presented will be connected to learning plans.

Workshops four and five. The fourth and fifth workshops in the series were designed to contribute to participants' sense of confidence, motivation, and efficacy by focusing on learning successes occurring in their classrooms as a result of teacher experimentation. The primary activity in both workshops were presentations with peer feedback opportunities. Process data stemmed from field notes and audio recordings of presentations and feedback with a focus on examples of student work demonstrating success within the professional learning plans. Additionally, a group discussion question from the close of workshop five was recorded and analyzed below.

Planned activities and anticipated outcomes. Presentations intended to allow participants to evaluate themselves and others while using a feedback protocol designed to remove judgmental language and focus on the learning, thereby improving the sense of learning safety in the group. The following analysis used field notes, referring to audio recordings when necessary, to glean moments of learning and patterns in participants' experiences and growth.

Prior to the workshops four and five, participants were sent a reminder with simple suggestions to focus their presentations. The proposed outline for presentations minimized the description of the activity, suggesting a focus on the student evidence and even more time spent receiving feedback from colleagues. Feedback from colleagues was structured with sentence stems meant to diminish judgmental language and focus on low inference observations, such as: I noticed/heard (something directly from their presentation of evidence) that made me think (something regarding positive student learning).

In the first round of presentations, Marcus Claudia and Tony presented their strategies and student evidence. During workshop five, Nancy, Sara, and a teacher not participating in the design development study made their presentations.

Observations during presentations. During the presentations in workshop four, Marcus and Claudia both presented activities requiring students to complete lengthy worksheets. Both examples required students to provide specific answers demonstrating knowledge transmission.

Although Claudia is supposed to be focusing on collaborative structures for student success and Marcus on student discussions during a Kagan rally coaching session, both presentations focus on the technical functioning of the activities. For example, both focused on the questions that students answered incorrectly and how many of the questions were left incomplete. Emily attempted to redirect them both with questions like, "Claudia, what was the learning goal for you, the teacher, during this time." Marcus spent most of his presentation discussing the procedure used for discussion, instead of the observations of student participation, and the results on the exam he gave the next day. He concluded that the strategy seemed to help lower performing students on the test, but he lacked evidence to show this causal relationship.

Tony's presentation was more aligned to the desired outcomes for the teacher presentations, despite not completing a learning plan in the previous session. Tony guided his colleagues to look at student-generated models of ecosystem energy, noticing that many students are copying from the book. He raised this as a major concern and discussed attempting a peer review strategy during future modeling activities to avoid copying and enhance student created models. He did not expand on this idea in order to present his second set of artifacts, an online curriculum activity several grade levels below his course. Tony's colleagues spent the remainder of the session critiquing the rigor of the activity, describing it as a coloring activity and generally not producing evidence of student learning at grade level. The critiques forced Tony to justify his choice of activity by referencing the lower cognitive demand of the activity. This may indicate participants' use of student evidence to assess a classroom activity, but the comments revolved around the procedure and type of activity instead of the student evidence produced as a result of the activity.

During workshop five, Nancy, an additional teacher not fully participating in the design development study, and Sara presented their activities and student evidence. Nancy presented observations of students working collaboratively alongside quiz scores, concluding better quiz scores resulted from the groups working more collaboratively. She discussed issues with managing the group work with little reference to signs of success for students. Sara presented a collection of student work from an activity tasking students with internet research. While she highlighted one major student misconception regarding the scientific principle being taught, she spent nearly the entire time talking about the quality of drawings and aesthetic of the project. Neither factor helps determine the students' conceptual understanding. The other participating teachers suggested ways to improve the activity instructions and limitations of internet research, such as students copying internet images instead of producing original work that demonstrates conceptual understanding.

Following the presentations, participants responded to the question, 'What evidence did you see in the last two meetings that gave you insight into something that was working well to engage students?' Three of the participants gave responses highlighting the need for more in-depth instructions for students, alluding to the need for simplified, step-by-step instructions. An additional two participants complained that students were difficult to engage in lessons. Tony mentioned the need to have students working and producing something real. Nancy highlighted the increased use of academic language among students in her class.

Interpretation. The observations and responses to the reflection question detailed above illustrated the need for deeper connections between the planned design intervention, specifically the connections between a desired science instructional learning goal for the teachers and the evidence collected during classroom teaching. Additionally, the teachers' plans exhibited little structure and cohesive science pedagogical strategies, indicating a deficit in the implementation of the design. Further, many of the presentations featured a cursory look at student evidence while focusing on descriptions of how an activity is constructed, especially the instructions for an activity. This was also true of the feedback provided by colleagues. These four conclusions yielded insight into teacher learning needs and understanding of the intervention phases.

Due to the frequent use of general engagement strategies, such as Kagan strategies, seen in the presentation, I suspected a need for a more suggestive menu of strategies to encourage deeper, NGSS-aligned instruction. This conclusion was based on two assumptions. First, the teachers are choosing strategies for which they had received direct training. Second, using new science strategies is a higher order skill for the teachers and may require additional guidance, whereas engagement strategies were lower risk for teachers.

The second major challenge observed in teacher presentations and plans was the connection between activity and evidence. In too many cases, the evidence was related to the science content, such as a quiz at the end of a lesson, but not linked directly to the teaching strategy, using group discussion to increase academic language for example. Since the evidence connection to the strategy attempted held great importance for teacher learning, focusing on student need, and motivating teachers to continue with successful strategies, I strived to alter the design to provide archetypes of acceptable evidence.

Finally, the peer presentations aimed to foster a safe and collegial learning environment while observing possible learning strategies from other professionals. The examination of the colleagues' classroom experiments yielded safe and critical commentary about the detail of instruction or the type of activity needed to address specific science content. While this generated many strategies and activities for participants to try in the future, the feedback lacked attention to the student learning presented in the artifacts. Although this may have been a result of the poor selection of evidence, I suspected a need for teachers to develop skills related to providing low inference feedback based on evidence of classroom instruction and student artifacts. Workshop six attempted to clarify the purpose and acceptable foci of the learning plans and classroom experiments by using an activity with example strategies and corresponding preferred forms of evidence.

Additional feedback from participants. Following workshop five, the science department chair informally asked participants for feedback. Although this was unplanned, it provided valuable insight as to teacher thinking in the middle of the implementation. There was an increased chance of honest feedback as well, since the researcher-facilitator was not present. Here is what the department chair reported:

I asked the department how they were feeling about our work with you and they gave very positive feedback. You've managed to organize us in a way where we are functioning much better as a team. We are all getting a glimpse into each other's classrooms along with finding a way to grow ourselves. They're also getting better at questioning and looking for evidence of student learning. (Department Chair, personal communication, December 21, 2017)

This quote was emblematic of the intervention successes and shortcomings. The participants were deepening their ability to meaningfully participate in a cycle of inquiry based on student learning but did not deepen their actual science teaching practice.

Workshops six and seven. To begin the final two workshops of the design implementation, I inserted an activity on choosing evidence for the documentation of teaching improvement, due to the observations in Workshops Four and Five. Following the activity, participants created a second learning plan, and presented student artifacts to document success of these strategies enumerated in the learning plans. These activities aimed to build on the next level learning and motivation from deepening feelings of efficacy. Additionally, an anticipated increase in the sense of learning safety in groups was expected from the peer review process. To determine the discrepancies between anticipated process and participant experience, several forms of process data were analyzed with highlights described below including: field notes and audio recordings, analysis of learning plans, and a final group reflection question with a focus on the type and quality of evidence used in plans and presentations.

Planned activities and anticipated outcomes. The evidence choice activity was designed to help participants evaluate the quality of teaching and simultaneously improve their ability to choose next level strategies and define their sense of learning quality, by discussing connections between science-specific learning goals and the types of evidence collected from students. The presentation of artifacts had similar goals to those in workshops four and five: improve learning safety in the group, increase motivation through the experience of success, and evaluate instruction for next level learning. Four sources of process data were analyzed in the following section to determine the extent to which intended outcomes for the workshops were met. Analysis begins with a review of second learning plans and comments from an evidence strategy activity. The analysis continues with tracking important comments during the presentations of student artifacts. It concludes with an analysis of comments from a summary discussion of the design experience for the participants.

Observations of proceedings. Participants began the sixth workshop with an activity designed to make them consider possible ways that evidence matches to attempted strategies (Appendix E). In the activity, they matched strategies such as small group science discussions and providing science feedback on projects with evidence types including video clips or samples of student writing. While strategies and evidence were chosen intentionally as archetypes for the learning plans, the subsequent discussion remained open for participants to explore their own ideas regarding strategy and evidence. Marcus asked for clarification on a few of the listed activities. Participants read silently for five minutes and began chatting, signaling completion. During a whole group share out, participants commented on the activities as concrete ideas for implementation in their own classrooms and described the connection to the evidence. For example, Nancy mentioned wanting to try a phenomena first activity and using student writing samples and audio of group discussions due to the ability to identify student misconceptions from the activity. Participants concentrated on the ideas spoken by their peers and nod along in agreement.

During the next section of the workshop, participants created their second learning plan. I reminded them that the problem is students focused, but the learning goal is focused on the teacher learning. Additionally, I encouraged them to focus on a strategy that could be repeated quickly and tracked easily that may also build on their first learning plan. Claudia and Kathy called me over for personal assistance to help discuss their ideas. Following work time to complete the beginning of professional learning plans, participants shared their goals and strategies to help divide into two groups to receive feedback and ideas on their professional learning plans. One group was created to work on collaborative grouping strategies to enhance learning, composed of Claudia, Nancy, and Marcus. While Tony, Kathy, and Sara worked on eliciting scientific knowledge and thinking in discussion. Emily was absent for this workshop.

During the small group discussion, participants were instructed to present their problem, strategy and learning goal so that peers could then discuss the best type of evidence to collect during the classroom implementation. Marcus and Claudia expressed concern that there wasn't enough time to enact a plan, because the participants were still envisioning longer term projects, not experiments on an hourly or daily scale.

While participants shared their ideas for their next classroom experiment, colleagues provided feedback focused on the implementation of activities rather than the type of evidence to collect. For example, Claudia discussed wanting to see how models improved if she gave students a peer feedback protocol. Tony and Nancy suggested giving students a rubric and creating model drafts on paper before producing an in-depth model. Most frequently, the teachers discussed Kagan strategies as the base of their experimentation. The discussion of plan focused on the implementation of a strategy without careful attention to the type of evidence that would be most feasible and productive. This resulted in teacher learning which revolved around basic instructional changes, not student learning.

Interpretation. All of the participants logically connected a strategy to evidence. As teachers began to write learning plans, they were prompted to share their goals and activities planned for the next cycle. The most common strategies were still Kagan structures, but several of them were now thoughtfully connected to science activities, such as: using a Kagan discussion strategy to have students explain scientific phenomena visuals in class. Analysis of the second round of learning plans follows.

The second round of professional learning plan development had small but desirable improvements compared to the first round. Generally, each of the participants seemed to push themselves with a next level learning activity. Although two participants, Claudia and Emily, did not complete a plan, the remaining plans successfully connect a strategy to something the participants wanted to learn about their classrooms. For example, Kathy was interested in students' ability to explain a science concept seen in a model or visual. She planned to use three Kagan strategies to determine which strategy would elicit the deepest student science conversation. She was one of four participants to focus on Kagan strategies to explore science learning. Additionally, three of the participants were focusing on an NGSS Science and Engineering Practice in their plans. Marcus used an interesting science phenomenon to introduce and engage students. Although the plans exhibited many positive advancements, three plans did not name evidence, indicating a need to increase the quality of evidence being evaluated.

During workshop seven, six of seven participants brought student artifacts to share. Only Marcus was unable to prepare something due to his testing schedule. Overall, there were four NGSS-aligned strategies presented and more intentional data collection than in the first round of presentation. Logistically, the participants were split into two groups so everyone was able to present within the time allotted. While the goal of the workshop was to increase motivation by observing student successes, get feedback on next level learning, evaluate teaching, and increase the learning safety by presenting to groups of colleagues, mixed success was attained. Appendix F contains the protocol followed during the small group presentations. This protocol was intended to shift participants away from getting caught up in details or discussing student deficit, which often led to peer comments solely about changing the instructions provided for an activity, not the student learning. The protocol encouraged discussions of student evidence, often focusing on deficits in student learning, which is assumed to diminish motivation and feelings of efficacy.

One particular presentation serves as an exemplar for the possibility during this workshop. During Kathy's presentation, participants used the provided protocol to discuss the student evidence collected. Even though many moments exhibited one of the general patterns seen in all presentations, an overemphasis on the description of how a learning activity was designed instead of the evidence from student artifacts, Marcus used the protocol to redirect the group's conversation on two occasions. With the group refocused on the evidence, Emily remarked, "Looks like you were pretty darn successful with earth science kids. They got a lot of experience from the round robin (Kagan strategy)."

Kathy responded, "I never taught this unit like this, with so many models and so much engagement."

This serves as evidence of two positive outcomes from the design. First, a protocol used to investigate student artifacts shifted slightly towards student learning. The natural tendency of the participants was to discuss the procedural aspects of teaching, but the protocol was used to redirect conversation to student artifacts on a few occasions. Second, the designed workshops provided motivation for Kathy to learn about new strategies and track evidence of student learning. Although not every presentation was as successful, two other positive patterns emerged from the event.

First, there was an increase in the presentation of student artifacts with connections to the strategies being implemented. Tony included video clips of small group discussions in his classroom. Four others provided physical student work artifacts for review. The second, and potentially more significant observation was the amount of comfort and frequency of constructive feedback to colleagues. The conversations became more focused on ways to improve student learning and engagement.

Despite those gains, a few questions from this design development remained unanswered following the presentations. Although most participants brought in artifacts of learning, some of the artifacts were lacking in quantity or quality by focusing on a few successful examples, or only being described from the teachers' perspectives. Nancy, for example, only brought in one high quality example and explained characteristics of low quality examples. Providing low quality student work samples could have deepened the conversation. This indicated uncertain feelings of safety by sticking to a “safer” piece of student work. Finally, four of the participants used Kagan strategies in their professional learning plans and presentations. They deepened the connections to science, but a question remains about the need to provide explicit science teaching strategies during the intervention.

Following workshop seven, the participants had an opportunity to respond to the following questions in a whole group discussion format:

- What are your main takeaways from our work together during the first semester?
- How has this experience affected your teaching?
- How did working with your colleagues help you this semester?

The following collection of responses provides insight into the overall learning during the intervention.

> “Started rewriting lesson plans to incorporate more strategies that I was aiming to practice. Looking for ideas and more hands-on strategies for collaboration.” (Kathy)
>
> “Gives us more structure. We were sharing but all over the place. Let's us copy some strategies from each other.” (Nancy)
>
> “Get a little window into your world that we don't usually get. We were able to solve problems together.” (Emily)
>
> “Enjoy being introduced to how others are doing things. We shared a little before, but the organization lets us get more done.” (Marcus)
>
> “Look for ways to get evidence of lesson effectiveness. What can I do in the planning to be able to get real time feedback.” (Tony)
>
> “Got ideas from colleagues on doing my 3D model, that I should have criteria and a grade sheet to be more effective and get more of what I want out of the project. I got ideas that I didn't even think about.” (Claudia)

As seen in the comments above, participants have a sense of accomplishment based in activity and strategy sharing, reviewing evidence of student learning, and developing a more professional level of collegiality.

Summary of process data. During the seven workshops, the teachers in this design development study experienced several events leading to some shifts in their thinking about science instruction and in their approach to professional learning. Primarily, they increased their level of sharing actual teaching successes and struggles through the use of student artifacts. Additionally, participants began to open their ideas of successful teaching and learning from the perspective of student summative achievement to a more open inquiry based on their goals for their students on a daily basis. The structures and examples provided during the intervention, such as the learning plan, fishbowl conversation, and peer feedback questions were useful to the participants.

The shortcomings of design development studies present unique learning opportunities following reflection on the study. While this design began to shift the teachers' perceptions of their own professional learning and signs of growth, the study fell short in a few key areas. As exhibited by the quality of the artifacts and the questions teachers asked, the teachers dwelled on surface structures and procedures rather than deeper science learning and pedagogy. A focus on providing technical feedback, using language aligned to a tool like the TTM, could provide teachers with the confidence to analyze a lesson. Inherently, a focus would land on student learning. There was a tendency in the study to describe in detail the rationale and procedures for implementing specific strategies. Additionally, the strategies attempted stemmed from discrete strategies presented in other workshops. This illuminated an opportunity for me in subsequent iterations; I predict teachers to benefit from more examples of NGSS-aligned practices and explicit instruction, with follow-up practice, on ways to provide feedback and generally analyze a lesson for student learning.

Impact Data

In order to determine the overall effect of the design implementation, two sources of impact data were analyzed in this study. The first source, assessing all five dimensions, was a comparison of participant responses from baseline to outcome semi-structured interviews. In addition to self-reported assessment scores on a scale of one (not at all) to five (great extent), the interview responses were compared to a rubric and evaluated qualitatively for phrases that exhibit rationale for rubric evaluations.

The second form of impact data, the performance assessment, included questions designed to ascertain participants' abilities and growth in three of the domains by having them analyze and reflect on video observations of other teachers. Their responses were coded in vivo and ranked with the rubric and teacher typology matrix (Appendix A), noting the level of variation between the baseline and outcome data. In the following sections, analysis of the participants' performance assessments and structured interview responses are aggregated by domain.

Finally, impact and process data were compared to identify outcomes plausibly linked to implementation activities and outcomes that may be explained by forces outside of the design. This careful analysis was necessary to reduce the bias towards success often seen in design development studies. Each section discusses the anticipated shifts successfully made, those with little to no growth, and additional questions.

Analyzing teaching. Two components are analyzed in this section to ascertain shifts in the quality of participants' evaluation of their own teaching. The design attempted to improve teacher's skill at assessing the quality of their own teaching as well as others. First, responses to three structured interview questions and follow up probing questions are analyzed in two ways: a) participants provided a self-score for each question b) transcripts of the interviews were qualitatively analyzed using a three-stage rubric designed to show stages of growth in this dimension. While the analysis takes place at the individual teacher level, the reporting is limited to aggregated trends within the domain. Key examples representing each stage of the rubric and highlighting various amounts of growth are captured below.

Table 4.3: Self-reported quantitative analysis of Analyzing Teaching

	Claudia		Emily		Kathy		Marcus		Nancy		Sara		Tony	
To what extent are you satisfied with your teaching?	3.5	4	4	4	3	4	3	3	4	4	4.5	5	3	3
To what extent do you accurately assess your strengths and areas for growth in your classroom teaching?	nr	3	4	4	4	4	2	2	3.5	3	4	4.7	4	5
To what extent would you classify yourself as an effective teacher?	3.5	4	nr	4	3	4	3	4	nr	3.5	4	4.7	nr	3
AVERAGE CHANGE = sum of outcome scores - sum of baseline scores / number of scores	+0.5		0		+1		+0.5		-0.25		+0.63		+0.5	

Note: Self-reported scores were on a Likert scale of 1 (Very limited) to 5 (great extent). Domain with baseline scores on the left and outcome scores on the right of each participant's column. An entry of "nr" indicates the participant did not provide a quantitative evaluation nor a verbal equivalent according to the scale for the question.

Quantitative analysis. During the semi-structured interview, three questions were asked of all participants (Table 4.3). Although the questions are summative in nature, they were designed to elicit a rating to observe any large changes in the feelings participants have towards their teaching and ability to assess teaching. Subsequent probing questions, analyzed below, allowed for deeper explanation of the ratings. As seen during the structured interview responses, five of seven participants (Tony, Claudia, Marcus, Sara, and Kathy) reported small amounts of positive growth over the three questions; one participant (Emily) had no net change; and one participant (Nancy) reported a slight decrease in the metrics for evaluating teaching. Further qualitative analysis of participants' responses to probing questions reveals patterns useful to determine the types of changes seen.

Qualitative Analysis. The rubric in table 4.4 was used as the basis for analysis of the "Analyzing Teaching" domain. As seen in Table 4.4, the domain was analyzed based on participants' level of description of successful teaching with general descriptions involving compliance and management on the initial side, while thorough conceptions of teaching in the developed category involve technical critique aligned to the subject. Within this domain, the TTM was developed as a metric for the depth of technical description used, it was not explicitly taught, therefore participant responses and performance assessments are evaluated for signs of technical language used in the self-evaluation of teaching, such as hands-on, science and engineering practices, small group discussion, lecture, and checks for understanding. Although the structured questions did not immediately elicit comments aligned to the rubric and TTM, probing questions such as, "What do you think an outside observer would see?" and "How did you choose that rating?' gave insight into the participants' internal rationale and thought process when analyzing teaching.

While analyzing interview responses with regard to the rubric below, several patterns emerge as common participant responses. Five participants (Emily, Kathy, Marcus, Sara, and Tony) exhibited improvement from the baseline to outcome interviews. Claudia and Nancy made comments that aligned to the emerging category in both the baseline and outcome interviews.

Table 4.4: Analyzing teaching rubric

INITIAL	EMERGING	DEVELOPED
Does not describe teaching in terms of the teacher typology (TTM) (Appendix A). Refers to completion and compliance issues more than issues of understanding. Places blame on students for teaching challenges.	Identifies challenges but externalizes blame for those challenges. Occasionally refers to student experience but focuses on compliance and general engagement rather than deeper learning. Describes successes from a teacher perspective, rating the quality of the activity, without considering the impact on students.	Accurately and consistently uses TTM (or similar language) domains to describe teaching. Identifies challenges and successes based on student experience and artifacts. Locates the challenges within the teacher's control.

One salient shift between impact and process data occurred in that area of externalization. As detailed in Chapter Two, externalization is the positioning of outcomes and events in the classroom as effects of anything outside the teacher's immediate control. Overall, the outcome interviews exhibited far fewer mentions of externalization; externalization was only noted five times in the outcome interviews, with difficult to teach students being mentioned most frequently. During the baseline interview probing questions, similar student-based externalization was mentioned eight times in conjunction with other externalization references, such as poor work environment and limited resources mentioned four times. The reduction in frequency of externalizing forces mentioned demonstrates a shift from the initial towards the developed categories of the rubric. Key characteristics described by participants in baseline interviews related to issues of compliance and student attributes aligned with the emerging and developed categories.

Additionally, the interview responses were dominated by descriptions of teacher-centered approaches with very few descriptions of\ student experience and learning. Although the process data showed a greater focus on students, when prompted to discuss their own teaching, descriptions remained on superficial activities not connections to student learning. One teacher provides an example of externalization by saying, "It used to be a better school. The demographics changed. And also it's now chemistry for all... I've got third graders and it's just that's teaching in California, that a school like this. You're going to have you know third through 13th grade like in a chemistry class." Although an extreme case, the interview response highlights student experience as the external force reducing the quality of learning, thereby demonstrating an initial level of analyzing teaching, as the focus was on the effects of student ability on the quality of the teaching rather than identifying what was happening in the classroom.

Table 4.5: Qualitative analysis of Analyzing Teaching

	Structured Interview Analysis		Performance Assessment Analysis	
	Baseline	Outcome	Baseline	Outcome
Claudia	Emerging	Emerging	Emerging	No response
Emily	Emerging	Developed	Emerging	Emerging
Kathy	Initial	Emerging/Developing	Initial	Emerging
Marcus	Initial	Emerging	Emerging	Emerging
Nancy	Emerging	Emerging	Emerging	Emerging
Sara	Emerging	Developed	No response	Emerging
Tony	Emerging	Developed	Emerging	Emerging/Developed

Note: Summary of teacher responses in baseline and outcome data from the evaluation of structured interviews and the performance assessment.

The shift in participant ability to evaluate teaching is noticeable in the movement towards more developed descriptions of teaching appearing in structured interview responses. Descriptions of teaching aligned to categories in the TTM (Appendix A) are mentioned 11 times during the baseline interviews and 26 times in the outcome interviews. In addition to a doubling in frequency of mentions, the participant comments shift from a focus on procedural compliance and basic student engagement to identification of new practices that deepen student learning of academic skills and content as exhibited in the quotes below.

In the baseline interview Kathy stated, "I would say 80 percent of them get it or they're on task and accomplish the objective...It's 60 percent on average that the kids are working...I want to design a lesson that can help students be engaged completely. I'm working on it. I break down the content into activities and whatnot, everything that goes with it."

Whereas, in her outcome interview she stated, "They are on their task whether it's a hands-on activity, or finding answers from the textbook, or trying to plot the graph. Any assignment I'm giving them, they are able to at least try, because of the strategies that I'm using. Giving them enough information and guiding them before I'd like them work independently."

Kathy's quote is emblematic of the shift to evaluate teaching with more technically descriptive language about the types of activities rather than simply discussing the amount of student engagement. Many of the participants exhibited a similar change in language used to describe teaching. These changes indicated teacher learning in the area of "Analyzing Teaching" due to the deeper and more technically descriptive language used to analyze teaching.

Four unexpected responses were mentioned by multiple participants during the baseline and outcome interviews, which may provide insight into traditional teachers' views when analyzing teaching styles. First, three participants in both baseline and outcome interviews mentioned needing feedback from outside observers to understand what they should improve in their classroom. Two teachers in each set of interviews discussed the importance of test scores to evaluate their teaching. This observation does not align to an improvement in the ability to self-analyze teaching style, areas of strength and areas of improvement, due to the reliance on external sources of feedback.

Teachers' ability to plan an effective lesson as a skill and the teachers' knowledge of science subject matter being taught as qualities of effective teaching emerged as the third and fourth types of comments. While both areas are highly relevant to effective teaching and teaching improvement, the rubric for this session focuses on observable aspects of the teaching process, since this is seen as the foundational skill for the subsequent aspects of the design study.

While the structured interviews demonstrate a decrease in externalization and increase use of technical language, the results of the performance assessment were not as clear. This analysis was conducted based on responses to three questions asked in both the baseline and outcome video observations: 1. What did you find effective in the lesson? Why is this effective? 2. What is an area of growth for this teacher? What makes you say this? 3. If you were giving this teacher advice, what would you recommend? What should she try to improve? What might she learn about to improve? Only Kathy exhibited signs of improvement by focusing more on the student outcomes presented in the videos as evidence of her evaluation. The balance of the participants exhibited comments that aligned to the emerging category in both baseline and outcome data. Additionally, few participants use descriptive language of the technical aspects of teaching and the student experience, such as those listed in the teacher typology matrix (Appendix A).

Overall, the teacher self-reported scores match well with the qualitative coding of interview responses. The noticeable shift in self-reports represented the changing perspectives through which participants viewed teaching. However, the relative stagnation in the performance task represented the nominal shifts in actual teaching performance, as observed in the presentations of teacher learning goals and student artifacts in the process data section. The qualitative analysis of the structured interviews provides salient insight into potential shifts in participants' skill when analyzing and evaluating teaching. Participants exhibited more technical language and analytic precision to describe teaching within their locus of control, as opposed to describing extrinsic barriers and feedback. In short, teachers appeared to shift their ability to analyze teaching by recognizing their role in creating meaningful activities for students and focusing on the technical aspects of student achievement and science activities.

Connections between impact and process data. Improvements in the ability to analyze teaching domain were seen both quantitatively with teacher self-reports and in the qualitative analysis of the interviews and performance task. In particular, positive shifts were noticed in the location of challenges within teachers' control (a reduction in externalization) and the use of technical language to describe teaching style and relevance to students (TTM, Appendix A). With regard to the shift in perceived locus of control, three experiences in the design process may have contributed to the shift. First, thoughtfully creating a plan for improvement (professional learning plan) pushed participants to identify a teaching challenge and design a strategy to address it, rather than externalizing the causes of diminished student learning. Second, presenting artifacts of student learning shifted the accountability for success to the activity rather than student ability or behavior, especially as several participants noticed increased engagement from their activities. Finally, providing colleagues with feedback based on student work began to open teachers to critically analyze their teaching. The second shift, analyzing teaching description using more technical language about specific science pedagogy, benefited from similar phases of the design process. Although the TTM was not explicitly taught to teachers, the activities they attempted were aligned to it, especially in areas of teaching style, learning formats, and responsiveness to students.

Ability to set next level learning. The second domain of the design attempted to shift participants' ability to determine their current level of teaching and set a reasonable learning goal for improvement. The impact data for this domain consisted of two self-reported responses to the questions seen in Table 7 along with qualitative analysis of the interview responses. Additionally, the final two questions of the performance assessment were analyzed. 1)What new strategies might the teacher try? 2) What could be a source of evidence to determine if that strategy was effective for her students? Qualitative analysis of recordings and transcripts was analyzed for alignment to the rubric components in Table 4.7 and the Teacher Typology Matrix (Appendix A).

Quantitative analysis. The two questions below serve as a foundation for deeper discussion through subsequent probing questions. The questions are valuable in determining general feelings about the participants' goal setting and tracking summative feelings of improvement. As seen in Table 4.6 below, participants claimed limited growth in the domain. Only Emily and Sara report an increase in their goal setting, with Kathy reporting a slight decrease and the others remaining at the same level. In the cases of "nr" responses, the participants gave qualitative descriptions of goal setting that did not correspond to a quantitative response, increasing the importance of the subsequent qualitative analysis. The following qualitative analysis will highlight trends and characteristics of participants' responses that may illuminate this lack of change.

Table 4.6: Summary of self-reported scores in Ability to Set Next Level Learning

	Claudia		Emily		Kathy		Marcus		Nancy		Sara		Tony	
To what extent do you set goals regarding your teaching?	nr	nr	4	5	5	4	nr	2	nr	4	4	4.2	4	4
To what extent do you achieve your goals? If you don't set goals, to what extent does your teaching improve?	nr	4	nr	3.5	3	3	nr	3	nr	3	4.5	4.5	3.5	3.5
AVERAGE CHANGE = sum of outcome scores - sum of baseline scores / number of scores	NA		+0.5		-0.5		NA		NA		+0.1		0	

Note: Domain with baseline scores on the left and outcome scores on the right of each participant's column. An entry of "nr" indicates the participant did not provide a quantitative evaluation nor a verbal equivalent according to the scale for the question.

Qualitative analysis. The qualitative impact data related to this domain consisted of two parts, analysis of semi-structured interview responses and responses to two questions on the performance assessment. This discussion begins with analysis of the interviews along the gradient of the rubric seen in Table 4.7. The following analysis discusses the alignment between participant self-reported ranking and the transcripts, while identifying underlying trends in the responses. Examples are provided to illustrate the rationale for qualitative rubric analysis.

Table 4.7: Rubric for coding impact data in Ability to Set Next Level Learning

	INITIAL	EMERGING	DEVELOPED
Ability to set next level learning	Does not set goals for improvement. Goals are set for students but not related to teaching. New activities attempted do not represent a move up the TTM	Sets performance goals and/or confuses learning goals and performance goals. Goals set that do not address improvements to teaching. Weak alignment to the TTM or only one new activity attempted represents a shift up the TTM. Goals are not used for reflection.	Accurately and consistently identifies a learning goal to shift in one or more of the teacher typology domains. Uses learning goals to make improvements to instruction. Three strategies attempted align to the TTM strongly.

While little change is observed in self-reported scores, Emily, Kathy, Nancy, Sara and Tony exhibit signs of improvement from baseline to outcome in the structured interview analysis. The structured interview with Marcus demonstrates a decrease in ability to set next level learning, while Claudia mentioned using summative assessment and “trying new lesson types” without specifying types or decision processes, thus demonstrating “emerging” status in baseline and outcome metrics.

The performance assessment exhibits increased development in “ability to set next level learning” for Emily, Marcus, Kathy, and Nancy. Claudia and Sara did not participate in one portion of the assessment. Overall, the structured interview analysis yields a shift from an initial condition of lacking goals or basing goals on external sources (principals or student summative performance) to more immediate learning goals surrounding teaching practices with measurement based on student performance during or immediately following lessons. Overall, participants provided responses aligned to the “emerging” category of the rubric for this domain, similar to their self-responses. Some trends and characteristics of the responses were incorporated into the following qualitative analysis.

Table 4.8: Summary of qualitative responses in Ability to Set Next Level Learning

	Structured Interview Analysis		Performance Assessment Analysis	
	Baseline	Outcome	Baseline	Outcome
Claudia	Emerging	Emerging	Emerging/developed	No response
Emily	Emerging/developed	Developed	Emerging	Emerging/developed
Kathy	Initial	Developed	Emerging	Emerging/developed
Marcus	Emerging	Initial	Initial	Emerging
Nancy	Emerging/Developed	Developed	Emerging	Emerging/developed
Sara	Emerging	Emerging/developed	No Response	Emerging/developed
Tony	Emerging/developed	Developed	Initial/emerging	Initial

Overall, the structured interview analysis yields a shift from an initial condition of lacking goals or basing goals on external sources (principals or student summative performance) to more immediate learning goals surrounding teaching practices with measurement based on student performance during or immediately following lessons. These shifts are exemplified when comparing Sara’s comments from baseline and outcome data:

> “I have complete goals for the entire year that I'll be covering so and so.” (Baseline)

> “The students actually get the concept or not, and how do you know they're different. Some students can get it conceptually. Some students actually understand it mathematically. Sometimes some students get it, you know, when they do experiments…There are different ways how to get the concept.” (Outcome)

Further evidence of a shift in next level learning and goals emerges from analysis of the responses with regard to the TTM. The structured interview transcripts reveal a higher degree of alignment to characteristics in the TTM with regard to “next level learning.” Participant comments contained mention of TTM characteristics 10 times in the baseline interviews compared to 22 times in the outcome interviews. The majority of these shifts occurred in the areas of teaching style and responsiveness to students, in order of decreasing frequency. Teaching style is referred to with comments such as “I’m using hands-on labs,” (Sara) and “Doing more collaborative learning and revising of science models” (Kathy). With regard to responsiveness to students, Nancy and Emily mentioned tracking student conversations using academic language.

Emily states in her outcome interview, “I definitely set goals and I work on those and with my sheltered students I'm really trying to get them to talk. I mean I think-pair-share is crickets in my room right now… So, I'm really trying to encourage them to take a chance it's OK to be wrong. You know that's how you learn. You learn from your mistakes. You know I'm really trying to get them to think that way... You know they just want to get the right answer and they want the praise. You know so I have to really think about how to praise them. And make that more important than what they turn into me."

Another trend observed within the interview responses for this domain involved the type of indicators used to establish goals and next level learning. In the baseline interviews, there were six mentions of measuring the achievements of goals and improvement based on summative evaluations of students or principal evaluations, with only three mentions of more immediate student-centered metrics of success to guide next level learning. In the outcome data, there were twice as many mentions of student learning on the daily or more frequent level, being used to understand classroom instruction. Kathy provides an example of setting goals around student learning, "Recently I have started focusing on writing skills. For my students to analyze the data that they have and to use the vocabulary in scientific times and all the scientific thinking process. The thought process. In putting together the information to analyze the data."

Finally, the NGSS, as a representation of cognitively-demanding science teaching, appeared in several transcripts. In a few cases, the new standards were described as motivating forces. Tony described his experimentation, "The goal is trying to introduce engineering practices. How do I know I do that? You know you've done a couple of projects and made that the focus. Emphasis on models is another goal. So, we do the pond water models. We're doing all these different models. And I'm getting some evidence that students are seeing that you can have multiple ways of representing things."

Although Marcus named the NGSS as a point of frustration, "This year, this year quite honestly survive, straight up survive. Everything's been thrown up in the air and I don't know. I mean, I can create a curriculum. I've taught chemistry before, but I've been thrown into space science. I don't know what an H.R. diagram *is.*" Clearly, there is an incomplete understanding and assimilation of the NGSS that may have benefited from anchoring the learning in deeper science pedagogical tasks aligned to the NGSS.

Connections between impact and process data. Two primary changes were observed in the impact data that support the observed increase in goal setting. First, a noticeable rise in self-determination of goals was observed as opposed to those based on supervisor evaluations or test scores. Second, goals were increasingly set to respond to student needs with immediately observable metrics of success the following discussion returns to the process data to determine potential connections to these two changes.

As reviewed in the previous section (Table 4.1), next level learning and goal setting were the foci of activities in workshops one, two, three and six. Within those workshops, the creation of learning plans and feedback from peers created lively discussion or the participants. These activities likely contributed to both shifts in goal setting that were observed. However, a few of the participants who showed positive changes in setting next level learning may have benefited from individual consultation with me. Claudia, Emily, Sara and Tony had the opportunity to individually discuss their plans with me. All but Claudia showed improvements in the dimension. Further discussion, in Chapter 5, will consider the role of the facilitator in changes seen during implementation.

Sense of confidence, motivation, and efficacy. In the domain titled "Sense of Confidence, Motivation, and Efficacy", four questions (Table 4.9) were analyzed to determine potential shifts in participants' view of themselves as effective practitioners who are willing to attempt new strategies. Additionally, the domain investigated motivational forces, especially rooted in perceived successes. Following analysis of the participant self-reported quantitative data, a qualitative analysis of the questions looks for evidence of progress along the rubric in Table 4.10 below, while also identifying relevant trends and notable unique responses. Analysis of change in this domain and the remaining two domains relies on the structured interview process alone, as the performance assessment was not created with a metric for these domains.

Table 4.9: Summary of self-reported scores in Sense of Confidence, Motivation, and Efficacy

	Claudia		Emily		Kathy		Marcus		Nancy		Sara		Tony	
To what extent do you see improvements in your teaching practice as time goes on?	nr			5	4	4	nr	3	nr		4	4		4
To what extent do improvements in your teaching motivate or inspire you to keep improving?	5	5	4	5	5	5		4	5	5	nr	5	5	5
To what extent are you comfortable with learning and trying new things in the classroom?	4	4	5	5	5	5	2	3	4	5	4.5	4.5	5	5
To what extent do you believe that you can effectively teach all students in your classes?	3.5	3	4	4.5	4	4	3	4.5	3.5	2	4.5	5	3	5
AVERAGE CHANGE = sum of outcome scores - sum of baseline scores / number of scores	-0.25		+0.5		0		+0.75		-0.17		+0.17		+0.67	

Note: Domain with baseline scores on the left and outcome scores on the right of each participant's column. An entry of "nr" indicates the participant did not provide a quantitative evaluation nor a verbal equivalent according to the scale for the question.

Quantitative analysis. The participant self-reported quantitative values yield small changes between pre- and post-interviews, with Marcus, Tony, and Emily reporting the largest growth. Kathy reported no changes throughout all questions, while Nancy and Claudia exhibited small decreases in their reported scores. Notably, Nancy had large but opposite changes in two questions, yielding a small overall negative change. The remaining participants reported relatively similar levels when analyzing each individual question. Transcripts of the structured questions and probing questions in this domain were analyzed to identify trends and verify the self-reported scores.

Table 4.10: Rubric for coding impact data in Sense of Confidence, Motivation, and Efficacy

	INITIAL	EMERGING	DEVELOPED
Sense of confidence, motivation, and efficacy	Participants rarely try new strategies in the classroom. Participants do not perceive an ability to affect outcomes for students.	Attempt new strategies but do not persist to improve those strategies or use them in the future. Takes some responsibility for student improvement and engagement	Try new, risky and innovative strategies frequently. List concrete evidence of student learning and improvement. A great sense of ability to change student outcomes.

TABLE 10. Rubric for coding and quantifying structured interview responses and performance assessment in the area of "Sense of Confidence, Motivation and Efficacy".

Qualitative analysis. The overall changes between baseline and outcome interview responses aligned with the self-reported scores. Small positive changes were observed for most participants, excepting Claudia and Emily with responses matching the "emerging" qualities in the rubric (Table 4.11). Marcus, following a baseline interview with particularly low demonstration of characteristics of the domain, exhibited the greatest growth in the outcome data. Further analysis of the trends is detailed below.

Two main patterns were identified as shifts from the baseline to outcome data. A noticeable shift in the motivational reason for attempting new strategies happened between the beginning and end of the design implementation. During the baseline interviews, three teachers mentioned student skills and behaviors as primary obstacles to teaching success. In the outcome interviews, a few extreme student cases were mentioned as obstacles without generalizing to the entire student population. This indicated movement towards a greater sense of confidence and ability to affect change according to the rubric. Strategies that were being implemented by the teachers were aligned to their perceived student need rather than for compliance to the NGSS or management of the classroom. For example, Marcus stated, "I'm getting a better idea of how students see the world, so I can better plan lessons that mesh with their perception."

The second major trend in the "Sense of Confidence, Motivation, and Efficacy" domain was in the quality of the strategies being used with the students. For example, Kathy identified a strategy as reducing rigor so that students can accomplish the task, exhibiting low confidence and efficacy in the ability to teach grade level material. In the outcome interview, she provided examples ranging from intentionally grouping students for collaboration to "storming around the room to check student learning". Emily recognized the "need to pique student curiosity" in the baseline interview and exhibited a more complex understanding of the student needs by providing an example of strategies attempted with some success to increase English Language Learners' understanding of science.

Table 4.11: Summary of qualitative responses in Sense of Confidence, Motivation, and Efficacy

	Baseline	Outcome
Claudia	Emerging	Emerging
Emily	Emerging	Emerging
Kathy	Emerging	Emerging/developed
Marcus	Initial/emerging	Emerging/developed
Nancy	Emerging	Emerging/developed
Sara	Emerging	Emerging/developed
Tony	Emerging/developed	Developed

In subsequent sections, a third noticing from the structured interview analysis is discussed. Four of the participants mentioned using Kagan strategies to address learning needs with mixed success. While the reflections demonstrated logical connections between the strategies and student success and an interest in altering the strategies to provide the best possible student experience, the process data was consulted for connections between this design implementation and the technical training received on Kagan strategies held by the school at the beginning of this study.

Overall, the shifts in this domain demonstrated increasing levels of confidence and motivation to try new strategies. These shifts were also aligned to a deeper sense of efficacy to work with all students. As stated in Chapter 3, one of the key activities to build motivation and a sense of efficacy relied on reviewing evidence of signs of success with students. Scant evidence suggested this was accomplished during the implementation. The second area of concern in this domain stemmed from the type of strategies being attempted. A reliance on structural and procedural strategies took precedence over deeper, more cognitively demanding science pedagogical strategies. Although precise causal relationships were not established, these two areas of deficit point to weaknesses in the design and my responsiveness to participant need during the intervention. Course correction could have included more modelling of the expectations for identification of successful teaching strategies and repeatedly connecting learning back to the opening vignettes with the addition of a menu of deeper science strategies to attempt. The following analysis compares these key findings to the participants experience during the implementation.

Connections between impact and process data. Overall, observed shifts in the domain of Sense of confidence, motivation, and efficacy, demonstrated a movement by several participants toward an increased level of confidence and feeling of efficacy to reach more of their students while also feeling motivated to attempt more cognitively demanding lessons with their students. The primary activities meant to bolster this domain occurred during workshops wherein participants shared activities and received feedback. Although they expressed great interest in sharing their classroom experiences, workshops four, five, and seven exhibited relatively little discussion of positive experiences with students. Therefore, the shifts should not be attributed to portions of the implementation designed to shift this domain. However, an unintentional effect of simply sharing artifacts and stories about increasing cognitive demand of activities may have created the shift. This effect is discussed in the discussion of the fifth domain, sense of learning safety in groups.

Sense of learning quality. In the "Sense of Learning Quality Domain", participants provided scaled responses to two questions used as a metric for the shift in practice that has happened over the course of the last three months. The subsequent semi-structured interview questions were analyzed qualitatively to find patterns and feelings regarding the types of changes made in teaching practice and reflection practices that lead to learning and intentional changes in practice.

Table 4.12: Summary of self-reported scores in Sense of Learning Quality

	Claudia		Emily		Kathy		Marcus		Nancy		Sara		Tony	
To what extent have you intentionally changed your practice in the last three months?	3	4	1	3	5	5	nr		nr	3	5	5	nr	2.5
To what extent do you try new strategies in your classroom?	nr	4	nr	4	3.5	4	nr	2	4	3	5	5	3	3.5
AVERAGE CHANGE = sum of outcome scores - sum of baseline scores / number of scores	+1		+2		+0.25		NA		-1		0		-0.5	

Note: Domain with baseline scores on the left and outcome scores on the right of each participant's column. An entry of "nr" indicates the participant did not provide a quantitative evaluation nor a verbal equivalent according to the scale for the question.

Quantitative Analysis. The two questions used to assess this domain were designed to determine a specific area of learning, new teaching strategies, while additional probing questions were used to determine learning quality in other areas, such as recognizing student cues, planning, or evidence collection. Given the small number of questions in this domain and the limited use of the scale during self-reported responses, trends in change between the baseline and outcome are difficult to ascertain. However, it is important to qualitatively analyze the semi-structured interview responses to provide insight in to the participants who responded with very high self-score (Kathy and Sara) compared to the remainder ranking their sense of learning quality with "to some extent" or "to little extent".

Table 4.13: Rubric for coding the impact data in Sense of Learning Quality

	INITIAL	EMERGING	DEVELOPED
Assess the quality of what they have learned (strategies and reflection on improvement)	Learning is not related to classroom strategies. Rarely reflects on student successes. Reflection is about how they felt about the lesson taught. Little to no reflection given.	Desire to learn about strategies to improve teaching loosely related to student academic performance Sometimes reflects on implementation successes. Sometimes builds on activities that seem successful.	Choose pedagogical strategies to test in their classrooms and reflect upon. Consistently reflect on daily lesson implementation successes. Frequently try new strategies in response to observations of student need and challenges.

Qualitative Analysis. In terms of "sense of learning quality," five of seven participants exhibit signs of improvement during structured interview responses. Even though the participants did not demonstrate strong signs of systematically experimenting with strategies nor a habit of daily reflection, a few noteworthy trends are observed between baseline and outcome data. These analytical observations stemmed primarily from teacher explanation of their responses to the structured questions and discussion in subsequent probing questions (Appendix C).

In the baseline interviews, teachers exhibit signs of the "initial" category for a few reasons. First, teachers are generally evaluating their learning quality by tracking curriculum completion and summative test scores. Additionally, participants admit a focus on student compliance and behavior management as signs of success. Claudia mentions, "Things are pretty much the same lesson, just maybe a little bit improved where I thought it was good enough the way I did it... Well for example with the lab that we're doing with models, I feel like more and more people are figuring out how to make it unable to make it faster than they used to and not take as long to try and figure it out."

In the outcome interview, one salient change is most frequently observed; participants are determining the strategies they need to learn about and the success of strategies that they attempt, based on student needs and academic experience. In responding to the extent that practiced has changed in the last three months, Kathy describes, "Thanks to you. Because we are really geared up to set goals that we want to achieve. I've changed strategies a lot. I've changed to collect more student evidence. To make it more fun for them to give me the evidence, and to write a plan that concludes in evidence with active student engagement."

Table 4.14: Summary of qualitative responses to interview questions regarding Sense of Learning Quality

	Baseline	Outcome
Claudia	Emerging	Emerging
Emily	Initial	Developed
Kathy	Emerging	Developed
Marcus	Initial	Emerging/Developed
Nancy	Emerging	emerging/developed
Sara	Emerging	Developed
Tony	Emerging/developed	Emerging/developed

Connections between impact and process data. In the domain "Sense of Learning Quality," the intervention attempted to increase the professional orientation of teachers by encouraging them to reflect on their learning opportunities and successes. During workshops three and six, participants designed learning plans with the help of peers. The learning plans presented some challenging instructions at first, with many participants confused regarding the purpose of the learning goals. During the sixth workshop, conversations deepened during peer feedback sessions for learning plan development and strategies attempted. Given the improvement in the depth and engagement around planning, the portion of the learning plan implementation was likely related to the increase in the sense of learning quality exhibited in this section.

The "sense of learning quality" domain was given the least attention during the design process and subsequently showed minimal to no signs of shift in the impact data. The design process intended to instill the professional learning plan as a catalyst for daily planning and reflection. It is uncertain that providing more iterations would help, but the design could benefit from a deeper and more explicit explanation of the importance of professional learning plans, rather than the implemented cursory explanation relying on teacher in situ use of the plan. However, quality of learning could be linked to future motivation to continue with a professional learning plan such as this, so it is worth exploring expansion of this domain. Possible alteration or removal of this domain from future iterations of the project is discussed in chapter five.

Sense of Learning Safety. The final domain, "Sense of Learning Safety", investigated the amount of trust and growth that happens as a result of sharing with colleagues. Additionally, this domain contained a question regarding specific collegial sharing of student artifacts of learning. Following scaled self-responses, participants' semi-structured interview questions were analyzed qualitatively. The following analysis explains potential shift in the learning happening due to group collaboration.

Table 4.15: Summary of self-reported scores for structured interview questions in "Sense of Learning Safety

	Claudia		Emily		Kathy		Marcus		Nancy		Sara		Tony	
To what extent is there an open and honest relationship among science teachers, based on trust?	5	4	3	4	5	5	3	5	4.5	5	4.5	4.5	5	5
To what extent are you comfortable sharing your ideas and experiences with regard to teaching with other science teachers?	5	5	4	5	3	5	5	4.5	5	5	nr	5	5	5
To what extent do you share student artifacts (work samples, videos, observations) with other science teachers?	4	5	2	5	5	4	nr	2	nr	4.5	4	3.5	3.5	5
AVERAGE CHANGE = sum of outcome scores - sum of baseline scores / number of scores	0		+1.67		+0.33		+0.75		+0.25		-0.25		+0.5	

Note: Baseline scores appear on the left and outcome scores on the right of each participant's column. An entry of "nr" indicates the participant did not provide a quantitative evaluation nor a verbal equivalent according to the scale for the question.

Quantitative Analysis. The "Sense of Learning Safety" domain exhibits a large degree of improvement in self-reported scores between the baseline and outcome interview responses. Although all three questions were designed to elicit ideas regarding collegial interaction, the third question (Table 4.15) aimed to determine the type and depth of exchange between participants. The only exceptions are Claudia and Kathy who both score themselves near the top of the scale for every question. The subsequent qualitative analysis is important as a method of discerning between inflated self-reported scores and characteristics in the domain representing improvement from baseline to outcome.

Table 4.16: Rubric for qualitative analysis of Sense of Learning Safety

	INITIAL	EMERGING	DEVELOPED
Sense of Learning Safety	Sharing is often in one direction only and only of activities that have a low risk associated (e.g. activities from another source or activities that are associated with success) Advice not solicited. Interactions mostly social and supportive. Trust lacking or broken.	Shares some ideas and how those may be implemented in the classroom, but rarely shares artifacts of student learning. Advice is solicited for safe topics, such as student behavior, materials management, lesson ideas. Some interactions focused on pedagogy and other instructional issues.	Share more sensitive and revealing material about their teaching, such as student work and videos. Participants are receptive to constructive feedback from colleagues. Participants describe benefits of sharing student artifacts. Interactions push each other to think about ways to improve instruction.

Qualitative Analysis. During the qualitative analysis, responses to the structured questions (Table 4.15) and the probing questions (Appendix B) were transcribed, coded, and analyzed with regard to the rubric in Table 4.16. Analysis of interview transcripts for the "Sense of Learning Safety" domain exhibited similar levels of positive growth as the quantitative self-reported scores. Only two participants, Sara and Tony, remained at the same level between baseline and outcome, whereas the other five exhibited signs of positive growth in the domain. Unsurprisingly, the most common characteristic of interview responses was one-way sharing. Each participant mentioned providing lesson ideas, worksheets, resources, and the like to their colleagues. While one-way sharing persisted throughout the design study, two notable trends appear in the outcome data summarized in Table 4.17.

The mentions of sticking to safe topics reduced from four to two between the baseline and outcome interviews. As Nancy exemplified, "Whenever we do something that was like ahh this was so cool. We share that." The outcome data contains four times the mention of sharing student artifacts and asking for constructive feedback from colleagues.

Table 4.17: Summary of qualitative analysis for Sense of Learning Safety

	Baseline	Outcome
Claudia	Initial/emerging	Emerging
Emily	Emerging	Emerging/Developed
Kathy	Initial/emerging	Emerging
Marcus	Initial/emerging	Emerging
Nancy	Emerging	Emerging/developed
Sara	Emerging	Emerging
Tony	Emerging	Emerging

For example, Claudia mentioned, "I may not share test scores that are low, because I'm embarrassed," in the baseline interview. During the outcome interview, she recounted, "You know we're pretty honest about how it's going for us. I'm willing to show artifacts. They give advice on how to do it better. Give me new ideas I didn't think about. procedural suggestions. Work more with them on scaffolding ideas."

Finally, a level of distrust exists between some of the teachers. Most comments were targeted at one member of the team, not participating in this study, who was criticized for a lack of rigor and participation. Marcus' stated, "I'm the only one (of the chemistry teachers) who assign homework." Other teachers mention the lack of participation and personal disagreements from this colleague.

Connections between impact and process data. Within this domain, positive shifts were seen in the frequency and topics of materials being shared and critiqued. Although the participant conversations during workshops four five and seven focused more on the pedagogical strategies and science activities from the teacher standpoint, this represented a more vulnerable level of sharing than the initial condition of discussing socially safer topics such as: student behavior, one way sharing of activities, general social conversation, and the state of the school and public education. However, the depth of conversation did not get to the level of sharing the most sensitive types of student artifacts, classroom videos and audio. Given more time or a forced emphasis on these types of artifacts, participants may have accelerated to that level of sharing. As the design unfolded, it was obvious that participants were not connecting to more valuable student artifacts, creating the impetus for the strategy-evidence activity (Appendix E) from workshop 6. In addition to the opportunity to intentionally meet and established norm of sharing some student artifact, the fishbowl activity and subsequent first round of sharing served to increase the learning safety. The following chapter serves to continue a discussion of the valuable components of the design process that are worth implementing in future iterations.

Summary of key finding from impact data. While not necessarily attributable to this design development process, several strong patterns emerged within the five domains. Although self-reported scores are fraught with reporting bias towards success, scores in this design generally show improvement for four of the five domains, with more than half the participants awarding higher scores in the outcome interview than the baseline. Only "Ability to Set Next Level Learning" had zero or negative impact for most participants. This finding may be attributed to a lack of knowledge of the trajectory for science teaching improvement. The qualitative analysis yields detailed information gleaned from semi-structured interview transcript analysis. The following list represents the critical changes between the baseline and outcome data.

- A decrease on the emphasis given to external forces as a barrier to learning, motivation and success.
- Increases in the descriptions of teaching style and responsiveness to students.
- Shift from goals set by others to goals set to learn about students
- A greater sense of immediacy to observe student performance and outcomes in classes
- Increasing the amount of science-specific skills focused on during classroom experimentation
- Increase in the observation of student work and sharing of student work to learn about instructional improvements
- Move from safer, social interactions addressing common complaints to more discussion about student learning and classroom observations in the group

Additionally, several informative and unexpected observations exist in the impact data.

- Large number of Kagan strategies, specifically intentional grouping and discussion strategies, being implemented as general pedagogical approaches.
- Need for outside evaluation to understand and validate professional growth.
- Participants perceiving lesson planning and subject knowledge as teacher skills needing development
- Attention to procedure, organization, and classroom management
- Motivation stemmed from talking about improving lessons rather than focusing on positive student outcomes

Finally, the design was implemented in attempt to move participants within five domains. In several crucial areas, there was no apparent movement or a change of a lesser extent than anticipated. The following represented the most significant of these areas:

- Evidence of learning goals being used for teacher improvement
- Systematic use of high quality learning artifacts for reflection and improvement
- Advancement in student-centered science pedagogical practices

Theses primary shortfalls in the design intervention create deep questions in contrast to the listed positive outcomes of the design. Overall, successes can be explained by providing an environment for intentional cycles of inquiry. While participants responded favorably to the design, a deeper connection to science pedagogy escaped the scope of this intervention, suggesting a need for science-specific pedagogical practices to be included in the intervention. The intention of the design was to generate motivation for deep learning and reflection on a frequent, if not daily or hourly, basis grounded in artifacts of student experience from the classroom. With these greater goals left outstanding, careful reflection by the researcher/practitioner can provide alterations to the design in subsequent iterations.

As seen in the previous section, impact data was not always explained by the implementation of the workshops. Therefore, the following concluding remarks review the entirety of the design development study with thoughtful critique of the implementation design and suggestions for further iterations based on the implementation experience.

Chapter Five: Discussion

The discrepancy between practice and expectations represented by the new standards, CCSS and NGSS, is thoroughly outlined in the research (NASEM, 2015). This presents a significant learning and pedagogical change requirement for teachers. While not insurmountable, deeply rooted teacher attitudes, societal perceptions, habits of mind, and other motivational factors present significant challenges to the learning needs of teachers, especially science teachers. Working at the nexus of research, design and facilitation, design development studies provide a powerful academic platform to begin to address these challenges.

In the following sections, I discuss the key learning, the design challenge, theory of action, limitations, remaining questions, suggestions for further iterations and implications for research. In so doing, I unveil the strengths and weaknesses of this design, which serves to illuminate key learnings by simultaneously connecting advancements made by participants to the design principles and highlighting shortcomings of the design explained by the design components, context-specific challenges not addressed, and issues with my conflated role in the research.

Summary of Key Findings

The overarching objective of this intervention design was to shift science teachers from a repetition of status quo teaching and over-reliance on summative assessments into an open learning stance, characterized by classroom inquiry focused on deeper, NGSS-aligned science pedagogical strategies with timely reflection based on student artifacts of learning. Five focal levers attempted to shift teachers to the desired state; each is discussed below. Although this section mentions several positive outcomes from the intervention, the most salient lesson resulted from reflection on the shortcomings of the design. In the subsequent sections of this chapter, I name probable causes of the shortcomings and the implications for the design and design development.

As outlined in Chapter Four, the strongest positive shifts occurred within Ability to Evaluate Teaching, Sense of Motivation and Efficacy, and Sense of Learning Safety. Participants exhibited shifts in behavior toward desired states in the following ways. Most importantly, they began to internalize control of the academic content of the classroom, as opposed to continuing with traditional instruction techniques due to low student abilities. At a nearly equal level, participants opened their instruction to others with detailed descriptions of activities and sharing of student artifacts of learning. However, the depth of analysis of teaching and cognitively-demanding, student-centered science pedagogical techniques used remained more basic than anticipated in the design.

Although teachers were willing to implement Kagan collaborative learning strategies, suggesting a willingness to experiment for the benefit of student learning, they did not demonstrate scientific pedagogical content knowledge learning through experimentation. In conjunction, the peer feedback sessions exhibited minimal low inference descriptions of teaching and learning using technical language. Both the depth of science practices and technical feedback may be related to the potential for the chosen artifacts of learning to provide meaningful insights into the science classroom experience of students. Resulting conversations dwelled on the performative aspects of teaching, such as the clarity of instructions and the ease of student understanding. In the following section, I critically analyze the design development process to illuminate probable causes of these key findings

Design Development Analysis

As discussed in Chapter Three, Design Development Research presents an emerging style with promising informative potential within educational systems. By extending action research with more thoughtful and well-planned activities and data collection, research practitioners may gain salient insights about the success of interventions in educational systems. A crucial step in design development research is reflection on the aspects of the design related to the observed outcomes of the intervention. During this process, I connected aspects of the intervention success and challenges to my identification of the problem at Bay Vista, problem and promising approaches in the literature, and the context at Bay Vista, before turning to suggestions of intervention activities to keep and those to change.

Problem of practice reflection. This design development study pivoted on the need for intense teacher learning at a time of new standards implementation and lingering effects of high stakes accountability and managerial control. While the Bay Vista teachers exhibited attitudinal and behavioral patterns consistent with this need for teacher learning, the needs assessment conducted for this study proved inadequate for a few reasons. First, the participants were not completely unmotivated to learn, as evident by their willingness to try Kagan strategies following a schoolwide workshop. However, I did not anticipate the need to explicitly teach many deeper scientific learning strategies, assuming the participants had been exposed to a sufficient list of advanced science pedagogy. As the intervention unfurled, I noticed a relative lack of technical language being used by participants to describe the artifacts of teaching. This was not a salient need in neighboring districts but presented as an obstacle in Bay Vista. Additionally, I assumed prior experience with cycles of inquiry would be a benefit to the learning planned in this intervention. Alternatively, the notions of cycles of inquiry relying on long term projects analyzing summative student data proved difficult to shift. A more thorough assessment of these experiences and attitudes could have helped me adjust workshops to address this need. Despite the challenges created by the inadequate needs assessment, I remain confident that greater perceived challenge, motivating teachers to engage in deep learning through classroom experimentation and observation, was the correct focus.

Theory of action critique. In this section, I outline crucial learning in three areas of the theory of action before turning to a discussion on how to shift the intervention activities for future workshops. First, I discuss my use of the knowledge base to outline problematic behaviors and root causes of those behaviors, focusing on key areas of alignment and dissonance with the observations in Chapter Four. I then turn to an analysis of the design principles used to construct the intervention before closing with an analysis of the feasibility of this intervention at Bay Vista High School.

Alignment to problem descriptions. During the design of the intervention in this study, I raised three salient issues from the literature and observations of teachers. First, the science teachers at Bay Vista lacked opportunities to engage in worthwhile professional development, evidenced by the unstructured nature of their collaboration time prior to this study and their accounts of opportunities for professional learning. Additionally, the participants frequently mentioned constraints on their time and the focus of the administration of the school, leading me to believe there were motivational factors preventing deeper learning by the teachers. However, following the schoolwide implementation of a workshop on Kagan strategies, performed by a provider outside the scope of this study, I realized the science teachers yearned for instructional strategies that could help improve their practice. The participants would likely benefit from high-quality examples of secondary science lessons involving inquiry practices and sense-making opportunities for students. A motivation to improve was present, but a motivation to invest time and energy into their own learning and persist through necessary stages of learning was absent. In the subsequent reflection on workshop activities, I explore potential intervention steps to address the need for a greater focus on example strategies as a path to deepening science instructional practices.

Choices of design principles. When consulting the research for strategies to shift the perceived persistent problems in teacher learning, several principles guided the intervention design. First, I attempted to create an intervention aligned to the most recent thinking on high quality professional learning experiences. While the design successfully engaged teachers in collegial learning, as evident by the presentations and reflective discussions in the second half of the intervention, I could have enacted a design with attention to other aspects of the research. First, the NASEM (2015) report suggests an experimentation and focus on high quality examples of lessons and pedagogical practices representing deeper science learning. Second, I ignored the notion of minimum required time, or engagement level, for successful professional learning due to feasibility concerns, discussed below, and limited consensus in the literature.

Third, I addressed issues of teacher motivation through goal setting, self-efficacy, experiencing successes, and collegial learning. I am confident with the influence of self-efficacy and collegial learning as the participants in this study began to shift their described challenges as student deficit to recognizing their control and ability to reach more students. These conversations were deepened by conversations with their colleagues. However, very limited discussion focused on the successes of students. Further practice with using student successes could have been helpful. However, focusing primarily on positive aspects of student learning represents a major shift in behavior that may not have been possible within this brief intervention.

Additionally, the design progressed without participants commanding a deep understanding of setting learning goals for themselves. Since this study did not have sufficient data to assess these two areas, I am unable to evaluate the principles as correct or not. The sections below discuss possible intervention alterations and ides for future iterations of the design.

Feasibility challenges. Although the implementation of the intervention workshops proceeded within an acceptable time frame, a few challenges to the design feasibility were not anticipated before commencing the design. Principally, the engagement of the participants was problematic. The success of professional development programs seems to correlate to the number of hours teachers spend participating, but a minimum threshold value has not been agreed upon (Desimone et. al., 2002). Alternatively, the authors of the NASEM (2015) report suggest that focused and committed participation in PD may substitute for a minimum number of hours. This design attempted to implement a short intervention of approximately 12 hours. I made this decision based on the relatively limited amount of time allocated to PD in most schools. However, the number of absences and missed opportunities for participation by the teachers in the study led me to question the feasibility of the design. Implementing an intervention consisting of more sessions of a slightly longer duration may have created an environment more conducive to teacher participation and learning by allowing them to fully explore their classroom experiments and develop the necessary skills for reflections and analysis of teaching and learning.

Intervention activity reflection. For the purposes of improving aspects of the designed intervention for future iterations and applications to other teacher learning experiences, I analyze some of the key components of the intervention workshops and suggest activities to keep, those to alter, and activities to remove from future iterations. Given the conclusion that teachers needed more examples of high quality science teaching, I recommend adjusting the opening engagement activity and vignettes to include more explicit examples of the types of lesson and activities that could form the basis for participant's learning plans. This could be further supported with an activity reviewing high quality science lessons, ideally through video analysis. and Additionally, explicit use of the Teacher Typology Matrix (Appendix A) to analyze videos and vignettes could provide participants with a necessary insight into the planned trajectory of learning and growth towards desired science pedagogy.

An additional point of divergence from the implemented plan concerns the use of the personal professional learning plans. Specifically, the participants exhibited a need for instruction around types of goals and evidence used in the plans. Two activities were added as design alterations in the middle of the intervention: the fishbowl discussion and the activity-evidence pairing discussion. These two activities proved instrumental in shifting the focus to teacher learning and using more informative artifacts of learning as evidence, respectively. This was seen in the quality of plans between the first and second feedback sessions. Participants may benefit from moving these activities to the beginning of the intervention. Finally, I would construct a more detailed peer feedback tool that would allow teachers to structure their components in a way that guided each participant to extend their learning for the second iteration. Many of the participants treated the two sessions as stand-alone learning opportunities.

That final workshop component was the peer feedback sessions. During workshops four, five and seven, participants provided feedback. In workshops four and five, this feedback occurred in a whole group setting. While interesting for everyone to hear their colleagues' presentations, the depth of discussion and feedback did not merit the amount of time used for the discussions. During workshop seven, groups had more effective conversations for two reasons. First, small groups were formed, enabling each participant to present their work during the workshop. Additionally, each member of the small group was working on a similar aspect of science teaching. Second, a facilitator's guide was constructed for peer facilitation of the small groups, making it more effective for them to stay focused on the artifacts of learning.

In summary, key changes to the intervention would be the replacement of the opening engagement activity with more detailed examples of high quality science teaching and learning; explicit instruction on the TTM, teacher-focused learning goals, and types of evidence to choose based on learning plan strategies; and a more structured venue for peer feedback to encourage deeper discussion of student learning based on attempted strategies. From a design perspective, this raises the question of time and commitment which will be explored when analyzing the design process below.

Research design. Design development research offers a unique opportunity to combine the expertise and experiences of practitioners with insights and research methodology from academia to investigate deep and persistent problems with educational systems. I believe the benefits of this methodology outweigh the challenges. However, two challenging aspects were encountered during the process of designing and implementing this study.

Contextual considerations. One surprising challenge was locating a study site for this research. While the aim to work with middle career teachers removed a few sites as potential locations, working one step removed from my position in the organization proved more challenging. I planned a design to intervene in teacher learning while working at the district level with teacher leaders. If I had organizational control over the teachers work, as a principal, it would have been easy to mandate implementation of this intervention. However, without any immediate supervisory responsibility to the participants, I was able to implement a design with a decreased amount of expectancy bias from participants who were not dealing with a supervisor.

The difficulty in this approach was finding a site willing to implement the design. After completing the needs assessment and first three workshops in one high school, the principal altered the schedule to implement school-wide professional learning during this time. It took several months to find another site that was willing to dedicate time to the professional learning in this study. This presents a general challenge with design development research, control of the time of the research participants. This must be balanced with the conflated roles of a lead in design development studies.

Role as researcher-facilitator-practitioner. During the implementation of this design, I served as the lead designer, researcher, facilitator of workshops, and practitioner with advanced skill and technical knowledge of the subject. With regard to my role as researcher, the key area of concern is with my biases during the collection of data. I mitigated this concern by using audio recordings to check bias in field notes and having a critical research friend at the university review my findings and check for bias towards positive outcomes in the study.

As mentioned in the previous section, acting as the designer and facilitator had limited influence on participants' responses, because I was not also serving as their supervisor, nor was I in the same supervisory chain as they were. This allowed me to keep their performance and ideas confidential from those above them in Bay Vista Schools. However, this study did not discuss and account for my expertise in the area of science teacher professional learning. Having led hundreds of hours of professional development for science teachers and administrators, my facilitation skills are well honed. Future studies should consider how to implement this study with less experienced facilitators.

Finally, the advantages of design development purported by Mintrop (2016) remained elusive during this design. Without a team of co-designers and coherence between the system learning and this specific intervention, I was not able to leverage the skills of partners and leaders in the school system. Working collaboratively with leaders in the school and district could allow for the creation of a coherent professional learning system that enables the type of deep learning needed by science teachers. A more powerful intervention would include planning and collaboration with partners across the educational system.

Implications for Practice and Research

The design development research performed during this study provides many insights into the specific context of science teacher professional learning and the potential power of design development studies. Given the professional learning literature is lacking deep analysis of teacher experience, this study provides a glimpse into the process. For example, the context for learning from managerial attitudes of administrators to the daily experience of the teachers alters the perspective and amount of engagement teachers bring into their learning experiences. While large-scale studies with generalizable tools are more frequently cited in the academy, studies, such as this one, illuminate the importance of catering the learning to the specific learner. In grounding the learning with specific learners, this study revealed that tools and foci of learning needed more explicit instruction to be powerful for the learners at Bay Vista High School. Some of the approaches highlighted in this study may be useful to any professional learning provider. However, the findings may prove most valuable to facilitators working with a diverse group of mid-career teachers in similar high-poverty contexts with many ELL students. Additionally, this study exhibited limited success within the established outcome goals indicating the importance of a high-quality needs assessment. This study would have benefited from information about specific teachers' learning needs. Despite these larger lessons, several aspects of this research methodology limit the scope of learning and leave several unanswered questions.

Limitations of the study. In situ research, such as design development or action research, contains inherent challenges to generalizability. At the largest grain size, this design development study was based on the experience of seven teachers. While this allows deeper inquiry and closer analysis of their experience, many contexts do not have the capacity to implement such intensive designs for so few participants. This calls the utility of design development research into question. DDS is established as a powerful methodology for deeper learning, but it may lack the necessary practical utility for wide use by school administrators, given the investment in background research, needs assessment, and implementation time.

In this design, I made an assumption about teacher learning that was ill-informed due to my limited ability to conduct an ongoing needs assessment and lack of institutional knowledge of the participating teachers. Given a more intimate knowledge of the teachers participating in this study, I could have prepared a tailored program addressing their specific areas of interest in science pedagogy. I assumed that teachers had experience with student-centered, cognitively-demanding, NGSS-aligned pedagogical techniques. While a few had some experience, the result was a mid-implementation shift to general goals about the use of daily student artifacts and teacher-centered locus of control for classroom lessons. While insights were gleaned from the research, an understanding of teacher preparedness in the area of advanced pedagogical content knowledge escaped the scope of this intervention.

Finally, the use of self-reported data for this study may have complicated the outcome. Specifically, in areas that showed limited or negative change, positive results may have occurred. As participants develop and learn during the intervention, they may experience a recognition of their own learning needs, thereby assessing their development at a lower level in the outcome interviews. Reliance on multiple sources of evidence should be continued in further studies.

Remaining questions. Due to the named limitations, several questions remain unexplored with a few salient issues to the original objectives of this research remaining obscured from my understanding. Most importantly, the importance of a high-quality needs assessment is now established, but I am uncertain if a better understanding of the participants' attitudes and behaviors could have shifted the learning to involve more science pedagogy explorations. While participants did become motivated to learn about their classrooms, I am left wondering if the teachers in this study could be motivated to attempt riskier classroom strategies that deepen science learning for students. Additionally, the participants made limited advances in the use of technical language to describe teaching and learning alongside limited student artifacts of learning. The question remains, would development of low inference observation and use of a wide range of student artifacts enable powerful teacher learning? These questions and others could be answered through research describe in the following section.

Further iterations. In subsequent implementations of similar interventions, I suggest attempting three key modifications to the design. First, participants may benefit from explicit instruction on descriptions of teaching and classroom artifacts. Activities such as the fishbowl discussion and the strategy-evidence matching discussion suggested success for additional explicative activities. For example, participants lacked a deep description of high quality teaching and learning. An activity using the terms and ideas found in the TTM (Appendix A) may have provided the technical language necessary for teachers to deepen their focus on NGSS-aligned pedagogy. Second, a list of several archetypal examples of high-quality science instruction may guide participants away from safer and more general engagement strategies towards deeper science pedagogical strategies. The addition of more discrete, science-specific pedagogical techniques may guide participants to the type of NGSS-aligned instruction purported in the design desired outcomes. While example strategies were listed on the example professional learning plans attached to the vignettes, participants lacked a thorough understanding of these strategies. Reviewing classroom video and student artifacts of learning from classrooms using those strategies would serve the dual purpose of instructing teachers on the implementation of the strategies and reinforcing the use of technical language to describe the desired teacher learning arc. Finally, sustainability is a major concern in this design. Subsequent iterations should attempt to further develop the design with teacher leaders and department chairs. If they have a thorough understanding of the intent of the design and contribute some specifics to the planning, a wider group of facilitators could be developed in the school district. These actors are closest to the teachers and could become trusted and knowledgeable facilitators. Questions of feasibility would need to be addressed to ensure district-wide support for a program involving teacher leaders.

Conclusions

Within this design development study, I attempted to motivate teachers to experiment with science pedagogical strategies representing a shift towards deeper science learning and student-centered pedagogy focusing on the combination of science practices, knowledge, and concepts. The aim of this study addressed concerns from the literature and my teacher observations which illustrated a need to motivate teachers away from the repetition of status quo behavior patterns into a state of learning and experimentation with reflection. During the seven-workshop series, teacher participants exhibited shifts in their perception of control towards teacher responsibility for the outcomes of lessons and their openness to investigate their classroom teaching through artifacts. However, the depth of teacher learning and experimentation remained at a more general level on engagement strategies, in lieu of advancing science-specific pedagogy. I believe a need for more explicit instruction on technical observation of teaching and learning alongside specific science strategies are needed in this small population of teachers.

References

Achinstein, B., & Ogawa, R. (2006). (In) Fidelity: What the resistance of new teachers reveals about professional principles and prescriptive educational policies. *Harvard educational review*, *76*(1), 30-63.

Amabile, T., & Kramer, S. (2011). *The progress principle: Using small wins to ignite joy, engagement, and creativity at work*. Harvard Business Press.

Ashton, P. T. (1985). Motivation and the teacher's sense of efficacy. *Research on motivation in education*, *2*, 141-174.

Ashton, P. T., & Webb, R. B. (1986). Teachers' sense of efficacy, classroom behavior, and student achievement. *Teachers' sense of efficacy and student achievement*, 125-144.

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, *36*(5), 258-267.

Ball, D. L. (2000). Bridging practices intertwining content and pedagogy in teaching and learning to teach. *Journal of teacher education*, *51*(3), 241-247.

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). Report of the 2012 National Survey of Science and Mathematics Education. *Horizon Research, Inc.(NJ1)*.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher*, *33*(8), 3-15.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American educational research journal*, *42*(2), 231-268.

Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of school psychology*, *44*(6), 473-490.

Coghlan, D., & Brannick, T. (2014). *Doing action research in your own organization*. Sage.

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social–emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology*, *104*(4), 1189.

Cousins, J., & Walker, C. (1995, June). Personal teacher efficacy as a predictor of teachers' attitudes toward applied educational research. In *annual meeting of the Canadian Association for the Study of Educational Administration, Montreal*.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009). Professional learning in the learning profession. *Washington, DC: National Staff Development Council*.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational researcher*, *38*(3), 181-199.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational evaluation and policy analysis*, *24*(2), 81-112.

Finnigan, K. S., & Gross, B. (2007). Do accountability policy sanctions influence teacher motivation? Lessons from Chicago's low-performing schools. *American Educational Research Journal*, *44*(3), 594-630.

Gardner, H., Csikszentmihalyi, M., & Damon, W. (2002). *Good work: When excellence and ethics meet*. Basic Books.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of educational psychology*, *76*(4), 569.

Goddard, R. D., Hoy, W. K., & Hoy, A. W. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal*, *37*(2), 479-507.

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and teacher education*, *4*(1), 63-69.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Ingersoll, R. M. (2003). Who controls teachers' work.

Kelley, C., & Protsik, J. (1997). Risk and reward: Perspectives on the implementation of Kentucky's school-based performance award program. *Educational Administration Quarterly*, *33*(4), 474-505.

Kruse, S. D., & Louis, K. S. (1993). An Emerging Framework for Analyzing School-Based Professional Community.

Lewis, E., Baker, D., Watts, N. B., & Lang, M. (2014). A professional learning community activity for science teachers: How to incorporate discourse-rich instructional strategies into science lessons. *Science Educator*, *23*(1), 27.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, *57*(9), 705.

Luthans, F., Youssef-Morgan, C. M., & Avolio, B. J. (2015). *Psychological capital and beyond*. Oxford University Press.

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In *Examining pedagogical content knowledge* (pp. 95-132). Springer Netherlands.

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach: An interactive approach*. Sage.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.

Mintrop, H. (2004). *Schools on probation: How accountability works (and doesn't work)*. Teachers College Press.

Mintrop, H. (2012). Bridging Accountability Obligations, Professional Values and (Perceived) Student Needs with Integrity. Journal of Educational Administration, 50 (5), 695 – 726. (Highly Commended Paper in the 2012 volume of Journal of Educational Administration).

Mintrop, H. (2016). Designing for Equity: How Educational Leaders Create Research-based Interventions. Harvard.

Mintrop, H., & Trujillo, T. (2007). The practical relevance of accountability systems for school improvement: A descriptive analysis of California schools. *Educational Evaluation and Policy Analysis*, *29*(4), 319-352.

National Academies of Sciences, Engineering, and Medicine. (2015). *Science Teachers Learning: Enhancing Opportunities, Creating Supportive Contexts*. Committee on Strengthening Science Education through a Teacher Learning Continuum. Board on Science Education and Teacher Advisory Council, Division of Behavioral and Social Science and Education. Washington, DC: The National Academies Press.

National Governors Association Center for Best Practices, Council of Chief State School. (2010). *Common Core State Standards*. Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers

NGSS Lead States. 2013. *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

Patton, M. Q. (2005). *Qualitative research*. John Wiley & Sons, Ltd.

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American educational research journal*, *44*(4), 921-958.

Reiser, B. J. (2013, September). What professional development strategies are needed for successful implementation of the Next Generation Science Standards. In *Paper written for the Invitational Research Symposium on Science Assessment* (Vol. 24, p. 25).

Santos, M., Darling-Hammond, L., & Cheuk, T. (2012). Teacher development to support English language learners in the context of common core state standards. In *Understanding Language Conference, Stanford University, California. http://ell.stanford.edu/sites/default/files/pdf/academic-papers/10-Santos% 20LDH% 20Teacher% 20Development% 20FINAL. pdf.*

Seijts, G. H., & Latham, G. P. (2001). The effect of distal learning, outcome, and proximal goals on a moderately complex task. *Journal of Organizational Behavior*, *22*(3), 291-307.

Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher education*, *20*(2), 163-183.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, *15*(2), 4-14.

Skrla, L., Scheurich, J. J., Johnson Jr, J. F., & Koschoreck, J. W. (2001). Complex and contested constructions of accountability and educational equity. *International Journal of Leadership in Education*, *4*(3), 277-283.

Stone, D. N., Deci, E. L., & Ryan, R. M. (2009). Beyond talk: Creating autonomous motivation through self-determination theory. *Journal of General Management*, *34*(3), 75.

Supovitz, J. A., & Turner, H. M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of research in science teaching*, *37*(9), 963-980.

Thoonen, E. E., Sleegers, P. J., Oort, F. J., Peetsma, T. T., & Geijsel, F. P. (2011). How to improve teaching practices: The role of teacher motivation, organizational factors, and leadership practices. *Educational administration quarterly*, *47*(3), 496-536.

Tienken, C. H., & Zhao, Y. (2013). How common standards and standardized testing widen the opportunity gap. *Closing the opportunity gap. What American must do to give every child an even chance*, 111-122.

Tierney, P., & Farmer, S. M. (2011). Creative self-efficacy development and creative performance over time. *Journal of Applied Psychology*, *96*(2), 277.

Tschannen-Moran, M., Hoy, A. W., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of educational research*, *68*(2), 202-248.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and teacher education*, *17*(7), 783-805.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. Routledge.

Van Driel, J. H., Meirink, J. A., Van Veen, K., & Zwart, R. C. (2012). Current trends and missing links in studies on teacher professional development in science education: a review of design features and quality of research. *Studies in Science Education*, *48*(2), 129-160.

Van Veen, K. (2008). Analyzing teachers' working conditions from the perspective of teachers as professionals. *The profession of teacher in Dutch educational praxis. Sense Publishers, Rotterdam*.

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and teacher education*, *24*(1), 80-91.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

Webb, R. B., & Ashton, P. T. (1986). Teacher motivation and the conditions of teaching: A call for ecological reform. *Journal of Thought*, 43-60.

Woolfolk, A. E., Rosoff, B., & Hoy, W. K. (1990). Teachers' sense of efficacy and their beliefs about managing students. *Teaching and teacher Education*, *6*(2), 137-148.

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of research in education*, *24*, 173-209.

Wilson, S., Scweingruber, H., & Nielsen, N. (Eds.). (2015) *Science Teachers' Learning: Enhancing Opportunities, Creating Supportive Contexts.* The National Academies Press.

Appendix A: Teacher Typology Matrix (TTM)

The following matrix is used during the coding of observations, teacher self-evaluations, and video review activities to identify the characteristics of teaching style and lessons.

DOMAIN	INITIAL	EMERGING	DEVELOPING
Teaching Style	Teacher-centered, didactic, students receive knowledge and answer questions to demonstrate that knowledge, focuses on the Disciplinary Core Ideas of NGSS. Knowledge transmission: lectures, videos, forced response quizzes, teacher does most of the talking,	Some opportunities for students to ask questions and demonstrate learning during the lesson. Some science and engineering practices used. Activities and experiences only offer reinforcement of the knowledge being transmitted.	Student-centered, constructivist, integrates the crosscutting concepts, disciplinary core ideas, and science and engineering practices. Activity and experience based: investigations, hands-on activities, small group discussions, student talk; activities are open-ended; multiple responses are acceptable
Activity coherence	Activities are disjointed; There is not a clear progression of learning; some activities introduce new material.	Activities are loosely related but not presented in a way that facilitates students building of understanding.	Activities build from an engaging introduction and relate to each other building towards a preplanned conclusion; each activity contributes to student understanding of the topic
Relevance to students	Lessons focus on concepts and topics relevant only to scientists. Topics presented in a matter of fact way without context for students.	Lessons touch on interesting topics, but those topics are not the focus of the lesson.	Lessons framed in ways that captivate attention of students by highlighting connections to their lives, other studies, or futures.
Responsivenes s to students	Teacher does not acknowledge student questions or signs of interest in activities. Checks for understanding are missing from the lesson.	Teacher occasionally responds to student verbal and nonverbal cues. Response is inconsistent.	Teacher routinely responds to signs of interest from students, including questions. Teacher is aware of students' levels of understanding through a variety of checks for understanding.

Appendix B: Vignettes and Learning Plans

Vignettes. Descriptions of teachers' lessons introducing new science topics. Along with examples of pedagogy, an addition to vignettes will be a personal learning plan aligned to the desired learning trajectory constructed with regard to the theoretical framework of the project. Teacher typologies for vignettes:

Didactic lecture. lessons revolve around transmission of facts. Some examples are used. Student involvement is limited to recapitulating main ideas and discrete facts. Lessons have an obvious flow.

INTRODUCING DNA

Mr. Stevens displays a diagram of a DNA model with his projector. He asks the students, "What is this?" Several hands are raised, and he calls on one student who responds, "DNA." Acknowledging the answer, Mr. Stevens begins to describe the structure of DNA, highlighting vocabulary words such as nucleotide and double helix. Most of the students are taking notes as he talks. He continues to discuss the way the base-pairing rule was discovered. At one point he pauses to allow students to sketch a segment of DNA illustrating the base-pairing rule.

One student asks, "I notice the double rings always match with a single ring."

"That's right," responds Mr. Stevens. He continues to explain the differences between purine and pyrimidine bases. He pauses to ask the class which bases are purine. After the class responds, Mr. Stevens finishes explaining the hydrogen bonds between bases. He closes class with an exit ticket quiz covering the key vocabulary from the lesson.

Disjointed implementer. lessons are characterized by a string of activities, some teacher-centered, others hands-on. Students have difficulty drawing connections between different activities. Teacher perceives their teaching as high quality due to the inclusion of many "best practices."

INTRODUCING THE ORIGIN OF THE ELEMENTS

Ms. Lewis begins class by having students write down where they think carbon comes from. After a brief period, she turns on a video that provides a description of a supernova and the various products of fusion. Students spend the next 20 minutes on an interactive activity investigating the characteristics of elements on the periodic table. By using various models of atoms, they make conclusions about the arrangement of the elements on the periodic table, atomic mass, and atomic number. The day concludes with students breaking into jigsaw groups to read about and present the life cycles of stars of various sizes.

Front loader. Classes eventually have students engaging in activities with a purpose, but facts and instructions take the majority of the time. Teacher transmit facts through lectures and procedural instruction or textbook readings.

INTRODUCING NEWTON"S SECOND LAW

Mr. Fredericks is teaching a class on Newton's second law. He begins with a video of a rollercoaster. He pauses the video every few seconds to describe the motion of the rollercoaster car say things like, "See how much farther the car travels in one second the farther it is down the slope," or "Watch it slow down as it goes up the hill."

He continues the class with a brief lecture on the definition of Newton's second law and a description of the equation relating force to mass and acceleration. He works through several sample calculations on the board as students copy notes into their notebooks. For the last ten minutes of class, Mr. Fredericks distributes a data table showing the change in distance and velocity of a skydiver before they open their parachute and immediately after they open it. Students are asked to make a graph and explain how Newton's second law is seen in the example.

Open inquirer. Students are free to investigate or create in class. Little structure is given. Students discover scientific principles by investigation.

INTRODUCING HOMEOSTASIS

Mr. Howard begins class with a list of experiment ideas on the board, including effects of exercise, length of time holding breath, changes in blood pressure, effects of warm and cold temperatures. Students find groups that want to work on similar projects. Mr. Howard directs the students to a section in their textbooks that explains homeostasis and feedback in organisms.

He instructs the students to design an experiment on homeostasis that involves their human body phenomenon.

Some students read the textbook, while others being researching the phenomenon on the computers in the back of class. Another group begins charting their experimental designs on a dry erase board.

Half way through the class period, Mr. Howard passes out a sheet of requirements for the project. The full lab report requires students to track multiple variables, visually represent their data, build a model of homeostasis as it relates to the phenomenon they are studying, and write a conclusion connecting their work to homeostasis. Students ask several questions regarding the requirements, such as “How long does the report have to be?” Mr. Howard answers most of the questions and allows students to return to work time.

Near the end of class, Mr. Howard brings out a poster display of a previous student’s project on the effects of various beverages on urine pH. He highlights the aspects he appreciates about the example project. As he dismisses class, he reminds them that they should begin conducting their experiments tomorrow and complete projects in four days.

NGSS-aligned constructivist. Lessons are guided inquiry or phenomena first. Students engage in relevant activities and topics to pique their interest in the subject. Ample time is allotted for student questioning and sense-making.

INTRODUCING BOND ENERGY AND COMBUSTION REACTIONS

At the beginning of class, students are shown a demonstration where a small amount of alcohol is burned in a water bottle making a jet-like flame. The teacher, Ms. Edwards, prompts students to write down their observations about how the fuel burns and what are the likely reactants and products based on what they can observe. She then poses the question, which type of fuel is the most efficient: methane, propane, or coal.

To answer the question, students are given a list of formulas and diagrams representing the chemicals and a brief passage about the use of the three fuels in cooking. Students are required to support their claim about the fuels with evidence.

Ms. Edwards also provides alcohol lamps with various alcohols in them: methanol, ethanol, propanol, and butanol. Before beginning an experiment to test the energy of the alcohols, students build models of the molecular structures with gummy bears and toothpicks. Students make predictions about the energy in the chemicals and set up a calorimetry experiment to calculate the energy per mole of fuel. After students discuss their experiments with another group, individual students are asked to compare their findings to the original question about coal, methane, and propane.

Personal Learning Plan. Plans will include desired shifts in teaching, learning goals, classroom implementation ideas/inspirations, intended artifact collection showing evidence of progress, and a note catcher to record changes to plans that happen in the moment of teaching.

Name:	Course/Grade:
Problem or Area of Improvement:	Learning Goal:
Implementation Ideas / Resources / Inspirations for Change:	Strategy Attempted:
Evidence to be Collected:	Implementation Notes:
Signs of Success for Students:	Next Steps based on Evidence Analysis:

Example PLPs for Vignettes

Name: Mr. Stevens	Course/Grade: Biology/9
Problem or Area of Improvement: Not all students are engaged and raising hands to answer questions. Exit tickets only have a few questions and many students answer incorrectly.	Learning Goal: Determine types of questions that work for a class discussion strategy.
Implementation Ideas / Resources / Inspirations for Change: A neighboring teacher uses multiple choice response cards. The district science team discussed how to use think-pair-share in class. A recent PD covered the need to use questions that engage students' prior knowledge.	Planned Strategy: During the next lesson, I will build in a think pair share. I will add a question that is relevant to the students' background knowledge and a question that tests their scientific understanding to see if the strategy works for both.
Evidence to be Collected: Video or audio record one group's conversation. Walk around the room and record ideas heard from students. Gauge overall engagement, noting differences between the two types of questions.	Implementation Notes: Think about other opportunities to try the discussion strategy.
Signs of Success for Students	Next Steps based on Evidence Analysis

Name: Ms. Lewis	Course/Grade: Chemistry/10 and 11
Problem or Area of Improvement: Students seem confused between activities, possibly due to the amount of material trying to be covered.	Learning Goal: Figure out how to connect various ideas with an activity.
Implementation Ideas / Resources / Inspirations for Change: A recent PD presented a Revise, Refine, and Reflect form that has students keep referring back to one essential question after each portion of the activity.	Strategy Attempted: Determine an essential question and how a series of activities will relate to that question.
Evidence to be Collected: Checking how students document classroom experiences to	Implementation Notes:
Signs of Success for Students:	Next Steps based on Evidence Analysis:

Name: Mr. Fredericks	Course/Grade: Physics 10/11
Problem or Area of Improvement: Wondering if students are really getting the best use out of the activities. They don't seem to contribute much to their increased understanding.	Learning Goal: Determine the best types of activities to get students engaged in a topic before explaining the specific scientific. Also, learn from the student discussions during the activity to modify what parts of the explanation I need to emphasize and use some of their examples.
Implementation Ideas / Resources / Inspirations for Change: While reading a teacher's blog post, I came across the idea ABC, CBV. Activity before content and content before vocabulary.	Strategy Attempted: Upcoming Introductory Activities to try: - Students are given a 5-pound weight and 20-pound weight and asked to describe the differences between lifting each to the table 10 times in two trials, one fast lifting and one slow lifting to explore work and power. Students are asked to make two models of two cars of different masses colliding on a track, one where cars stick together and one where they bounce apart.
Evidence to be Collected: Compare video clips for student participation and discussion as well as exit tickets checking for understanding to see which type of activity works better for the majority of students.	Implementation Notes:
Signs of Success for Students	Next Steps based on Evidence Analysis

Name: Mr. Howard	Course/Grade: Physiology 11/12
Problem or Area of Improvement: Students work on their projects but can't seem to explain the scientific concepts/phenomena in a novel context.	Learning Goal: Determine two structures that help students learn from lab experiences to create learning about key science concepts.
Implementation Ideas / Resources / Inspirations for Change: Peer Review of each group's work, such as Expert Carousel (Sampson, 2014); Focus on a list of 3-5 claims that students focus on when collecting evidence to support or refute the claims.	Strategy Attempted: Using a guiding question, have students collect evidence to support and refute a list of claims. After students construct their claims with evidence from their experiments, they go through a peer review process where one member explains the group's thinking to another group.
Evidence to be Collected: Photos of student posters before and after revisions from peer feedback. Responses to open ended formative assessment. Documentation of group participation.	Implementation Notes
Signs of Success for Students	Next Steps based on Evidence Analysis

Name: Ms. Edwards	Course/Grade: Chemistry/11
Problem or Area of Improvement: Some students seem to be letting others do most of the work and talking during the lessons. It is unclear if they are just not participating or are not learning the material.	Learning Goal: Determine the best way to check for student understanding during the lessons.
Implementation Ideas / Resources / Inspirations for Change: Some colleagues suggested a few strategies: rotating to groups and asking questions of the students who seem to be avoiding participation, build in individual formative assessments throughout the lesson, assign certain roles (facilitator) to the reluctant students.	Strategy Attempted: Verbal checks with students of concern and individual assessment at the end of lessons to see if
Evidence to be Collected: Audio recording and notes on group responses to my questions and scores on the end of lesson paragraphs to determine if reluctant students are grasping key ideas.	Implementation Notes
Signs of Success for Students	Next Steps based on Evidence Analysis

Appendix C: Pre/Post Interview Questions

This is the primary source of impact data evaluating the five areas of potential growth.

On a scale of 1-5…	Example probing questions
Analyzing teaching	
To what extent are you satisfied with your teaching? To what extent do you accurately assess your strengths and areas for growth in your classroom teaching? To what extent would you classify yourself as an effective teacher?	How did you choose that rating? What do you think outside observers would see? How do you know those are strengths? What makes you effective or not? What characterizes an average day for you, in terms of effectiveness?
Ability to set next level learning	
To what extent do you set goals regarding your teaching? To what extent do you achieve your goals? If you don't set goals, to what extent does your teaching improve?	What type of goals do you set? Could you give an example? How do you choose goals? What do you think determines if you do or don't meet your goals/improve your teaching? Do you find goal setting motivating?
Sense of confidence, motivation, efficacy	
To what extent do you see improvements in your teaching practice as time goes on? To what extent do improvements in your teaching motivate or inspire you to keep improving? To what extent are you comfortable with learning and trying new things in the classroom? To what extent is your improvement or motivation related to your students' successes? To what extent do you believe that you can effectively teach all students in your classes?	How do you know you are improving? Can you provide an example? What would help you improve more? Do you find that you get into ruts and periods of consistent improvement? What determines those periods? Tell me about your rating. How do you measure success for your students? If not, how do you gauge your improvement? Will you give a specific example of that? What other factors motivate you? Can you give some examples of when your lessons don't seem to be working?
Sense of Learning Quality	
To what extent have you intentionally changed your practice in the last three months? To what extent do you try new strategies in your classroom?	Looking back at the last three months, how do you assess the kinds of progress you have made on the way teach science? Can you tell me what you do at the end of the day or lesson to reflect on your teaching of lessons? Can you give an example of something you have intentionally changed?
Sense of Learning Safety	
To what extent is there an open and honest relationship among science teachers, based on trust? To what extent are you comfortable sharing your ideas and experiences with regard to teaching with other science teachers? To what extent do you share student artifacts (work samples, videos, observations) with other science teachers?	Can you give an example to illustrate those feelings? What specifically have you shared with your colleagues? What types of strategies and experiences have you shared? Are there examples of sharing what students are doing this year? If a high rating, can you talk about the conversations that happened? If low rating, do you think it would be helpful to look at student artifacts? Tell me what you think could come out of those conversations. If not, what do you think would be a good way to really understand the challenges other teachers have in their teaching.

Appendix D: Impact Metrics Rubric

A list of descriptors to align participants' self-evaluation to descriptors and the information coded from probing questions and the performance assessment.

	INITIAL	EMERGING	DEVELOPED
Analyzing Teaching	Does not describe teaching in terms of the teacher typology (TTM) (Appendix A). Refers to completion and compliance issues more than issues of understanding. Places blame on students for teaching challenges.	Identifies challenges but externalizes blame for those challenges. Occasionally refers to student experience but focuses on compliance and general engagement rather than deeper learning. Describes successes from a teacher perspective, rating the quality of the activity, without considering the impact on students.	Accurately and consistently uses TTM (or similar language) domains to describe teaching. Identifies challenges and successes based on student experience and artifacts. Locates the challenges within the teacher's control.
Ability to set next level learning	Participants rarely try new strategies in the classroom. Participants do not perceive an ability to affect outcomes for students.	Attempt new strategies but do not persist to improve those strategies or use them in the future. Takes some responsibility for student improvement and engagement	Try new, risky and innovative strategies frequently. List concrete evidence of student learning and improvement. A great sense of ability to change student outcomes.
Sense of confidence, motivation, and efficacy	Participants rarely try new strategies in the classroom. Participants do not perceive an ability to affect outcomes for students.	Attempt new strategies but do not persist to improve those strategies or use them in the future. Takes some responsibility for student improvement and engagement	Try new, risky and innovative strategies frequently. List concrete evidence of student learning and improvement. A great sense of ability to change student outcomes.
Sense of Learning Quality	Learning is not related to classroom strategies. Rarely reflects on student successes. Reflection is about how they felt about the lesson taught. Little to no reflection given.	Desire to learn about strategies to improve teaching loosely related to student academic performance Sometimes reflects on implementation successes. Sometimes builds on activities that seem successful.	Choose pedagogical strategies to test in their classrooms and reflect upon. Consistently reflect on daily lesson implementation successes. Frequently try new strategies in response to observations of student need and challenges.
Sense of Learning Safety	Sharing is often in one direction only and only of activities that have a low risk associated (e.g. activities from another source or activities that are associated with success) Advice not solicited. Interactions mostly social and supportive. Trust lacking or broken.	Shares some ideas and how those may be implemented in the classroom, but rarely shares artifacts of student learning. Advice is solicited for safe topics, such as student behavior, materials management, lesson ideas. Some interactions focused on pedagogy and other instructional issues.	Share more sensitive and revealing material about their teaching, such as student work and videos. Participants are receptive to constructive feedback from colleagues. Participants describe benefits of sharing student artifacts. Interactions push each other to think about ways to improve instruction.

Appendix E: Strategy-Evidence Matching Activity

1. Choose one practice or strategy from the first column that seems like something you would like to try in your classroom.
2. Chose an evidence source from the second column that will allow you to determine the effectiveness of the practice or strategy.
3. Explain why you think this is the best evidence to collect for the strategy.
4. Give an example of what you might see or hear in the evidence that will be a sign of success for your students.

PRACTICE or STRATEGY	EVIDENCE
Small group (or pair) student discussion using a Kagan strategy to engage in conversation about a scientific idea.	Samples of student writing from students of varying ability and experience
Phenomena first to engage students - For example, showing a video of something thought provoking related to the topic and allowing students to ask questions.	Video clip of teacher questioning
Whole class summary/sense-making discussion using guiding questions and/or a protocol	Audio recording of students talking
Providing students feedback on projects, lab reports, or other explanations to correct misconceptions and allowing them to revise their work.	Photos of student group work products
Refining ideas - have students explain how their ideas about a topic or phenomena change as they experience various activities in class.	Video of the whole class completing a task
Peer feedback - using a protocol, allow students to learn from and give feedback to each other on a project or problem solution	Student self-assessment information
	Short quiz at the end of class
	Recordings of teacher moves and questions (video, audio, or notes)

Appendix F: Peer Feedback on Classroom Evidence

PURPOSE: To review evidence and have colleagues identify insights and areas of growth and next steps.

ROLES:
Presenter - presents evidence and receives feedback
Moderator - keeps track of time to ensure all members have a chance and redirects conversation to ensure all voices are heard
Participants - listen to the presentation of evidence and bring their ideas and perspectives into the conversation.

PROTOCOL:
Presentation (5 min) - the presenter provides an overview of the evidence, showing key artifacts, signs of student success, areas of concern, and main takeaways from the experience. Focus on direct observations from the evidence. Spend little to no time explaining the activity. You may opt to have participants observe your artifact for the whole time.

Signs of Positive Student Experience (3-4 min) - the whole group engages in a discussion citing specific pieces of evidence that seem to indicate success with students. For example: During the video artifact, I heard students using their own language to explain photosynthesis to their peers. OR In the work samples, I saw that ⅚ of students made improvements to increase the science content of their models, such as adding arrows and labelling those with processes.

Suggestions (3-4 min) - the whole group discusses areas of growth or potential learning directions for the presenter. At this time, you may also respond to concerns named by the presenter at the beginning of the protocol. For example: I wonder if you would have more students participating in the class discussion fi you gave them a chance to write their ideas down and discuss them with a neighbor before the whole group discussion.