UNIVERSITY OF CALIFORNIA, SAN DIEGO

CALIFORNIA STATE UNIVERSITY, SAN MARCOS

The Impact of Improvement Science Professional Development on Teacher Agency

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Education

in

Educational Leadership

by

Daisy Sharrock

Committee in charge:

California State University, San Marcos

Professor Sinem Siyahhan, Chair
Professor Brooke Soles
Professor Pat Stall

University of California, San Diego

Professor Alan Daly

2018

The Dissertation of Daisy Sharrock is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

Chair

University of California, San Diego

California State University, San Marcos

2018

DEDICATION

For lifelong learners everywhere.

Especially Karl and Kyna, two of my favorites.

EPIGRAPH

You cannot buy the revolution.

You cannot make the revolution.

You can only be the revolution.

It is in your spirit, or it is nowhere.

~ Ursula K. Le Guin

TABLE OF CONTENTS

Signature Page……...iii

Dedication……...iv

Epigraph…….....v

Table of Contents…….....vi

List of Tables……..……....xi

Acknowledgements……xii

Vita……..…….xiii

Abstract of the Dissertation ……..xv

CHAPTER 1: INTRODUCTION……....1

1.1 Statement of the Problem……...1

1.2 The Need for a Different Type of Reform…….2

1.3 Overview of the Study……...3

1.4 Significance of the Study……..……4

1.5 Limitations of the Study…….4

1.6 Overview of Dissertation……...5

CHAPTER 2: LITERATURE REVIEW…….6

2.1 Introduction……6

2.2 Different Conceptualizations of Teacher Agency……..6

2.3 Three Waves of Educational Reform and their Relation to Different Conceptualizations of Teacher Agency……...9

2.3.1 First Wave: The Progressive Era (1890 - 1930)……...10

2.3.2 Second Wave: The Civil Rights Movement (1960 - 1970)……..12

2.3.3 Third Wave: Standards Based Reform (1980 - 2015)……..14

2.3.3.1 Emergence of the common core……16

2.4 The Three Missing Components in Current Reform Efforts……...17

2.4.1 The Need to Foster Teacher Inquiry…….18

2.4.2 The Need to Promote Collaboration…….19

2.4.3 The Need to Develop a Professional Knowledge Base……....20

2.5 Improvement Science……...20

2.5.1 Principle 1. Make the Work Problem-Specific and User-Centered……..22

2.5.2 Principle 2. Focus on Variation in Performance……...23

2.5.3 Principle 3. See the System that Produces the Current Outcomes……24

2.5.3.1 Empathy interviews……...25

2.5.3.2 Chain of 'why' questions……...25

2.5.3.3 Fishbone diagrams…….26

2.5.3.4 Process maps……..26

2.5.3.4 Developing a working theory of improvement……..26

2.5.3.5 Driver diagrams…….27

2.5.4 Principle 4. We Cannot Improve at Scale What We Cannot Measure…….27

2.5.5 Principle 5. Use Disciplined Inquiry to Drive Improvement……28

2.5.5.1 Plan Do Study Act cycles……..29

2.5.5.2 Moving towards scale……..……..29

2.5.6 Principle 6. Accelerate Learning Through Networked Communities……..30

2. 6 Research Questions……30

CHAPTER 3: METHODOLOGY…….32

3.1 Introduction……..32

3.2 Context of the Study……32

3.3 Participants……...33

3.4 Improvement Science Professional Development……...33

3.5 Instruments and Data Collection……..36

3.5.1 Pre- and Post-intervention Teacher Surveys…….36

3.5.2 Interviews……..37

3.5.3 Improvement Science Documentation……..……38

3.5.4 Classroom observational data……...38

3.6 Data Analysis……...38

3.7 Role of the Researcher…….39

CHAPTER 4: RESULTS……...41

4.1 Introduction…….…41

4.2 The Impact of Improvement Science Professional Development on Teacher Agency……...…41

4.3 Teachers' Engagement with Improvement Science Tools and Methods……..…..45

4.4 How Teachers Use Improvement Tools and Methods to Define the Problem……47

4.4.1 Empathy Interviews……..48

4.4.2 Fishbone Diagrams……...50

4.4.3 Interrelationship Digraphs……52

4.4.4 Driver Diagrams……54

4.5 How Teachers Use Data to Inform Their Practice……...57

4.5.1 Outcome Measure Data: SBAC and Youth Truth……57

4.5.2 Data for Improvement……..……..59

4.5.2.1 MAP data……...59

4.5.2.2 PDSA level data…….61

4.5.2.2.1 Trying change ideas and choosing what to measure…...62

4.5.2.2.2 Collecting and analyzing data for iteration…….65

4.6 Structures that Support Improvement Science Implementation……..70

4.6.1 Protocols……...70

4.6.2 Improvement Teams…….…71

4.6.3 Regular Meeting Times…….72

4.6.4 Internal and External Capacity and Support…….73

4.6.5 Administrator Support……..73

4.6.6 School Improvement Coach Program……...73

4.6.7 Graduate School of Education School Resident Program……....74

4.6.8 Induction Program……74

4.7 Changes in Teacher Agency Over Time…….…74

4.7.1 Intentionality, Inquiry and Changes to Individual Practice….....75

4.7.2 Sharing and Spreading Change Ideas……...77

4.7.3 Identification as an Improver in an Improvement Community……...….78

4.8 Summary……..80

CHAPTER 5: DISCUSSION…….81

5.1 Overview of the Study…….81

5.2 Discussion of Findings…….82

5.2.1 Structures that Promote Teacher Inquiry……..82

5.2.2 Challenges Teachers Experienced Engaging in Inquiry……...84

5.2.3 Structures that Promote Collaboration……..87

5.2.4 Challenges to Collaboration……..88

5.2.5 Generating a Shared Knowledge Base……..88

5.2.6 Challenges to Generating a Shared Knowledge Base……...90

5.3 Implications for Improvement Science as a Professional Development Structure…..91

5.3.1 Supporting Improvement Science Professional Development…….....91

5.3.2 Challenges to Engaging in Improvement Science……93

5.4 Limitations of the Study…….…….94

5.5 Implications for Future Research……....95

5.6 Conclusion……..97

Works Cited……..99

Appendix A: Design of the Improvement Science Professional Development Intervention.....105

Appendix B: Improvement Science Tool Templates and Examples From EL Improvement Group..……106

Appendix C: Fishbone Diagram Protocol……110

Appendix D: Interrelationship Digraph Protocol…….111

Appendix E: Driver Diagram Protocol……113

Appendix F: Pre- and Post-intervention Survey……..116

Appendix G: Interview Protocol for Focus Teachers……..119

Appendix H: Observation Protocol for Focus Teachers……..……121

Appendix I: Digging into Data Protocol……..122

Appendix J: Word Wall Iterations…….……..124

LIST OF TABLES

Table 1. Improvement Science Professional Development Workshop Activities……..35

Table 2. Improvement Group Aim Statements……..36

Table 3. Percentage of Teachers Responding Agree and Strongly Agree on the Pre-and Post-intervention Survey Measures……...42

Table 4. Emerging Themes from the Analysis of Qualitative Data from Focus Teaches…….46

Table 5. Focus Teacher Demographics……..47

Table 6. Improvement Group Problem Statements and Fishbone Categories……...50

Table 7. Improvement Group Aim Statements and Primary and Secondary Drivers……55

Table 8. Improvement Group Plan Do Study Act Cycles……..62

Table 9. Improvement Group Plan Do Study Act Data Types……..62

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of my family, especially my husband Karl, who deserves his own doctorate in education, and my daughter Kyna, who generously gave her blessing for me to begin this journey three years ago. I'd also like to thank my parents, Steve, Alison, and Sally, for their optimism and love, and for continually encouraging me to be curious and resourceful.

I am also grateful for my colleagues at the Center for Research on Equity and Innovation – Stacey, Ryan, Ben D., Ben S., Mari, and Michelle – who breath life into improvement work every day through their passion for providing meaningful learning experiences for teachers and students. Most of the processes and protocols used in this intervention are the result of collaboration and iteration with this amazing team.

I would also like to acknowledge the members of my 'Heavenly Elevenly' cohort – Meg, Courtney, Jackie, Theresa, Pegah, Anna, Christina, Holly, Brooke, Scott, Adriana, Luke, and Leila – for their perspective, encouragement, and candor.

Special thanks to the team at the Carnegie Foundation for the Advancement of Teaching for integrating the ideas of quality improvement and the power of networks as a framework for educational reform.

Finally, I would like to thank my chair, Sinem, and my dissertation committee members, Alan and Brooke, for their guidance and support throughout this study. Their insight and experience has been invaluable.

VITA

1996 - 1999 Bachelor of Science, University of British Columbia

2009 - 2015 Chemistry Teacher, High Tech High

2009 - 2011 California Teaching Credential (Chemistry), High Tech High

2011 - 2013 Clear California Teaching Credential (Chemistry), High Tech High

2011 - 2013 Master in Education, High Tech High Graduate School of Education

2015 - 2018 Education Research Specialist, High Tech High

2015 - 2018 Doctor of Education, University of California, San Diego

PUBLICATIONS

Sharrock, D. & Rubenstein, R., (in press June 2018) Rehumanizing the mathematics classroom for student engagement. In Fredricks, J. (Ed.) *Handbook of Student Engagement Interventions: Working With Disengaged Youth.* Elsevier Academic Press

Sharrock, D., (March, 2018) *The Impact of Improvement Science Professional Development on Teacher Agency* (doctoral thesis). University of San Diego, California and California State University, San Marcos, San Diego, USA

Sharrock, D., (2017, June) Forging connections: supporting authentic learning in Sonoma County. In Hamilton S., & Scherer, R., (Eds.), *Inspiration, not Replication: How Teachers are Leading School Change From the Inside (*pp. 64-101). Available from Amazon.com

Sharrock, D., (2017, June) How can we abolish the Phrase *"I'm not a math person"*? *Edweek Deeper Learning Blog.* Retrieved from: http://blogs.edweek.org/edweek/learning_deeply/2017/06/how_can_we_abolish_the_phrase_im_just_not_a_math_person.html

Sharrock, D., (2015). Tackling the ambiguities in action research: Advice for the goal oriented practitioner on how to stay sane. In H. Lattimer and S. Caillier, (Eds.)*, Surviving and Thriving with Teacher Action Research,* (pp.155-160). New York, NY: Peter Lang Publishing Inc.

Sharrock, D. & Perry, L., (2014, June) A Reel-y authentic project. *Unboxed: A Journal of Adult Learning, issue 11.*

Sharrock, D., (2013, June) Down the rabbit hole: an exploration of student grit, motivation, and scientific inquiry (masters thesis). High Tech High Graduate School of Education, San Diego, USA. Available from Amazon.com.

Sharrock, D. & Jana, P., (2012, June) Unraveling the knot: critical thinking in presentations of learning. *Unboxed: A Journal of Adult Learning, issue 9.*

ABSTRACT OF THE DISSERTATION

The Impact of Improvement Science Professional Development on Teacher Agency

by

Daisy Sharrock

Doctor of Education in Educational Leadership

University of California, San Diego, 2018

California State University, San Marcos, 2018

Professor Sinem Siyahhan, Chair

The educational field is riddled with ineffective top-down reform initiatives that have failed to address disparities in student learning outcomes for disadvantaged students. Over the past century, three successive waves of educational reform have situated decision-making power at the state and federal level, far away from the classroom where teaching and learning occurs. Reform efforts are often poorly funded and under resourced, leaving teachers frustrated and disengaged. There is a need for bottom-up educational reform that leverages teacher inquiry, promotes teacher collaboration, and supports teachers in building a robust pedagogical knowledge base. Improvement science, with its focus on small, iterative, inquiry cycles, networked learning, and evidence-based decision-making, is a promising bottom-up reform strategy that addresses these needs. This study utilized a mixed method approach to understand

how teachers engaged with the tools and methods of improvement science by exploring: (1) How do teachers use improvement science tools and methods? (2) What structures support teachers in engaging in improvement work? And, (3) In what ways do improvement science tools and methods impact teacher agency?

In this study teacher agency was conceptualized as teachers' sense of belonging to an improvement community – *I am a member of a community that believes it can learn and improve,* teachers' confidence that they can meet the learning needs of their students – *I believe I have the skills and tools I need to meet the learning needs of my students*, and teachers' perceived value of professional development support, – *the improvement science tools and methods I'm learning are valuable to me now and in the future*. The study found that engaging in improvement science professional development increased teachers' sense of agency along all three dimensions. Teachers reported an increased sense of belonging to an improvement community, an increased belief that they could meet the learning needs of all their students, and that learning about improvement science tools and methods was useful to them now and in the future.

The study also determined four key structures that supported teachers as they engaged in learning and using the tools of improvement science, and two challenges that bear further investigation. Administrator support, regular meeting times, enlisting teachers as co-facilitators, and protocols to scaffold using improvement science tools, all contributed to teachers developing an improvement culture at their school site. Challenges teachers experienced during the study included deciding what data to collect to determine if a change idea was leading to improvement and using data to inform iterative cycles of inquiry during Plan Do Study Act cycles. The findings of this study suggest that the tools and methods of improvement science learned through

professional development have a positive impact on teachers' sense of agency and the development of a school improvement culture. Considerations for adopting an improvement science professional development framework are also discussed.

CHAPTER 1: INTRODUCTION

Statement of the Problem

The educational field is riddled with top-down reform initiatives that have failed to address inequitable student learning outcomes for students from low socioeconomic backgrounds and students of color (NAEP, 2015; Lee & Reeves, 2012; Mathis, 2010). Over the past century, there has been a steady consolidation of decision-making power at the district, state, and federal levels away from the classroom, the context in which teaching and learning occurs (Darling-Hammond, 1994; Berube, 1994). These remote education reform decisions are often constructed as a 'one size fits all' solution and fail to take into account the significant variability in what counts as effective teaching and learning strategies from class to class and school to school. Consequently, many reforms fall short in fulfilling their promise to improve student achievement and high school graduation rates for disadvantaged students.

In many cases these externally mandated top-down reforms from the district, state, or federal level are also insufficiently supported by evidence and poorly implemented (Bryk, Gomez, Grunow, & LeMahieu, 2015; Mehta, 2013). A lack of resources and appropriate systems to comply with reform goals often leave teachers – the primary targets of reform initiatives – frustrated and disengaged. Ironically, individual teachers operating in the silos of their own classroom often develop significant pedagogical expertise through successive iteration and informal experimentation over time, but it is rarely shared widely with others. The top-down nature of most reforms significantly reduces their likelihood of success by failing to recognize and build on existing teacher expertise or to include teachers as collaborators. In addition, externally imposed reform initiatives send a tacit message to educators that they are not qualified to determine which teaching strategies to use in their classrooms, undermining educators'

professional agency (Chochran-Smith & Lytle, 2006) and further reducing their desire to collaborate or participate in developing a robust, shared pedagogical knowledge base. We need a reform effort that recognizes teachers as expertise generators and supports them in creating a solid knowledge base to meet the learning need of *all* students (Mehta, 2013).

The Need for a Different Type of Reform

To tackle persistent achievement disparities for disadvantaged students, reform efforts must shift from top-down mandates to a model that situates the power for change with those closest to issues of inequitable student outcomes (Darling-Hammond, 1994). Top-down reform efforts often suffer from a lack of teacher support (Elmore, 2004; Chochran-Smith & Lytle, 2006) and fail to address the following three areas of need: (1) the need to foster teacher inquiry, (2) the need to build collaboration among educators (Bryk, et al., 2015), and (3) the need to build a shared pedagogical knowledge base (Mehta, 2013).

Improvement science is a promising framework for bottom up reform that equips educators with the methods and tools to address these three needs. Drawing inspiration from across diverse disciplines including design thinking (Kelley & Kelley, 2013), continuous improvement (Deming, 1993), and lesson study (Doig, & Groves, 2011), improvement science as conceived by Bryk et al. (2015), provides teachers with methods and tools to explore the root causes of instructional challenges they face in the classroom, create a shared vision for how best to solve them, and to engage in iterative cycles of inquiry in order to learn quickly and scale practices that are evidenced to work well.

By focusing on the local context where teaching and learning occur, an improvement science framework provides tools and methods for teachers to explore and adapt existing craft knowledge and ideas from research into a useable, evidence-based set of shared pedagogical

tools to improve outcomes for disadvantaged students. What is not yet well understood, is how teachers engage with the tools and methods of improvement science, and whether or not it affects their sense of professional agency.

Overview of the Study

This study utilized a mixed method approach to understand how teachers engaged with the tools and methods of improvement science and the impact that learning and using the tools and methods of improvement science had on their sense of agency. The study follows an intervention design and explores teachers' attitudes and experiences as they engage with the methods and tools of improvement science through ongoing improvement science professional development (see Table 1 and Appendix A). The study took place within a charter school that serves a diverse student population in the southwestern United States. A total of 20 teachers participated in the study. Data was collected in the form of pre- and post-intervention surveys, focus teacher interviews, classroom observations, and artifacts produced during the professional development sessions. Pre- and post-intervention survey data was collected to determine if there were shifts in the three domains that are relevant to teacher agency over the course of the improvement science professional development: (1) teachers' sense of belonging to an improvement community, (2) teachers' confidence in meeting their students' learning needs, and (3) teachers' perceived value of professional development support. Four focus teachers were interviewed six times throughout the intervention to better understand how teachers engage with the tools and methods of improvement science and what affordances supported or hindered their efforts. Classroom observation data and documentation produced during the improvement process were used to corroborate survey and interview data.

Significance of the Study

This study explored how teachers experience improvement science methods and tools as they engage in educational reform efforts at their school site. Specifically, this study sought to understand how the use of improvement science tools and methods impact teacher collaboration and inquiry around pedagogical practices that promote equitable student outcomes, and how these inquiries and the generation of a pedagogical knowledge base affect teachers' professional agency. Results of this study provide important insights about implementing improvement science tools and methods as a bottom-up reform strategy and the impact such reform strategies have on teacher agency.

Limitations of the Study

This study has several limitations. First, only one middle school was included in the intervention and only 20 staff members participated in both pre- and post-surveys. While multiple interviews provided insight into changes in teacher experience over time, only four focus teachers participated in the interviews so data is limited to their experiences and may not be reflective of all teachers who participated in the study. In addition, there were only 11 hours of structured improvement science professional development, and 20, thirty-minute informal check-ins over the course of eight months, which is a minimal amount of time to integrate a complex new method of testing and refining classroom practice. Finally, only one formal classroom observation per teacher was performed; however, formal observations were corroborated by numerous informal classroom observations throughout the year.

Overview of Dissertation

In this chapter I have introduced the problem, outlined the research study, and discussed the study's limitations. In chapter two of this proposal I discuss the conceptualization of teacher agency, followed by an analysis of three waves of educational reform, their relation to different conceptualizations of agency and the impact on teacher agency. From there I discuss the concept of bottom-up reform and the three elements that reform efforts must include in order to promote teacher agency. I conclude chapter two with an overview of improvement science, a promising framework for reform that incorporates these elements. In chapter three, I outline the mixed method approach taken during this study to understand if and how improvement science professional development supports teachers' sense of agency and how teachers engage with the tools and methods of improvement science. In chapter four I share the results of my data analysis, presenting evidence from focus teacher interviews, classroom observations, and document analysis, as well as the results from the whole staff pre- and post-intervention survey. In chapter five I discuss the key takeaways from the study, the implications for improvement science professional development, and areas for further research.

CHAPTER 2: LITERATURE REVIEW

Introduction

Teaching is a complex, multifaceted endeavor requiring a deep understanding of subject-specific content knowledge, pedagogy, and the needs of diverse student populations. Most professions that require such a breadth of knowledge and skills have a clearly defined knowledge base and internally controlled structures to monitor the boundaries of the profession (Mehta, 2013). However, despite the breadth of knowledge and skills required to teach effectively, educators have struggled to develop a clear set of practices that work across varied classrooms and with all types of learners. As such, they have remained vulnerable to pressure from outside interests and external accountability measures. In this chapter, I explore how conceptualizations of teacher agency have evolved over time and review the historical factors that have contributed to the current system of external accountability measures and their impact on teacher agency and the development of the teaching profession. From there, I identify three elements that are missing from current reform efforts, and argue that if incorporated in reform efforts, they could promote teacher agency. These are: (a) fostering teacher inquiry, (b) promoting collaboration among stakeholders, and (c) building a professional knowledge base. I then introduce improvement science (IS) as a framework and a tool that addresses the shortcomings of current reform efforts and discuss it supports bottom-up reform efforts in education.

Different Conceptualizations of Teacher Agency

Our understanding of and how we conceptualize teacher agency has evolved significantly over time. During the Renaissance and Enlightenment eras an increased focus on scientific rationalism paved the way for the rise of individualism, a view of humans as "free agents" – capable beings with the power to direct their own destinies (Lukes, 1973). Philosophers such as

John Locke, Adam Smith, and John Stuart Mill, identified the individual as the locus of control for action and were instrumental in anchoring this conceptualization of agency as the bedrock for progress and freedom in Western cultures. At its extreme, individualism is the defining characteristic of rational choice theory, the concept that agency is situated within the individual and that collective behavior is an aggregate of individual actions determined through individual choices (Kiser, 1996).

A contrasting view of agency emerged in the mid twentieth century with the advent of constructionism. Berger and Luckmann's *The Social Construction of Reality* situated knowledge and understanding as an emergent property of social interactions. The authors argue that all interactions, even everyday routines, are a result of an ongoing social construction of meaning created by repeated mutual validation of behaviors through interactions with others. As these established interactions become habitual, people outside the system see the patterns of interaction as 'institutionalized' and accept them as part of the fabric of society (Berger & Luckmann, 1966). For example, the Common Core State Standards were written by a small group of people, predominantly from the standardized testing industry. Academics and a spattering of teachers were used in a subsequent feedback capacity, with ultimate decision making authority remaining with the original testing industry authors (Mathis, 2010). The process involved discussion and shared meaning-making between those directly involved in the writing of the standards, however, teachers entering the profession today do not have a similar ownership of the standards and instead view them as an institutionalized structure with which they must comply, exactly as predicted by the *Social Construction of Reality* formalism.

Contemporary views of agency have expanded the social constructionist conceptualization by adding a temporal dimension to account for the human capacity to plan for

the future based on prior experiences. In their seminal theoretical paper, Emirbayer and Mische (1998) conceptualize agency as:

> The temporally constructed engagement by actors of different structural environments – the temporal-relational context of action – which, through the interplay of habit, imagination, and judgment, both reproduces and transforms those structures in interactive response to the problems posed by changing historical situations.
> (p. 970)

In this quote, the authors describe the following temporal aspects of agency as: iteration, projectivity, and practical evaluation. The iteration component of agency is enacted when an individual reactivates past patterns of behavior or thought, reinforcing past concepts and adhering to a previously established, socially determined course of action. This element of agency helps "sustain identities, interactions, and institutions over time" (Emirbayer & Mische, 1998, p. 971). The projective element encompasses an individual's capacity to generate future plans of action in light of their "hopes, fears, and desires for the future" (p. 971) and is distally related to individualism and rational choice theory. This aspect of agency is future facing and underscores the ability of the individual to shape their future context. Finally, the practical-evaluative aspect of agency refers to the "capacity of actors to make practical and normative judgments among alternative possible trajectories of action, in response to the emerging demands, dilemmas, and ambiguities of presently evolving situations" (p. 971). This element recognizes the options open to individuals in any given situation and underscores the fluidity of those options based on the continually developing specifics of the context those individuals find themselves in. These three elements deconstruct the "internally complex temporal dynamic" (Emirbayer & Mische, 1998, p.964) of agency and provide an appropriately nuanced conceptualization that underscores much of the recent research on teacher agency.

Three Waves of Educational Reform and their Relation to Different Conceptualizations of Teacher Agency

The different conceptualizations of teacher agency – from purely individualistic to structural (purely contextual) to understanding agency as an interaction between the person and the environment – have informed the three waves of educational reform. The first wave of reform dates back to the beginning of the 20th century, emerging out of the country's strong individualist foundation and coinciding with the rise of scientific management and *Taylorism*. A group of businessmen and university elites sought to develop an efficient and rational system of schooling from the patchwork of existing school options and developed a system of superintendents to create a unified vision and management system for the disparate schools (Mehta, 2013, p. 3).

The second reform movement occurred during the civil rights era of the 1960's and 1970's, when the Coleman Report highlighted that family socioeconomic status had a significant impact on student achievement (Coleman, 1966). The individualistic views of agency were replaced with a socially mediated concept of success (Berger & Luckmann, 1966). Socially perpetuated inequities in student outcomes set the stage for standard-based reforms and increased state control over educational decisions in the name of student equity. Finally, the third reform movement linked educational outcomes to economic outcomes through the *A Nation at Risk* report, which claimed that the nation's schools were failing to create a competitive national workforce (United States, 1983). The era culminated in the *No Child Left Behind* legislation that linked federal funding to the enactment of state education standards and high-stakes accountability tests (No Child Left Behind [NCLB], 2002; Cochran-Smith & Lytle; 2006).

These three major shifts in educational reform moved education from a primarily private affair for families with means, to the inclusive K-12 publicly funded system we have today. Influenced by the public and academic conceptualizations of agency at the time, each reform movement resulted in a consolidation of decision-making power at higher and higher levels, and farther away from the classroom. The following sections will examine each of these reform movements in greater detail and their impact on teacher agency. In what follows, I explore how the different conceptualizations of agency played out across three different waves of educational reform and how they shaped the current education system in the United States.

First Wave: The Progressive Era (1890 – 1930)

During the 19th century schooling in the United States shifted from a predominantly private affair, with education handled through church parochial schools, family, community, and apprenticeship, to a widespread system of government sponsored free elementary schools largely responsible for learning and socialization (Iorio & Yeager, 2011). Each state passed laws making elementary schooling compulsory in the late 1800's and banned state funding of parochial schools. Despite the expanding network of schoolhouses and increasing number of students, there was no formal training program for teachers and consequently, education quality varied significantly across the country. Many teachers at the time engaged in repetitive call and response as the dominant lesson structure, which did little to foster student engagement or motivation for learning.

The progressive era was launched by an exposé of school practices by physician Joseph Mayer Rice who documented and criticized repetitive recitation as the main form of teaching in many of the nation's schools. In a series of publications, Rice described the existing patchwork system of schools as being largely nepotistic and corrupt. Drawing on the rational logic of

scientific management, he called for better teacher training, stronger oversight of instruction by superintendents, and measurement of student outcomes (Berube, 1994).

In the early 1900's scientific management, known as Taylorism, emerged as a method to increase productivity in the manufacturing sector and had spillover effects in the overhaul of the education system. Taylorism argues for strong central oversight, relies heavily on data analysis and places a premium on the economic bottom line, in order to identify and correct inefficiencies occurring in the lower ranks. Bolstered by public support for educational reform, power over schools was wrested away from local ward boards and consolidated into districts headed by a superintendent who functioned much like a modern-day CEO (Mehta, 2013, p. 10).

This centralization of oversight resulted in a separation of the education field along gender lines. Men were trained through emerging university superintendent and education doctoral programs to manage and make decisions about teaching that the predominantly female teaching force was expected to carry out (Brinkley, 2009; Cubberley, 1922). Doctoral theses of the time almost exclusively dealt with financial or personnel management issues associated with running schools. Perhaps surprisingly, none focused on the development of pedagogical practices or other issues at the heart of teaching (Mehta, 2013). During this time, the requirements to become a teacher were only a high school diploma and a two-year teacher prep program (Cubberley, 1922). This led to a generalized failure of pedagogy to garner serious academic exploration, which hindered the development of an empirically informed knowledge base that characterizes other professions such as law, medicine, or engineering (Mehta, 2013). Instead, control of teaching and teaching practices remained largely in the hands of those farthest away from students, and teacher agency was localized to classroom decisions that fell in line with their superintendent's vision.

Second Wave: The Civil Rights Movement (1960 – 1970)

The ideas and practices of the first reform wave persisted until the civil rights movement of the 1960's with the realization that there were racial disparities in academic achievement. The second reform movement was triggered by the Coleman Report (1966), which was commissioned by Lyndon Johnson and taken up as a rallying cry for creating greater state accountability measures for education by Richard Nixon. Consistent with prevailing economic rational theory sentiment at the time, the Coleman Report was commissioned to determine the extent to which unequal distribution of resources was affecting the racial and socioeconomic differences found in student graduation outcomes. What the report actually found was that differences in incoming socioeconomic/family status had the most significant effect on student outcomes, and that increased resource allocation did not even move the needle on closing the gap between minority students and their more affluent white counterparts (Coleman, 1966). Institutionalized views of rational individualism were confronted with the overwhelming power of pre-existing cultural and social ramifications of growing up in poverty. The Coleman Report provided empirical evidence for a structural view of agency. Greater access to resources did not shift students' likelihood of academic success. Instead, their cumulative past experiences and their capacity to navigate options available to them in their given contexts impacted their ability to succeed to a far greater extent.

The academic and public shift in understanding of agency had a significant impact on how the public viewed education. The educational system as a whole was now a target of increased public scrutiny and the actual act of teaching became an important focus. States became increasingly interested in how to measure student achievement to determine the best use of financial resources and many states shifted from analyzing per-student costs, such as materials

and teacher salaries, to setting up systems to analyze the cost of *student learning* and *skill acquisition* (Light, 2003). In droves, states legislated planning, programming and budgeting systems to monitor the economic costs of producing a variety of educational results. The underlying assumption was that data provided by these systems would allow the public to hold schools accountable for student outcomes (Lessinger, 1970). Aligning with the emerging academic structural view of agency, schools were viewed as both a major contributor to systemic poverty and, perhaps confusingly, as the potential solution. The result was an increased public focus on teachers, those who were most directly responsible for student learning. Public opinion had shifted from an individualistic view of agency, which held individual students ultimately accountable for their own success to viewing teachers as a contributing environmental factor directly responsible for student outcomes.

Ironically, teachers were not afforded the same socioconstructivist view of agency and instead were held to a more individualistic concept of agency – the public considered them directly responsible for closing the achievement gap. Unfortunately, due to the management focus of the progressive era reforms, the still largely female teaching corps did not have a unified professional voice to push back on the public critiques and increased state scrutiny – no clear canon of professional practices and supporting evidence of how to achieve optimum student learning existed to present an alternative narrative to the unrest playing out through the civil rights movement (Mehta, 2013). The stage was set to develop a method to measure student success, but due to the lack of consideration of teacher professional agency, teachers were not involved in shaping the direction of the dialog.

Third Wave: Standards Based Reform (1980 – 2015)

The civil rights era highlighted disparities in student achievement, and it also exposed the compounding issue that across districts, and even schools, there was little consensus on exactly what students should be achieving or how best to measure it. States set about developing standardized tests in order to hold schools accountable for student acquisition of content knowledge. Perhaps the most influential report leading to the current state of accountability reforms was *A Nation at Risk* (1983). The authors of the report squarely placed the blame for the "rising tide of mediocrity" being produced from America's schools and the assured doom for the country's global competitiveness at the feet of educators (United States, 1983). This report provided a launching pad for the federal government to enter the educational domain in a way that had hitherto been avoided, and the decades that followed saw an unprecedented concentration of power at the level of state and federal education departments.

Despite a call for a bottom-up approach to reform efforts with top-down support, which would locate reform initiatives with those who were closest to students and best positioned to develop a professional knowledge base (Darling-Hammond, 1994), the federal government enacted *No Child Left Behind* (NCLB), legislation that updated the Elementary and Secondary Education Act of 1965 (NCLB, 2002). The 1965 Elementary and Secondary Education Act was designed to provide financial assistance in the form of Title I funds to local education agencies to support the education of students from low-income families and students with disabilities. NCLB used Title I funds as a financial lever to require that states develop standards-based accountability measures to ensure schools were meeting Adequate Yearly Progress (AYP) for all students (NCLB, 2002; Bloomfield & Cooper, 2003). AYP was based on yearly student

achievement goals as measured by state-based standardized tests that culminated in all students achieving grade level reading and math proficiency by 2014.

This top-down mandate resulted in a narrowing of teaching curriculum; particularly in urban schools that were the most at risk of having their students fail to meet AYP (Elmore, 2004). In some cases teachers were required to 'teach' daily scripted lessons that marched through the standards, effectively eliminating professional agency entirely (Chochran-Smith & Lytle, 2006).

In a scathing report on NCLB, Elmore (2004) astutely calls attention to the assumptions about educators embedded in the legislation:

> Underlying [NCLB] was a sense that the key problem was motivational and that by tightening the strings, incentives would be clarified and school personnel would find a way to increase scores – or else. (p. 243)

Proponents of high stakes accountability tests were operating on the assumption that the problem stemmed from a lack of teacher effort. However, despite the lofty NCLB goal to have 100% of students performing at grade level by 2014, the reality was quite different. At the end of 2013, less than 50% of students were proficient in reading and math, and in some states upwards of 70% of schools were failing to meet their AYP goals (NEAP, 2015). An alternative view is that "the key problem is not a lack of effort, but lack of skill; the people in the schools are the solution not the problem" (Mehta, 2013, p.261). This view is supported by the fact that at end of the decade of NCLB reforms there were no sustained generalizable gains in student achievement and no evidence found to support the hypothesis that high-stakes accountability policies, if continued, would improve student outcomes and close the achievement gap (Lee & Reeves, 2012; Mathis, 2010).

During the decade of NCLB, teachers were held to standards born out of individualistic notions of agency and educators and academics criticized the legislation for deprofessionalizing teaching and ultimately under-serving students (Chochran-Smith & Lytle, 2006):

> NCLB's conceptions of teachers and teaching are flawed – linear, remarkably narrow and based on a technical transmission model of teaching, learning and teacher training that was rejected more than two decades ago and that is decidedly out of keeping with contemporary understandings of learning. (p. 669)

Teachers who felt their professional responsibility encompassed supporting students' socioemotional development as well as their academic progress felt constrained by the increase in content to be covered under NCLB (Lasky, 2005). The negative political climate, decreased resources, and strident consequences for failure to comply, damaged their ability to form important relationships with students and increased personal feelings of failure as they struggled, along with their students, to meet the externally imposed mandates (Lasky, 2005).

School-based structures were found to mediate how teachers reacted to the demands of NCLB. Pre-existing school cultures and how schools defined successful teaching and learning shaped the professional positions open to teachers, with more pedagogically progressive institutions providing more opportunities for teachers to flexibly adapt (or not) to the increased demands of the state tests (Buchanan, 2015). These findings suggest that resisting pressures to comply with externally mandated reforms requires a strong collegial and collaborative vision of teaching, which doesn't always match the "dominant frame of teaching as an individualistic, isolated endeavor" (Buchanan, 2015, p. 714).

Emergence of the Common Core. During the decade of NCLB, each state was responsible for creating a set of standards to determine student achievement. These different standards varied considerably and made it difficult to compare student results from one state to another (Taylor, 2010). In 1996, the National Governors Association and the Council of Chief

State School Officers founded Achieve, Inc., an "independent, nonpartisan, nonprofit education reform organization dedicated to working with states to raise academic standards and graduation requirements, improve assessments, and strengthen accountability" across the country (Achieve, 2016). Achieve hired David Coleman of the College Board and other educational research analysts interested in "achievement based" assessment standards to create the Common Core, a set of national standards that could be adopted by all states. In 2010, the Common Core was released and adopted by 45 of the 50 states. Once again, teachers were held to a standard developed by those far removed from their classrooms and were expected to adapt their teaching to meet these new goals.

The Common Core is a continuation of the different conceptualizations of agency afforded to students and teachers. Student achievement is inextricably linked to their environment, in which teachers play the most significant role. Therefore, aligning with the sociocultural view of student agency, teachers are viewed as a systemic factor responsible for student success. Teachers are not afforded the same sociocultural understanding of agency, and are instead held personally accountable for student achievement.

The Three Missing Components in Current Reform Efforts

The reform efforts since NCLB have important implications for the development of teaching as a full-fledged profession and for promoting a sense of agency among teachers. The increase in external accountability structures made teachers more vulnerable to public scrutiny, positioned teachers as semi-professionals, and left them feeling less agentic. The underlying assumption of external accountability reforms is that teachers know how to meet the learning needs of all their students and have the resources to do so, *but are not sufficiently motivated to do so.* Said differently, it is the individual teachers who are the ones who are solely responsible for

low student achievement. An increase in the external accountability structures should ensure that all teachers perform at the same level. The following quote from Mehta (2013) captures how this approach reduces teachers' sense of agency:

> This external accountability under No Child Left Behind: a higher power asks a lower power to do something that neither the higher power nor the lower power knows how to do and then proceeds to publicly embarrass the lower power for failing to achieve it. The result has been predictable levels of demoralization in schools, particularly those serving high-poverty kids, with little of the hoped for widespread improvement. (p. 260)

The current reform efforts fail to address the three areas of need: (1) the need to foster teacher inquiry, (2) the need to build collaboration among stakeholders, and (3) the need to build a pedagogical knowledge base. To increase teachers' sense of agency, we need a reform effort that recognizes teachers as expertise generators and supports them in creating a solid knowledge base to define the boundaries of the teaching profession and elevate its status. Reform efforts that foster teacher inquiry and support collaboration have the potential to improve teacher agency, professionalize teaching, and ultimately improve outcomes for students.

The Need to Foster Teacher Inquiry

Change often comes from a desire to improve. Data from NCLB, the Common Core, and other standardized tests have highlighted the persistent achievement gap, but provide little guidance about how to actually improve the day-to-day aspects of teaching and learning. In *Getting to Scale with Good Educational Practice*, Elmore and Elmore (1996) point out:

> The feedback teachers receive on the effects of their practice usually comes in the form of generalized test scores that have no relationship to the specific objectives of the new practice. In other words, the conditions under which teachers are asked to engage in new practices bear no relationship whatsoever to the conditions required for learning how to implement complex and new practices with success. Why would anyone want to change their practice under such conditions? (p. 24)

In order to make progress on developing effective teaching practices, a different type of data is needed: data for improvement (Bryk et al., 2015; Elmore & Elmore, 1996). Standardized test results, disseminated long after students have left a particular classroom, is data for accountability and has little value for effecting change.

Data for improvement is actionable data that can be collected and analyzed by those close to the work in order to assess the effectiveness of a particular process or intervention. If we want to determine what teaching structures work, for whom, and under what conditions, we need to enlist the help of teachers, i.e. those who actually *do* the work. Reform efforts that support teacher inquiry, recognizing them as knowledge creators – agents that synthesize and integrate relevant information from different contexts into their own practice – are needed to help teachers take ownership of their own improvement process (Cochran-Smith & Lytle, 2006; Day, 2002).

Reforms that focus on cycles of inquiry promote ownership of both the knowledge and process in which new knowledge about teaching and learning is gained (Bodman, 2012; Bryk et al., 2015). However, it is not enough to only focus on individual inquiry; in order to generate a shared knowledge base teachers will need to share information and try out each other's practices to determine if they work in their context.

The Need to Promote Collaboration

Teacher isolation has been identified as a significant barrier to the implementation of effective reform efforts (Eisener, 1992). Effective teaching practices are often developed by individual teachers, but fail to scale past a few classrooms, if at all. Reforms that reduce teacher isolation and build learning communities have emerged as one way to spread effective teaching practices (Little, 1993; Simmie, 2007; Gomez, Gomez, Rodela, Horton, Cunningham, & Ambrocio, 2015). Collaboration and opportunities for dialog help create a shared vision of

teaching and learning. These collegial relationships serve as a structure to promote the development and spread of effective teaching strategies, and also create a shared understanding that allows teachers to push back on reform efforts that feel inauthentic or quixotic to their learning goals (Buchanan, 2015).

The Need to Develop a Professional Knowledge Base

In order to maintain decision-making power over how students are taught and assessed on their learning, teachers need to have a shared knowledge base delineating the boundaries of the profession. A shared understanding of how learning occurs and its dependence on the individual, the collective group, and the learning environment, is an important framework for developing solid pedagogical practices and classroom structures that promote student achievement and equity. Challenges to developing a shared knowledge base include teachers' past experiences and current beliefs about the purpose of schooling, and the isolated nature of teaching. Each educator brings a unique perspective and set of goals to their practice. At times, those goals are at odds with the broader institutional and cultural conversations around the purpose of education. Since shared knowledge is co-constructed, creating a unified goal such as equitable student outcomes require ongoing dialog and collective consideration (Biesta, 2015).

Improvement Science

Improvement science offers a way to address these three areas of need through a 'bottom-up' process (Bryk, et al., 2015). In this study we have adopted Bryk et al.'s conceptualization of improvement science as a promising framework, including a set of methods and tools, for bottom-up reform. Borrowing from the design industry Bryk, et al., envisioned individuals at any level within a system using the tools of improvement science to tackle localized problems of practice in order to learn quickly on a small scale and gather evidence of success. Once a practice

is evidenced to work in one context, it can be shared through collaborative learning communities to be tested in other contexts. In this way those who are closest to the problem are instrumental in the problem solving process and reform occurs from the bottom up rather than from the top down.

New to educational contexts, improvement science has it's roots in the healthcare industry, where it has been used to great effect to reduce variability in care and improve outcomes for patients (Hannan, Russell, Takahashi & Park, 2015; Martin & Gobstein, 2015; MacConnell & Caillier, 2016). Recognizing that variation in patient outcomes is likely a manifestation of the different contexts in which care is provided, Dr. Paul Batalden of the Institute for Healthcare Improvement is quoted as saying "every system is perfectly designed to get the results it gets." From this apt tautology, it can be appreciated that improvement science is used to help improvers to 'see their system [for what it is]' in order to systematically dig into the root causes of their problem and identify high-leverage areas to target for improvement. Improvements, or change ideas, are implemented on a localized scale using quick, iterative *Plan Do Study Act* (PDSA) cycles. Data for improvement is collected to determine whether the change idea should be abandoned, adapted, or adopted. The goal is to fail early, and therefore cheaply, in order to learn quickly.

Improvement science provides educators with methods and tools to engage in inquiry around improving teaching and learning, to collaborate to share promising practices, and to learn from variation and scale practices that lead to improvement (Bryk et al., 2015). Improvement science replaces top-down reform initiatives that strip educators of their professionalism with a localized strategy for improvement that situates control over practice with the educator. By focusing on the local context where teaching and learning occur and attending to variation

through collaborative networks, improvement science provides a powerful tool for teachers to explore and adapt existing craft knowledge and research, and adapt those pieces that are contextually useful into a useable, evidence-based set of pedagogical tools to aid their profession and consequently improve student achievement.

Improvement science seeks to answer the question, "What works, for whom, and under what conditions?" This question necessitates that educators adopt an improvement mindset and engage in inquiries related to their classrooms and schools. There are six principles that have been identified that are helpful to guide improvement science work in education (Bryk, et al., 2015). These are: (1) make the work problem-specific and user-centered, (2) focus on variation in performance, (3) see the system that produces the current outcomes, (4) you cannot improve at scale what you cannot measure, (5) use disciplined inquiry to drive improvement, and (6) accelerate learning through networked communities. Each of these principles is discussed below.

Principle 1. Make the Work Problem-Specific and User-Centered

The problem-oriented focus of improvement science represents a significant shift from the traditional goals of educational research, which were often focused on developing prematurely nuanced theoretical frameworks to inform top-down reform policies. However, as Bryk, et al., aptly note, "documenting that there is a problem with a new policy or intervention is not the same as knowing how to solve it" (p. 32). Improvement science places the focus on solving a problem, necessitating a shift to a more user-centered approach. This shift focuses attention onto a different set of questions, such as: What problems are encountered at the classroom or school level? What can be learned about these problems from those who are engaged in the work? How can stakeholders collaboratively design improvements that target the practical problems they encounter? This is a significant departure from the current idea that one

reform initiative can be rolled out across multiple districts (Mehta, 2013). Differences in teacher priorities, students, or school cultures, often impedes effective uptake of one-size-fits-all top-down solutions. By making the work specific to the problem at hand, the work remains relevant to the immediate stakeholders and helps increase motivation for engaging in the work (Ryan & Deci, 2000). In addition, this focus recognizes and builds off of educator expertise, bringing educators in as collaborators to move their profession forward.

Principle 2. Focus on Variation in Performance

Improvement science recognizes the ubiquitous variation in socially created systems. This variation is used both as an entry point to learn about the system and why it produces the results it gets, and is also used as an opportunity for identifying procedures and processes that can be targeted for improvement. Teaching and learning consist of a multitude of micro-processes, from classroom routines and management strategies, to discussion and group work protocols that combine to form the rich tapestry of an entire lesson. Each of these processes adds another opportunity for variation to enter the system. Attending to variation and sharing the results of change ideas tested in different contexts allows work to be standardized, enabling educators to more reliably achieve quality results.

At first blush, standard work could be misinterpreted as a call to create scripted lessons and uniformity across contexts, creating teacher automatons and stripping educators of their agency. This is the opposite of what is intended. Standardization of practices and processes reduce cognitive load for complex tasks allowing educators to more effectively engage in "the deeply intellectual work of educating children." (Bryk, et al., 2015, p.50)

The now widely used 'think-pair-share' is a good example of this. Questions posed to the whole class are often met with a terrified silence, or the same eager hands that always shoot up.

Inserting a quick think-pair-share protocol provides every student with an opportunity to think about the question for themselves and then share their thoughts with a partner. This helps students feel prepared with ideas for a full class discussion, increasing the probability that they will raise their hand to participate. The core elements remain the same – individual think time and sharing with a partner – however, the details of execution vary widely depending on the nature of the question, the goals of the lesson, the follow up activity, the grade level of students, and many other factors. Teachers make decisions using their professional judgment gauging the length of time required to think about the question, deciding whether or not they want to have students write for a few minutes first to help students articulate their thoughts, and determining in what way they want to capture student thinking after partners have shared.

Principle 3. **See the System that Produces the Current Outcomes**

Given the complexity of factors impacting education, a rigorous analysis of the system is crucial before taking action (Wilson, Dell, & Anderson; 1993). Inevitably, each stakeholder has a particular view of the problem based on his or her relative positionality, and each perspective offers new insight to the problem at hand. Bryk et al. (2015) note that the temptation to skip this system interrogation can be powerful, as those involved often have to let go of pet theories or solutions in order to step back and clearly see the system as it actually exists.

> We note that the discussions entailed in a rigorous causal system analysis can make some participants uneasy, as it requires them to be brutally honest about how and why unsatisfactory outcomes continue to occur. Because this collective self-reflection can be painful, participants may find it easier to skip the analysis and jump to the happier task of brainstorming about solutions. But short-circuiting the discussion of root causes - another manifestation of solutionitis - is a mistake. Although the analysis can be trying, it is a critical prelude to tactical action. (p. 70)

A series of improvement science tools exist to clarify the root causes of a particular problem and avoid *solutionitis* – the tendency to jump to solutions before thoroughly

understanding the problem. The first task is to narrow the problem to a manageable target. Once a clearly identified problem is defined, *empathy interviews*, *fishbone diagrams,* and *process maps* are used to expose the underlying factors and processes that contribute to the problem. These factors and processes are then used to identify areas of breakdown to target within the system. A *driver diagram* is created to visualize the *working theory of action* that will guide the improvement work to follow. The following sections explore these tools in greater detail.

Empathy interviews. In order to check our own positionality and let go of our own pet theories about why a problem occurs, it can help to conduct a series of empathy interviews with stakeholders. For example, if a group of educators is working on improving 3rd grade literacy, it would be important for them ask some of their third graders about their experiences with reading. What makes it challenging? What makes them feel successful? In order to improve their students' literacy, educators must first understand their students' experiences.

Chain of 'why' questions. To identify the root causes of a problem it is helpful to engage in a chain of *why?* questions. For example, school leaders interested in reducing chronic absenteeism at their school site might ask: Why are students missing school? Answers might bring up transportation challenges, student disengagement, health concerns, and other factors. To dig deeper, another why question can be asked: Why do students have transportation issues? Why do students feel disengaged from school? Engaging in a series of why questions holds solutionitis at bay and allows for a detailed dissection of the problem. Answers to these questions illuminate the myriad of factors that can cause students to be chronically absent, rather than falling into the trap of only considering just one favored hypothesis, and consequently provide a more thorough understanding of all of the facets of the problem.

The chain of why questions can also help to drill down into deeper levels of understanding. Many issues in education are overwhelmingly complex, such as the need to increase the number of underserved students in college, or to increase student literacy. These problems need to be narrowed down to more manageable goals, such as streamlining the college application process for typically underserved students, or increasing third grade literacy at one school site.

Fishbone diagrams. Following the why questions, a fishbone diagram is used to identify the most significant root causes of a problem (see Appendix B). These diagrams are composed of five or six 'bones' that are intended to represent key factors thought to contribute to the problem. Additional details flesh out each of these factors. A fishbone diagram is the first step in documenting a shared understanding of the problem.

Process maps. A process map is another helpful tool for identifying where systems fail. Documenting the processes involved in a system from a user perspective helps to illuminate areas to target for improvement. For example, in an improvement project focused on building new teacher effectiveness, a group of administrators determined from empathy interviews and process maps of their system that there was no consistent system of feedback for new teachers. The administrators came up with a protocol to complete biweekly observations and a quick debrief in order to provide feedback on an area requested by the new teacher. Different sites iterated on the details of the protocol and timing to determine the most effective strategies for teacher feedback across the diverse schools (Hannan, et al., 2015).

Developing a working theory of improvement. Root cause analyses; including empathy interviews, fishbone diagrams, and process maps, are used to identify previously obscured areas for further research. Relevant theory, empirical studies, and pre-existing craft-knowledge are

consulted to inform possible change ideas to try. The goal is to learn just enough about the problem at hand in order to generate change ideas, and then start testing. At every stage of developing a working theory of action care must be taken to avoid *analysis paralysis* – the tendency to get stuck trying to develop the perfect solution before taking action.

> So when evidence about what works under varied conditions is gathered, the working theory of improvement evolves. This reflects a simple fact about systems: one's understanding of a system continues to deepen through efforts to change it. It is learning by doing. (Bryk et al., 2015, p. 79)

Improvers often qualify their working theory of action with the adage, "possibly wrong, definitely incomplete" (p. 79) as a reminder to embrace the fact that their theory will likely evolve as they collect data from trying out their change ideas (Bryk, et al., 2015).

Driver diagrams. Driver diagrams are a visual representation of a working theory of action and include a measurable aim and the key levers hypothesized to affect the problem (see Appendix B). A measurable aim is stated in the affirmative and specifies *what* needs to be accomplished for *whom*, by *when*. For example, "By June 2018, 80% of traditionally underserved students in participating classrooms will demonstrate increased mathematical agency and success." This is followed by the key leverage areas, or drivers, within the system where change ideas can be implemented in order to make progress on an aim.

Principle 4. We Cannot Improve at Scale What We Cannot Measure

Measurement plays a critical role in improvement efforts. The measurement most valued is one that can be used in a timely manner for continual improvement, which is referred to as *practical measurement*. Outcome measures, such as student test scores or final grades, are too large a grain size to be useful for the bulk of improvement work. This type of accountability data can generalize about student learning over the year, but it tells educators little about which teaching strategies were or were not effective, how students experienced the different strategies,

or what could be done to improve them. Measurement for testing whether or not a change idea is an improvement must be fine-grained enough to inform whether an idea needs to be adopted, adapted, or abandoned. Survey data in the form of student work, exit cards, weekly attendance reports, observations, or other forms of practical measurement are much more useful for a teacher to determine if a particular practice is working, than a validated research-oriented survey designed for statistical analysis (Bryk et al., 2015).

Principle 5. Use Disciplined Inquiry to Drive Improvement

Reform initiatives rolled out by district leaders and school administrators are usually done on a large scale and often play out according to the following paradigm, "implement fast, learn slow, and burn goodwill as you go" (Bryk, et al., 2015, p. 113). Reform efforts, even when based on sound evidence from research, are frequently implemented too broadly to be adequately supported by available resources, and ultimately report poor or nonexistent results (Lee & Reeves, 2012; Mathis, 2010). This repeated cycle of failed reforms has eroded teachers' trust in new initiatives and culminates in low teacher buy-in for new reform efforts (Buchanan, 2015; Elmore & Elmore, 1996).

Improvement science situates control of reform efforts with educators and their students. The following questions guide improvement work: (1) what specifically are we trying to accomplish? (2) What change might we introduce and why? And (3) how will we know that a change is actually an improvement? The first question is addressed through digging into the problem using empathy interviews, fishbone diagrams, and process maps. This exploration helps clarify the problem and provides a solid understanding to construct a driver diagram and working theory of action. Deciding what improvements to try requires identifying high-leverage change ideas that can be implemented on a small scale in order to learn quickly. Small-scale

implementation also minimizes damage in the case of failure, which is in stark contrast to top-down policy initiatives such as NCLB. Once a change idea has been selected, a method of measuring whether the change idea is actually an improvement needs to be determined and implemented.

Plan Do Study Act cycles. Once change idea and measurement methods have been identified, a PDSA cycle is carried out. PDSA cycles emulate the systematic experimentation of scientific research, but are adapted to fit everyday practices and are at the heart of improvement science. Hypotheses about how best to improve are generated and tested. Evidence of their effectiveness is collected and used to inform whether a change idea should be adopted, adapted, or abandoned (see Appendix B for PDSA template). Based on the evidence collected a new PDSA cycle is started and an iterative process of learning about the problem and its possible solutions occurs. In many cases an idea may fail completely and need to be abandoned, while in others the change idea will require numerous cycles testing different adaptations before working reliably.

Moving towards scale. Once an idea works reliably in one context it can be tried in others. A successful classroom procedure can be tested by other teachers and the iteration process continued in order to learn how the change idea best works in the new context. As a change idea is tested across diverse contexts, elements emerge that must remain consistent for successful implementation, and other elements are recognized as context specific. For example, in the think-pair-share protocol, the think time is crucial to its success, and the beginning teacher effectiveness study found that the two-week feedback cycles were considered optimal by new teachers across contexts (Hannan, et al., 2015). Evidence of reliable implementation across a variety of contexts can be used to convince new testers from outside the immediate community,

creating the opportunity to learn from iteration in even more diverse contexts and effectively scaling the practice (Bryk et al., 2015).

Principle 6. Accelerate Learning Through Networked Communities

Networked improvement communities (NICs) allow for the resourcing of social capital. As NIC participants meet to discuss improvement efforts and share learning, trust is built and members are more likely to adopt innovations being tried by their colleagues for testing in their own contexts (Bryk et al., 2015). Key support for a NIC is provided by a hub, a central person or group responsible for providing structural and analytical support for NIC participants. The hub orchestrates opportunities for NIC members to convene and discuss progress, serves to collect and analyze NIC data for patterns or bright spots, and feeds the information back to participants to distribute new learning across the NIC efficiently (Bryk et al., 2015).

Research Questions

This research study explored the experiences of 20 middle school teachers at a charter school in the southwestern United States as they engaged in learning and implementing improvement science methods to improve their practice. The following research questions were investigated in this study using a mixed methods approach:

1. What is the impact of improvement science professional development on teacher agency?
    a. To what degree does the intervention improves teachers' sense of belonging to an improvement community?
    b. To what degree does the intervention improves teachers' sense of mastery and confidence in their ability to meet the learning needs of their students?
    c. To what degree does the intervention improves teachers' perceived value of professional development supports?

2. How do teachers engage with improvement science tools and methods?
    a. How do teachers use improvement science tools and methods?
    b. What structures support teachers in engaging in improvement work?
    c. In what ways do improvement science tools and methods impact teacher agency?

CHAPTER 3: METHODOLOGY

Introduction

This study utilized a mixed method approach to understand the impact of improvement science professional development on teacher agency across three dimensions: (1) to what degree does the intervention improves teachers' sense of belonging to an improvement community? (2) To what degree does the intervention improves teachers' sense of mastery and confidence in their ability to meet the learning needs of their students? And (3) to what degree does the intervention improve teachers' perceived value of professional development supports? The study also explores how teachers engage with the tools and methods of improvement science, specifically: (1) how do teachers use improvement science tools and methods? And, (2) what structures support teachers in engaging in improvement work?

In this chapter, I first discuss the broader organizational context of the study in order to identify the norms and expectations that were embedded within the culture of the organization to situate the research. Second, I share the structure of the teacher professional development, which introduced improvement science as a tool and provided heuristics to help teachers engage in improvement work in their own practice. I then describe the participants of the study and the procedures that were used to collect both quantitative and qualitative data. I close this chapter with a discussion about how data was analyzed to answer the research questions.

Context of the Study

The study was conducted at a charter-based middle school in the southwestern United States. The student population consisted of 334 students and was 51% White, 36% Hispanic/Latino, 4% African American, 3% Asian and 3% Filipino. Fifty percent of students qualified for free and reduced lunch and 8% were classified as English language learners.

The school's twenty-five staff members consisted of one administrator, fifteen teachers, five inclusion teachers and four support staff. The school employed a project-based pedagogy where teachers were encouraged to collaborate on project design and taught in cross discipline teams. Each teaching team, composed of a humanities teacher and a math/science teacher stayed with a cohort of roughly fifty students, split into two classes of twenty-five students each for the duration of the year.

Several school structures exist to encourage and support teachers in their development of curriculum and project design, including eight summer staff days, half days every other Wednesday to create two additional two hour professional development meetings per month, and weekly morning meeting times. The majority of the teachers appreciate the autonomy, but also acknowledge that it often takes more time than allotted and significant effort to develop meaningful and rigorous curriculum for all their students.

Participants

Twenty teachers participated in the study, fifteen core subject teachers and five inclusion teachers. The teachers were evenly split by gender, with ten women and ten men participating. Of the twenty teachers, one was Asian-American, two were African-American, two were Hispanic/Latino, and fifteen were European-American.

Improvement Science Professional Development

As part of the study, teachers participated in improvement science professional development during the academic year. The improvement science professional development (PD) included five, two hour workshops, weekly (first semester) and bi-weekly (second semester) 30-minute check-in meetings, and a final one-hour workshop at the end of the academic year. Table 1 summarizes the activities covered in the different workshops. The

purpose of the PD was for teachers to learn and use the tools and methodologies of improvement science to address a felt need at their school or a problem of practice.

In the first PD session, teachers explored previous year's data from the Smarter Balanced Assessment Consortium (SBAC) statewide standardized tests and Youth Truth survey data. The Youth Truth survey measures student and stakeholder perceptions about school culture, academic rigor, and college and career readiness. The teachers used the data to identify potential areas for improvement and formed groups based on improvement interest. Five improvement groups emerged: (1) student belonging, (2) math, (3) literacy, (4) English Language Learner (emerging bilingual student) support, and (5) assessment (Table 2). In their improvement groups teachers generated interview questions for students in order to collect more information about their chosen problem.

The second PD workshop took place two weeks later. Teachers compared data from empathy interviews and used a fishbone diagram protocol (Appendix C) to develop a problem statement and explore potential root causes. After completing fishbone diagrams for their problems of practice, they used an interrelationship digraph protocol (Appendix D) to complete an interrelationship digraph to determine which of their root causes was the highest leverage and should be the area of focus for subsequent improvement efforts.

In the third PD workshop teachers developed a theory of action using a driver diagram protocol (Appendix E) and their deeper understanding of the problem generated through completing a fishbone diagram and the interrelationship digraph during the root cause analysis. Teachers developed a clear and measurable aim statement for their improvement work (Table 2), identified primary drivers – *what* they would have to focus on to drive improvement, and

secondary drivers – *where* in their system they could focus on the primary driver. Finally they selected high leverage change ideas and planned their first Plan Do Study Act (PDSA) cycle.

During the fourth PD workshop teachers analyzed their PDSA data and determined if they should adopt, adapt, or abandon their change idea, and planned for their next PDSA cycle (see Appendix B for a model PDSA template).

In addition to the PD workshop sessions, teachers met weekly in improvement groups for a thirty-minute morning check-in during the first semester. Second semester these meetings shifted to biweekly due to the addition of a collegial coaching initiative and teacher feedback that weekly improvement group check-ins were too frequent. During weekly improvement group check-ins teachers looked at data from student work such as quiz scores, exit cards, or survey data and planned change ideas.

Table 1. Improvement Science Professional Development Workshop Activities

Professional Development Workshop	**Activities**
1st Workshop September 9th, 2016	Digging into data protocol Deciding on improvement groups Prepping for empathy interviews
2nd Workshop October 5th, 2016	Comparing themes from empathy interviews Fishbone diagram protocol Interrelationship digraph Change idea brainstorm
3rd Workshop October 19th, 2016	Driver diagram protocol Change idea brainstorm Planning a Plan Do Study Act (PDSA) cycle
4th Workshop November 2nd, 2016	Looking at PDSA data Planning next PDSA
5th Workshop April 19th, 2017	Revisiting driver diagrams and fishbone diagrams Planning a PDSA Goal setting for end of the year
6th 1 Hour Workshop June 6th, 2017	Success analysis protocol
30 min weekly/biweekly check-ins	Looking at data Change idea selection Planning next PDSA

Table 2. Improvement Group Aim Statements

Improvement Group	Aim
Literacy	By June 2017, 100% of HTMNC Students will show growth in their reading comprehension
Math	By June 2017 80% of students will be on grade level as assessed by MAP (adjusted to 50% after looking at baseline MAP scores in December)
English Learner Support	By June 2017, 100% of EL Students will show growth in their academic skills/mindset and report an increased sense of belonging
Belonging	By June 2017, 100% of HTMNC students will feel an improved connection to the school in some fashion
Assessment	N/A

Instruments and Data Collection

Quantitative data collected during this included pre- and post-intervention surveys. Qualitative data collected during the study included improvement science tools documentation, such as system maps, fishbone diagrams, and Plan Do Study Act (PDSA) forms that were generated by teachers as part of the improvement science professional development. Four focus teachers were also selected through personal invitation to participate in a series of open-ended semi-structured interviews and classroom observations. The multiple data collection methods were used to develop and corroborate categories and themes.

Pre- and Post-intervention Teacher Surveys

An anonymous pre- and post-intervention survey was given to the teachers before and after the improvement science professional development (Appendix F). The survey consists of 17, five-point Likert scale questions that range from strongly disagree to strongly agree. The questions on the surveys were organized around three domains relevant to teacher agency: (1) teachers' sense of belonging to an improvement community, (2) teachers' mastery and

confidence in meeting their students' learning needs, and (3) teachers' perceived value of professional development support. The survey questions are adapted and modified from student agency surveys developed by the Carnegie Foundation for the Advancement of Teaching (2013).

The first domain, belonging to an improvement community, draws upon the sociocultural view of agency that suggests human agency is about doing things with other people in social settings with the cultural tools (in this case improvement science) available to them. The second domain, teachers' mastery and confidence in meeting their students' learning needs, draws upon the individualistic view of agency that suggests human agency is about exercising actions related to a person's goals, in this case teachers' goal of meeting their students' learning needs. Finally, the third domain, teachers' perceived value of professional development support in achieving their own goals, draws upon the structural view of agency that suggests human agency is about opportunities that are available in the environment, in this case the professional development provided by the organization to teachers, that allow people to see themselves as agents. Twenty teachers took the first survey at the beginning of the improvement science professional development intervention and again at the end of the intervention.

Interviews

Four teachers agreed to be interviewed over the course of the improvement science professional development. Interviews are an established and practical method for data collection and can illuminate important themes about a person's lived experience (Valenzuela & Shrivastava, 2009). Before the initial interview the purpose of the study was reviewed and participants signed a written consent form to participate in the study as well as to be audio taped. Each teacher engaged in six semi-structured interviews conducted throughout the year (see Appendix G for interview template) to explore teacher engagement with improvement science

(Plano Clark & Creswell, 2010). Interview questions were based on where participants were in the scope of the ongoing professional development and explored how teachers used the tools and methods of improvement science, whether or not they valued the tools and methods they were learning about, and whether or not they felt like they belonged to an improvement community. An external transcription agency transcribed the interviews. Each interview was listened to in order to check quality and correctness of the transcription.

Improvement Science Documentation

Documents generated by teachers during the improvement science professional development, including fishbone diagrams, driver diagrams and completed PDSA forms, were collected for analysis. Other forms of documentation that were also collected include professional development agendas, meeting agendas, protocols, survey data, exit card data from staff professional development sessions, and classroom tools created by teachers during their improvement work.

Classroom Observation Data

One formal classroom observation of each of the focus teachers was conducted over the duration of the study and several informal classroom observations of shorter duration (see Appendix H for observation template). The purpose was to collect descriptive data that would supplement and corroborate the interview data. The observations focused on teacher implementation of change ideas and teacher data collection methods to determine if their change idea was working.

Data Analysis

In order to answer the research questions below, the following data was collected and analyzed:

- Research Question 1: What is the impact of improvement science professional development on teacher agency?

Pre- and post-surveys were analyzed with exploratory techniques to understand the impact of the intervention on teachers' sense of belonging to an improvement community, sense of mastery and confidence in their ability to meet the learning needs of their students, and perceived value of professional development supports. Exploratory techniques were used to measure frequencies, percentages, and averages of the data.

- Research Question 2: How do teachers engage with improvement science tools and methods?

A constant comparative method was used for qualitative data analysis (Glaser & Strauss, 1967). Interview transcripts and observation notes were open coded in an iterative process to uncover significant themes (Esterberg, 2002; Saldaña, 2009). After an initial round of open coding, a second round of coding was completed using the research sub-questions as a lens to group existing codes and identify emerging themes: (1) how do teachers use improvement science tools and methods? (2) What structures support teachers in engaging in improvement work? And (3) in what ways do improvement science tools and methods impact teacher agency? Studying improvement documentation, professional development feedback surveys, and the pre- and post-intervention survey confirmed and corroborated these themes.

The Role of the Researcher

In this study I held the role of researcher-facilitator and co-facilitated the six professional development workshop sessions with teacher leaders from the school. Since one of the goals of the professional development was for teachers to have ownership over the improvement process, every effort was made to share leadership responsibilities and the weekly and biweekly check-in

sessions for improvement teams were teacher led. Feedback was requested after all improvement science professional development sessions and focus teachers were encouraged to be candid in their interviews about their experiences engaging with improvement science methods and tools. Despite efforts to acknowledge and reduce any perceived power imbalances related to my position as researcher-facilitator, my role may have impacted how teachers experienced and used the tools of improvement science during the professional development sessions as well as what the focus teachers chose to share during interviews.

CHAPTER 4: RESULTS

Introduction

This research study used both quantitative and qualitative methods to explore how teachers engage with improvement science methods and tools as provided through a series of six professional development sessions over nine months. Pre- and post-intervention surveys were used to document shifts in teacher agency as conceptualized as (1) an increased sense of belonging to an improvement community, (2) an increase in teacher perceptions of their ability to meet the learning needs of all their students and (3) an increase in the degree to which teachers valued the improvement science professional development opportunity at their school site.

Interview data from focus teachers were used to elucidate themes around how teachers learn, use, and share improvement science methods and tools and what professional development or other school structures help or hinder their engagement. Data from observations and document analysis were used to triangulate with emerging themes from the interviews and survey responses to thoroughly answer the research questions. In what follows, I share the findings of this research study. I first share the findings from the pre-and post-surveys. From there I discuss the findings from interview and observational data from the four focus teachers.

The Impact of Improvement Science Professional Development on Teacher Agency

Results from the pre- and post-intervention survey data indicate positive shifts in 9 of the 13 measures related to teacher agency. The survey tool consisted of five questions assessing teacher perceptions of belonging to an improvement community, five questions assessing teachers' sense of confidence that they could meet the learning needs of all their students, and three questions assessing whether teachers felt the improvement science professional development was of value to their practice (Table 3).

Table 3. Percentage of Teachers Responding Agree and Strongly Agree on the Pre-and Post-intervention Survey Measures

Measures	Pre-intervention Survey Percent Agree or Strongly Agree	Post-intervention Survey Percent Agree or Strongly Agree
Measures of teacher perceptions of belonging to an improvement community.		
My school feels like a community that is always trying to improve to get better at meeting the learning needs of our students	96%	95%
My colleagues actively test out new teaching practices to meet the learning needs of all their students	83%	80%
My colleagues use data such as student work, exit cards, or student surveys to determine if their teaching practices are working for all their students	71%	80%
I share new teaching strategies with my colleagues	79%	85%
I learn new teaching strategies from my colleagues	75%	95%
Measures of teacher perceptions of confidence that they can meet the learning needs of their students.		
I feel confident that I can meet the learning needs of all my students	71%	70%
I feel confident that I can learn new pedagogical tools to meet the learning needs of all my students	96%	100%
I feel confident that I can create new pedagogical tools to meet the learning needs of all my students	79%	75%
I actively test out new teaching practices to meet the learning needs of all of my students	78%	90%
I use data such as student work, exit cards, or student surveys to determine if my teaching practices are working for all my students	75%	80%
Measures of teacher perception of the value of improvement science professional development for their practice.		
The professional development I receive at my school helps me meet the learning needs of my students	63%	75%
The professional development I receive at my school helps me improve my teaching practice	66%	80%
The professional development I receive at my school will be useful to me in the future	79%	90%

Teachers reported an increased sense of belonging to an improvement community on three of the five belonging measures. In both the pre- and post-surveys the percentage of teachers who felt that they were part of an improvement community were high, 96% and 95% respectively, despite a slight decrease over the year. A similar decrease was observed in teachers reporting that that their colleagues were actively testing new teaching strategies, 83% on the pre-intervention survey compared to 80% on the post-intervention survey. In contrast, the percent of teachers choosing agree or strongly agree increased on post-intervention survey for the final three measures. When asked whether they felt that their colleagues used data to determine if their teaching practices were working for all students, 80% of teachers agreed or strongly agreed post-intervention as compared to 71% at the beginning of the year. More teachers also reported sharing new strategies with colleagues after the intervention, 85%, as compared to 79% of teachers at the beginning of the year. In addition, 95% of teachers felt that they learned new strategies from their colleagues compared to 75% of teachers at the beginning of the year. Taken together, these data indicate that teachers felt a strong sense of belonging to an improvement community and more teachers engaged in sharing data and learning teaching strategies with one another.

With respect to confidence in their ability to meet the learning needs of all their students, the number of teachers who agreed or strongly agreed on the post-intervention survey increased on three of five measures. The number of teachers who felt confident that they could meet the learning needs of all their students decreased slightly from 71% to 70% over the course of the intervention, and the number of teachers who felt confident that they could design new pedagogical tools to meet the learning needs of their students decreased from 79% to 75%. In contrast, the percent of teachers who felt confident that they could *learn* new pedagogical tools

to meet the learning needs of their students increased from 96% to 100% and the percent of teachers who reported actively testing out new teaching practices increased from 78% to 90%. There was also an increase in the percent of teachers who reported using data to inform their practice from 75% to 80% after the intervention. These data suggest that while teacher confidence that they could meet the learning needs of their students remained the same after the intervention, teachers were actively using data to test out new pedagogical practices and felt confident they could learn new tools to meet the learning needs of their students.

The percentage of teachers reporting that they valued the improvement science professional development increased across all three measures on the post-intervention survey. The percentage of teachers who reported that the improvement science professional development they received would help them meet the learning needs of their students increased from 63% to 75% and the percentage of teachers reporting that the professional development helped them improve their teaching practice increased from 66% to 80% over the course of the intervention. And finally, 90% of teachers felt that the improvement science professional development would be useful to them in the future.

From the data in Table 3, it is clear that the teachers rated the measures quite high in the pre-intervention survey (the lowest percent of teachers indicating agree or strongly agree on any of the pre-survey measures was 63% and the highest was 96%), indicating significant levels of teacher agency to begin with. From informal observations of the teachers, it was clear they had a strong culture of collegial support and teachers had a strong sense of belonging and confidence that they could meet the learning needs of all their students at the beginning of the year. In light of these observations it is possible that the five-point Likert scale was insufficient for capturing shifts in teacher beliefs and *response shift bias* – the phenomenon that you don't know what you

don't know – may have occurred (Howard & Dailey 1979). Response shift bias is common in situations where participants answer questions before engaging in a lesson or educational activity that changes the participants' internal frame of reference for the questions, and then they answer the questions again in a post-intervention survey. In this case, teachers' initial sense of agency may have been inflated as compared to post-intervention survey responses, after they developed a deeper understanding and appreciation for their common problems of practice through engaging with improvement science tools and methods.

Interview data throughout the year corroborated this idea and it was apparent that significant shifts in perceptions occurred over the improvement science intervention. As evidenced after one professional development session where teachers examined the root causes of a problem of practice, and one teacher exclaimed, "That was the best professional development I've ever had. That was so useful!" The following findings sections explore interview and observation data from case study teachers and provide additional insight about how teachers engage with the tools and methods of improvement science.

Teachers' Engagement with Improvement Science Tools and Methods

To develop an in-depth understanding of teachers' engagement with improvement tools and methods, the researcher selected and focused on four teachers as cases. Analysis of interview transcriptions, classroom observations, and documents generated through participation in the improvement science professional development, revealed four emerging themes around the research sub-questions (1) how teachers use improvement tools and methods, (2) what structures support teachers in engaging in improvement work, and (3) how improvement science informed their sense of agency (Table 4).

Table 4. Emerging Themes from the Analysis of Qualitative Data from Focus Teachers

How teachers use improvement tools and methods
Theme 1: Using Improvement Tools to Define the Problem ● Empathy interviews ● Fishbone diagrams ● Interrelationship digraphs ● Driver diagrams Theme 2: Using Data to Inform Practice ● Outcome measure data ● Data for improvement
Structures that support teachers in engaging in improvement work
Theme 3: Structures that Support Improvement ● Protocols ● Improvement teams ● Regular meeting times ● Internal and external capacity and support
How improvement science inform teachers' sense of agency
Theme 4: Changes in Teacher Agency Over Time ● Intentionality, inquiry, and changes to individual practice ● Sharing and spreading change ideas ● Identification as an improver in an improvement community

Table 5 lists the focus teachers, identified by pseudonyms, and their demographic data. The four focus teachers are Zara, an African-American woman, Colby a Mexican-American man, and Jennifer and Stephanie, two European-American women. Each focus teacher had a minimum of two years teaching experience.

Table 5. Focus Teacher Demographics

Demographic Category	Jennifer	Colby	Stephanie	Zara
Gender	F	M	F	F
Grade level	8th	8th	7th	8th
Subject	Math/Science	Math/Science	Math/Science	Humanities
Years Teaching (as of beginning of the study)	2	5	2	2
Years at Current School (as of beginning of the study)	1	3	2	0
Ethnicity	European-American	Mexican-American	European-American	African-American

How Teachers Use Improvement Tools and Methods to Define the Problem

A series of improvement science tools exist to clarify the root causes of a problem and to avoid *solutionitis* – the tendency to jump to solutions before thoroughly understanding the problem. Tools such as empathy interviews, fishbone diagrams, interrelationship digraphs, and driver diagrams (Appendix B) are used to conduct a root cause analysis and inform the development of an initial theory of action that can guide quick iterative cycles of inquiry in order to learn more about the problem (Bryk et al., 2015).

The first principle of improvement science – make the work problem specific and user centered – anchors improvement in the felt needs of a particular community. During the first improvement science professional development session teachers looked at student data from the Smarter Balanced Assessment Consortium (SBAC) test and Youth Truth survey data from the previous academic year, and then responded to the following journal prompts: *What areas for improvement for our school are emerging from this data? What are you passionate to work on this year?*

After journaling, teachers created a list of possible areas to improve and then voted on five areas to focus on over the year: literacy, math, student belonging, bilingual student support[1], and assessment. Staff members then each decided which of the five focus areas they felt most passionate to work on and formed five improvement groups. The four focus teachers joined three of the improvement groups; Jennifer decided to focus on math, Colby and Stephanie decided to focus on bilingual student support and joined the English Learner support improvement group, and Zara joined the literacy improvement group.

Empathy Interviews

Empathy interviews are an approach for collecting qualitative data about a person's lived experience and were used by the teachers formally and informally throughout the year. Teachers found the interviews useful for understanding students' experiences and perceptions related to their improvement goals. In addition, data from empathy interviews also challenged teacher beliefs about the nature of the problem they were trying to improve. To understand their chosen area for improvement more deeply, the four focus teachers and their improvement group members started out by conducting empathy interviews with students.

After looking at the math SBAC data, Jennifer's math improvement group began to wonder about the low scores for many of their students. One hypothesis the group put forth was that student test anxiety might be a factor contributing to the low scores. Perhaps students were getting anxious because test taking was not a big part of their project based learning culture and the SBAC assessment felt unfamiliar to them. However, empathy interviews with students didn't corroborate their hypothesis. When asked which tools were most useful for learning about their problem, Jennifer shared:

[1] Staff called themselves the *EL group* short for English Learner support group and the term *EL*, and *EL group* will be used throughout this document to refer to emerging bilingual students and the staff improvement group focused on supporting them.

> Interviewing students. Having empathy interviews at the very beginning of the process, because we have all these preconceived ideas of what's happening. We think that because we don't do tests that [students] are stressed when they take a test. My kids said they're not – there's no stress at all. They just take the test, they do their best – and it was one of those eye-opening moments where maybe what we're saying is happening is a cover-up for what could be missing.

Jennifer's reflection highlights how empathy interviews helped her improvement group understand the problem more accurately. This quick method of checking assumptions by asking stakeholders how they experience a particular situation helped the math improvement group avoid wasting time and energy on a test anxiety problem that didn't exist in their context.

Empathy interviews were also used by the focus teachers throughout the year as a method to identify particular students' unmet needs. Colby claimed that he and another member of his group were "planning on doing some interviews with the same people we started the year with. Kind of to reset ourselves." For one emerging bilingual student in particular he wanted to determine whether she still felt that the group she was sitting with in class was working for her, and whether she was feeling challenged enough. "This is why I wanted to do another empathy interview. Am I just keeping them in that spot together? Is that keeping them from stretching with the rest of the class?"

In addition to the student empathy interviews conducted at the beginning of the year, both Stephanie and Zara also used empathy interviews throughout the year to better understand their students' experiences. Stephanie held one-on-one check-ins with her bilingual students to see which support strategies were working for them, and Zara conducted interviews with students who demonstrated low growth on midyear standardized tests to better understand how they experienced taking the assessment.

Teachers used formal empathy interviews and informal check-ins over the course of the year to get a sense of individual student experiences related to their chosen improvement focus.

The term 'empathy interview' became a commonly understood idea and practice for the teachers. When asked what improvement tools she might use the following school year, Jennifer replied "empathy interviews."

Fishbone Diagrams

During the second improvement science professional development session the four focus teachers shared their initial empathy interview data with their improvement groups and used the new insights about their issue to collaboratively craft a problem statement. Table 6 details the problem statements and major root causes generated by all five of the improvement groups.

Table 6. Improvement Group Problem Statements and Fishbone Categories

Improvement Group	Problem Statement and Identified Root Causes
Literacy	*There are discrepancies in student literacy across different student subgroups.* ● Different ideas of what good literacy instruction is across classes ● Some students lack access to content ● Lack of good literacy instruction ● Some students are English Language Learners ● Varied parent support for literacy at home ● Uneven support systems within the school ● Student motivation factors
Math	*Students do not feel confident in math class.* ● Difficulty completing homework ● Tension between doing projects and devoting time to practice math skills ● Students have conceptual gaps/lack foundational math concepts ● Students lack confidence and don't persist at problem solving ● Grading practices ● Uneven support for math at home
English Learner Support	*English Learners scored low on SBAC test and reported a lower sense of belonging on the Youth Truth survey.* ● Fear of participation ● Support systems at home ● Part of school minority group ● Vocabulary ● Teacher strategies ● Content ● Test taking problems ● Metacognition
Belonging	*Some student groups do not feel like they belong* ● School wide culture fits some students and not others ● New students still learning to navigate the system ● Latino students are still a minority
Assessment	*We can't tell from data which of our practices are working and which are not working* ● Results don't clearly match back to teaching methods used ● We don't have assessments that measure growth ● Data compares schools with different pedagogical philosophies ● Data we do have seems skewed towards things we find less meaningful ● Tests we use change often

During the rest of the professional development session the focus teachers used a protocol to dig into the root causes of their identified area for improvement and constructed a fishbone diagram with their improvement groups. The protocol prompted the teachers to brainstorm as many causes as possible that contribute to their problem statement and then to group them into categories. The broader categories or 'bones' represent key factors that contribute to the problem. A fishbone diagram is the first step in collaboratively creating and documenting a shared understanding of the problem and the contributing contextual factors (see Appendix B for a fishbone diagram template and models from the EL improvement group, and Appendix C for the fishbone protocol).

Stephanie's group were concerned that emerging bilingual students scored low on the SBAC test and reported a lower sense of belonging on the Youth Truth survey (Table 6). Stephanie found that "the fishbone diagram really helped because I was able to see the five or six key areas that we're focusing on. That process was really helpful for me." From the fishbone diagram, Stephanie and her improvement group were able to determine one area that was within their locus of control to focus on. "We chose one part of the fishbone which was teacher strategies, and that's what we're really focusing on for our change ideas."

Stephanie also shared how the fishbone protocol helped to clearly identify areas that would be high impact for the teachers to focus on and areas where they might have less influence. She said, "An example on our fishbone was at-home support. We can't really impact that as much as we can impact what happens in our classroom. Going through that with the group was helpful."

Colby corroborated Stephanie's comments, "I think the fishbone was probably the [tool] that stands out the most. [It] really honed down what our group wanted to focus on."

After completing the fishbone diagram, Zara's group decided to focus on using common literacy instructional practices in order to provide more consistent literacy instruction across the school. Whereas for Jennifer, the fishbone diagram yielded two compelling and competing root causes. The first was that students had gaps in their conceptual understanding of foundational math concepts and the second was that students lacked confidence and did not persist at problem solving. The group was divided about which root cause to focus on. Jennifer wanted to focus on building up student foundational knowledge, but some of her group members felt that focusing on student confidence in problem solving would be more productive.

The fishbone diagram generation process facilitated group discussion around a shared problem, supporting group consensus on the major factors impacting the problem. The next step was to determine which of the root causes to focus improvement efforts around. Which root causes were contributing the most to the problem but were still within the group's locus of control?

Interrelationship Digraphs

An interrelationship digraph tool is used to determine causal relationships between different root causes. It helps identify which root causes are more significant and which may actually be symptoms of other root causes. The dialogical process involves determining the direction of causality between each pair of root causes, those with more arrows pointing to them are likely more superficial and may actually be symptoms of other root causes, whereas those with more arrows pointing outwards are likely to be at the heart of the problem. In order to have the best chance of solving a problem, improvement efforts need to focus on the root causes with the most outgoing arrows.

The focus teachers used a protocol to construct an interrelationship digraph with their improvement groups to come to a common understanding of which root causes to target first in their improvement efforts (see Appendix D for protocol).

Colby and Stephanie's improvement group used the interrelationship digraph tool to determine that the root cause t*eacher strategies* – a lack of specific teaching strategies for supporting bilingual students – was likely negatively impacting some of the other root causes, such as vocabulary development, fear of participating in class, and student metacognition about their own learning. The group determined that testing teaching strategies that promoted vocabulary development, supported class participation, and helped students think about how they learn, was not only within their locus of control, but also likely to result in improvement of the other root causes.

Zara's group also identified classroom instruction as the most significant root cause to focus on and started to discuss which literacy structures within their classes might have the largest impact on student literacy development.

Of all four focus teachers, the interrelationship digraph probably helped Jennifer the most. With her math improvement group caught in a debate about which root cause to focus on, Jennifer highlighted how the tool helped them choose between focusing improvement efforts on student confidence in problem solving and student's lack of foundational math skills:

> We kept saying, 'oh we think it's their confidence that is affecting their foundations' and 'confidence is what's affecting being able to do homework' and then we realized that we don't actually think it is confidence. We think it's the foundations and their understanding of the foundations that affects their confidence. And so that was a really cool way to step back and see that maybe we're looking at the wrong piece. As teachers we're interpreting it this way, but maybe that's not what's actually happening with the student.

Using the interrelationship digraph to determine the more significant root cause, the group ultimately decided to focus on incorporating more foundational math concepts into their lessons.

For each of the focus teachers the interrelationship digraph was a useful tool to determine which root cause they should ultimately make the focus of their improvement efforts.

Driver Diagrams

Using information from the root cause analysis – including empathy interviews, fishbone diagrams, and interrelationship digraphs – the focus teachers used another protocol to develop a driver diagram to guide their improvement efforts (Appendices B and E).

Driver diagrams are a visual representation of a working theory of action and include a measurable aim and the key drivers – areas in a system – on which to focus improvement efforts (see template example in Appendix B). A measurable aim states *what* needs to be accomplished for *whom*, by *when*. Table 7 summarizes the aim statements generated by each improvement group during the third improvement science professional development meeting.

Table 7. Improvement Group Aim Statements and Primary and Secondary Drivers

Improvement Group Aim Statement	**Primary and Secondary Drivers**
Literacy: *By June 2017,100% of [school] Students will show growth in their reading comprehension.*	● Classroom reading instruction ○ Reading strategies
Math: *By June 2017 80% of students will be on grade level as assessed by MAP. (adjusted to %50 after looking at baseline MAP scores in December)*	● Student problem solving skills ○ Classroom instruction ● Content alignment across grade levels ○ Discipline meetings ● Student confidence ○ Classroom culture and routines
English Learner Support: *By June 2017, 100% of EL Students will show growth in their academic skills/mindset and report an increased sense of belonging.*	● Home - school relationship ○ Communication ● Teacher strategies ○ Professional development ○ Teacher talk, response and questioning ● Student Participation ● Access to content ● School relationships ○ Classes, group work, recess, halls
Belonging: *By June 2017, 100% of [school] students will feel an improved connection to the school in some fashion.*	● Peer to peer interactions ○ Advisory culture ○ Welcoming school environment ● Student to teacher interactions ○ Clubs ○ Sports ○ Lunch groups ● School culture ○ Advisory Olympics ○ Restorative circles ○ Classroom culture
Assessment: *Did not generate a driver diagram*	N/A

Stephanie found the sentence frames on the driver diagram helpful. *If we want to achieve our aim, then we need to focus on primary drivers, through secondary drivers, and one way to do that is through a change idea.* "When we were making the driver diagram, once I got the hang of

it, it made a lot of sense with the flow of things. Then there was the sentence stems that we could plug things into, and that really helped."

Jennifer also found the process of engaging in creating the driver diagram helpful for learning more about their topic. "I think it was really difficult to do the fishbone and the driver diagram, but then it became easier once we understood our topic more in depth. So it really seemed difficult from the outside, but then it became easy."

Colby shared how the driver diagram helped focus his improvement group's conversation; "I think it put them at ease, because teachers tend to want to go straight to the action to make the changes. [However] we had a lot of investment in the aim – building up to it. So we could continuously return to the aim." By continually refocusing the group on their aim they were able to identify a path through their driver diagram to guide their improvement efforts. Colby's group decided to focus on the 'teacher strategies' primary driver by implementing change ideas related to the secondary driver – teacher talk, response, and questioning. They decided to try out different Specifically Designed Academic Instruction in English (SDAIE) strategies and collect data about which strategies were leading to improved outcomes for their bilingual students (see Appendix B for the EL group's driver diagram).

Throughout the root cause analysis phase of the improvement science professional development, teachers felt that the improvement science tools – empathy interviews, fishbone diagrams, interrelationship digraphs, and driver diagrams – were useful for developing a deeper understanding of their focus problem, for determining which causes were most likely to be at the root of their problem, and for choosing a high leverage area on which to focus their improvement efforts.

How Teachers Use Data to Inform Their Practice

At the heart of improvement science is user-based inquiry, including collecting and analyzing data to inform next steps. The focus teachers used two categories of data to guide improvement efforts. Outcome measure data, including SBAC assessment results from the previous year and Youth Truth survey data from the previous and current year, was used both to launch improvement and to determine if their efforts over the year were resulting in improvement. Whereas, data for improvement was collected on a more frequent basis and used to determine if change ideas were leading to improvement. The four focus teachers all engaged in determining, collecting, and analyzing Plan Do Study Act (PDSA) data and also used the Measure of Academic Progress (MAP) assessment at three points during the year to track student growth related to literacy and math proficiency.

Outcome Measure Data: SBAC and Youth Truth

The improvement science professional development intervention was launched using SBAC and Youth Truth data from the previous academic year to identify possible areas for improvement. The SBAC is a computer based adaptive assessment tool aligned to the Common Core standards adopted by California in 2010. Students at all California schools that rely on public funding take the SBAC in third through eighth grade and again in eleventh grade. The test assesses grade level proficiency in mathematics and English language arts/literacy.

Youth Truth is a student perception survey of school climate factors including student engagement, perceptions of academic rigor, school culture, and incidence of bullying. The Youth Truth survey results are disaggregated by grade level and student demographics to provide insight into the collective experience of different student subgroups.

The focus teachers used a protocol to discuss the data and to surface areas for improvement during their first professional development session (see Appendix I for a looking at data protocol). Stephanie felt that the SBAC data from the year before clearly identified bilingual student support as an area for improvement. “That process leading up to the fishbone diagram I think was really useful. Looking at the data. We were looking at the SBAC scores, specifically students with special needs and then students who were English language learner (ELL) identified.” Over 60% of students categorized as ELL did not meet grade level standard on the SBAC, compared to 30% of students who identified as ‘white’.

While the focus teachers appreciated that the data suggested general areas for improvement, the lagging timeframe and aggregate nature of the data made it difficult to use the data to formulate change ideas for specific students or classroom specific interventions. Jennifer shared “I think it’s hard, because we’re looking at a lot of data. But then we don’t know what kids that data is from.”

The focus teachers felt that looking at data was informative and helped identify areas of improvement, however the aggregate and lagging nature of these yearly assessments was not useful for tracking improvement efforts done in classes over the course of the year. The surfaced need of a more timely assessment to track student growth over the year culminated in the school purchasing access to the MAP assessments in math and literacy for all students. The focus teachers planned to have students take it at the beginning of the year, at the beginning of second semester, and again at the end of the year.

Data for Improvement

The focus teachers used multiple forms of data to guide improvement efforts throughout the intervention. Each of them had their class take the MAP assessment at three points over the year, and also collected data on change ideas they were testing in their classrooms.

MAP data. Despite previously looking at SBAC data that indicated the majority of the schools students were not proficient in mathematics or reading during the first improvement science professional development session, the freshly collected MAP data from current students seemed to trigger a more visceral response from the focus teachers. Jennifer, in particular, felt taken aback by what the data said about her current students,“ …then we looked at our data and a majority of our school is below proficient. … But it kind of blew our mind how low the students actually are.”

The midyear (second) round of MAP data allowed teachers to determine which students had demonstrated growth as measured by the assessment. Stephanie appreciated the specificity of the MAP data as opposed to the aggregate data from the SBAC, “it personalized it a little bit more for my students and allowed me to contextualize it a bit more.” She could put names and faces to the MAP growth data.

Jennifer also appreciated the specificity of MAP data. “What I like about MAPs is that it breaks it down into categories, so it’s not just one overall score, but you see in the individual categories how they are doing.”

Zara, reflected on using MAP data with her current class of students, specifically ones with independent education plans (IEPs), “Because MAP will […] set goals for the students, especially students with IEPs, we were able to use those goals in their actual IEP.” She also felt it was useful to determine where the class as a whole could use more practice and to determine

next steps for instruction. Specifically, she found MAP data useful "to see where other students are struggling that aren't on the Special Ed caseload. … Informational text is something that most of the students are struggling with. Which is why I started the articles of the week."

Colby also found the MAP data useful and shared that, "The fact that the MAP scores were based on a student's growth was really helpful. I still had some room to grow for some of my ELs. [MAP scores] told me that they're improving, but I think that there are some where their growth could be more. Same thing with some of our higher level kids too. I noticed that some did not improve in as much of a jump as some others." In response to looking at the midyear MAP data, Colby shared that he rearranged his seating plan to move students who needed more support closer to and facing the front of the class. "It sounds small, but I think just not having to turn your body at all… it really does help some kids. And just surrounding them by other people that are going to be strong and helpful." In addition to modifying the seating arrangements, Colby also felt that the MAP data prompted him to connect with his bilingual students more frequently. "I think I just paid a lot more attention to some of our ELs individually."

The four focus teachers found the specificity of the MAP data helpful for making instructional decisions such as focusing more on foundational math concepts (Jennifer), or on information texts in their class (Zara), or to adjust seating arrangements (Colby). MAP data also provided an opportunity for the focus teachers to reflect on which students were meeting their math and literacy growth goals, and which students needed additional support. The focus teachers also found reason to celebrate when their midyear MAP data indicated that students were meeting their targeted growth goals in math and literacy.

PDSA level data. A key component to using improvement science to guide inquiry is using a Plan Do Study Act (PDSA) cycle to test out a change idea. Once a theory of action has been developed and a high leverage change idea chosen for implementation, a suitable measure needs to be selected or developed to determine if the change idea is resulting in improvement. All together, the teachers documented 35 PDSA cycles and collected two different types of data, perception data and performance data. Perception data measures people's beliefs, feelings, and thoughts about a particular topic and is collected from surveys, empathy interviews, exit slips, written reflections, and observations. Performance data included student work and assessment data from MAP tests. Perception data made up the bulk of data collected through PDSA cycles, 26 instances compared to 10 instances of performance data collection. Table 8 summarizes the PDSAs completed by improvement groups and Table 9 summarizes the types of data each group collected.

The EL improvement group engaged in the most PDSA cycles over the year, 22 compared to five or fewer for all the other improvement groups. A number of factors may have contributed to this disparity. One possible factor was the group size. The EL improvement group had six group members, whereas the next largest groups had four members. A second factor was that one of the group members was a graduate student spending the year shadowing the school director and chose to study the EL improvement group for her graduate work. She helped facilitate during their check-in meetings and may have provided an additional layer of organization and focus for the group. A third factor was the strong motivation from the group members to support their bilingual students. Stephanie shared, "Well, we were all really interested in academic vocabulary acquisition for our ELs and I think it just kind of happened that we all had that same passion. And then that drove all of our choices."

Table 8. Improvement Group Plan Do Study Act Cycles

Improvement Group	Number of PDSAs	Percent of Total PDSAs
Assessment	1	3%
Belonging	4	11%
English Learner	22	63%
Literacy	5	14%
Math	3	9%

Table 9. Improvement Group Plan Do Study Act Data Types

Data Type	Times collected	Percent of Total Data Type Collected
Perception Data (n=26)		
• *Surveys*	4	15%
• *Empathy interviews*	1	4%
• *Exit slips*	4	15%
• *Reflections*	2	8%
• *Observations*	15	58%
Performance Data (n=10)		
• *Student work*	9	90%
• *MAP data*	1	10%

How the focus teachers engaged with PDSA level data fell into two distinct subcategories: (1) trying change ideas and choosing what to measure and (2) engaging in iterative inquiry.

Trying change ideas and choosing what to measure. During the third improvement science professional development session, the four focus teachers and their

improvement groups brainstormed change ideas and planned for their first PDSA cycle. Stephanie and Colby each decided to choose a SDAIE strategy to implement in their classroom over the next week. Zara and her literacy improvement group decided to use ACTIVE reading journals (after reading a text students *A*sk questions, make *C*onnections, *T*rack down information, *I*nfer, *V*isualize, and record one *E*ureka moment of something they learned), and Jennifer chose to implement a lesson study focused on teaching foundational skills.

All of the focus teachers decided to try out a new structure or teaching practice. However, Jennifer and her math group struggled to implement their lesson study plan due to an inability to find coverage for the two teachers who would participate in the lesson observation. Change ideas that were smaller in scope and within the locus of the teacher's control were more likely to happen.

Another challenge was deciding what to measure, and the focus teachers wound up using a mix of perception data (surveys, observations) and performance data (student work samples, quiz scores) to determine if their change ideas were resulting in improvement throughout the year. According to Colby, the EL improvement group found it challenging to collectively decide on one SDAIE strategy to all test at first. "It was hard for us to pinpoint one change idea to do. So we started off each doing our own and just sharing about it and tweaking it for own classroom."

Colby's first change idea was to use Plickers to encourage his bilingual students to participate in class discussions. The Plickers app allows students to answer questions anonymously and displays class data so teachers can formatively assess student content knowledge. "It was nice because they [the Plickers app] got 100 percent of students responding, and the anonymity, I think made it a lot safer for all of them."

Colby also iterated on using word walls in his classroom to help students build academic vocabulary. Colby had each student choose a word from an upcoming unit and write out a student version of the definition and draw a picture to illustrate it. Then the class posted the words on the wall at the back of the class. Colby shared, "It was interesting to see some of the students, not just EL's, say "Oh!" Once we went over it in class. Like, "Oh, that was my word!" They were experts on it."

After informally noticing that students seemed to appreciate having prior exposure to vocabulary, Colby decided to move the word wall from the back wall onto the desks. When asked how he would measure whether this change idea was leading to improvement, Colby shared, "We were going to take a quiz. I haven't decided if I want to do a pre and post [quiz]. Probably because I haven't decided how that quiz is going to look." In the end, Colby decided against a formal quiz and used informal observations to determine if students were learning the unit vocabulary.

Unlike Colby, Stephanie used quiz scores to track her change idea of teaching different quiz study skills,

> Every two weeks I've been trying a different quiz study technique with the students, so we'll make note cards in class, or we'll go through a study guide together in class, or they wrote a quiz themselves one time. … at the beginning of the year, they were on average getting about a 7.2 to 7.8 on my 10 point quizzes and both classes have jumped up to 8.2, to like 8.6. I think it's based on that change because that's the only thing I've changed.

When asked which study technique seemed to result in better EL student performance data, Stephanie clarified that she wasn't formally tracking each set of quiz scores, but that on average the whole class quiz results were better near the end of the semester than they had been at the beginning, "I didn't think about taking data from every single study technique."

Zara struggled to engage in PDSA cycles. One member of her literacy improvement group was already using ACTIVE reading journals and suggested the rest of the group try them as well. Zara explained, "She was already using them with her students so we decided not to reinvent the wheel because all of us are already stressed and she's been using it for a few years. But it's new to the rest of the literacy group." However, adoption of ACTIVE reading journals required thought and time to determine the best way to incorporate them into the classroom. When asked about how she might collect data to determine if ACTIVE reading journals were improving her students' reading comprehension Zara replied, "I haven't figured that out yet. I haven't figured out how I'm going to actually grade the ACTIVE journals yet."

These data illuminate some of the challenges experienced by teachers as they determined which change idea to try and what data to collect to determine if it was an improvement. Colby found success with the Plickers app change idea, but struggled with how to formally determine if the word wall was ultimately impacting student learning. Zara, a newer teacher, clearly felt overwhelmed and adopted a practice that she wasn't familiar with and struggled with how to determine if it was improving student literacy or not. Stephanie was invested in her students' academic language acquisition and tested multiple quiz study strategies, but struggled to determine which of the strategies was most effective for student learning.

Collecting and analyzing data for iteration. Throughout the improvement science professional development, the four focus teachers engaged in data collection and analysis, both individually and collaboratively with their improvement groups, to evaluate if a change idea was leading to improvement. While three of the four focus teachers engaged in active iteration of a change idea and eventually adopted a new teaching practice, the primary form of data collected throughout this inquiry process was observational data and anecdotal student feedback.

Colby and Stephanie's EL improvement group chose to analyze data collaboratively. Over the first semester, the group decided to try a different SDAIE strategy each week, using the weekly morning meetings to share data and discuss whether they wanted to adopt, adapt or abandon their respective change ideas. Colby recalls, "In our meetings all first semester we were pretty on top of having data and bringing it to the meetings." However, over time, the group decided that they might learn more about a particular change idea if they all tried the same one. Colby had tried a word wall, so the group decided to try word walls as well.

The intent of improvement science is to iteratively refine a change idea based on whether data collected indicates an improvement is occurring. Colby shared how word walls evolved over time to become part of his teaching routine.

> It started off by just having the kids create vocab words, the definition, an example, and a drawing that I just posted up on the wall for one project. Then for another project I improved on that by putting them around the tables. I don't think students were actually reading them, they were just there. Then the next project I put them around the tables, but before even doing that, I had the kids do an anticipatory set. Not a quiz, but just – try your best to explain what these words mean – and then we returned to it after. So that's where I actually found that it was useful, because kids weren't just doing an assignment. I could actually see them trying to understand the words initially, and then the 'a-ha' moments when we went over it again later on in the project.

Colby highlighted that since the words were new to most students, doing the anticipatory set seemed to level the vocabulary playing field (see Appendix K for Colby's final anticipatory worksheet). "Especially with our ELs, because they could see other people struggling right off the bat. And it was okay to make a mistake."

Colby used observational data throughout this iterative process, but did not formally track vocabulary performance data on a per student basis. For example, Colby's reasoning for moving the word wall to the tables was based off an observation, "I put it almost eye-level for these tables [tables at the back of the class near the wall], so these people would get to see them

frequently. But it only gets half the class. So, the thing we're doing to adapt is to put them on the tables. So everyone in our [EL] group is trying that."

Student performance data in the form of test scores was used to inform how a particular change idea worked *after the fact*, but not necessarily as a PDSA measure, where data is used to make a decision about what to try next. Colby found that the time frame of each class project did not lend itself to quick iterations. "I think gathering data in a timely manner is the tough part." He planned to use test data to determine if moving the vocabulary words from the back wall to the tables where students sat would help his bilingual students learn the vocabulary. "For example, doing the vocab. I would plan it and implement it over the course of a project. By the time I get that [test] data I've already begun planning the next thing."

Colby also noted that variation in the length of time students were exposed to the vocabulary might have impacted student vocabulary acquisition, highlighting some of the difficulties in isolating specific factors contributing to student outcomes in a classroom environment. "We took the final exam leading up to exhibition, so turning the room around they didn't have the words on their table for very long. And some of the folks, especially ELs who I was trying to target did not do so well."

By the time he had graded the test Colby had already decided on using the anticipatory word set at the beginning of his next project, so the assessment data was not used to inform the next change idea iteration. Colby found observations of student interactions with the new anticipatory word set promising and decided to adopt the practice:

> I think that's a much bigger impact – doing the anticipatory set. I had them write down, no matter what their experience with the word, what they think it is – even just from how it sounds. And already today I could see it kind of working with some folks. Cause some of the words came up in discussion, [and students were] like "oh, that's what crash means, okay.

Despite the lack of more systematic tracking of student vocabulary development, the observational data of how students interacted with the word wall and ultimately how they engaged with the anticipatory vocabulary set, resulted in Colby adopting the anticipatory set as a new teaching strategy.

Stephanie iterated on word walls along with Colby and the other members of her improvement group. She used both student quiz scores and informal classroom observations to determine how the practice was working. "I've found that it helps the students access the academic vocabulary significantly better. They did significantly better on the quizzes than I would have anticipated them doing without the word wall." Stephanie used the weekly morning meeting structure to analyze the quiz scores with her group, "I had three quizzes that had some element of vocabulary tied to them. I was physically looking at the quizzes with the group."

When the improvement group decided to move the words to the desks, Stephanie tried the iteration as well (see Appendix K for Stephanie's word wall template). When asked if it improved student vocabulary acquisition she said, "I feel like it shifted it to their immediate focus because they are looking at their desks more than this list on the whiteboard, but as far as measurable differences, not that I can say. But observationally I've noticed that students are using the words a lot more. That could have been because a variety of different things, but I think the words on the table helped too." Stephanie acknowledged that she was not using student performance data to determine whether moving the word wall to the desks impacted student vocabulary acquisition, but instead was relying on informal observation data.

Jennifer also used observational data to determine how her change ideas were impacting student mindsets around math. In an effort to shift students' perceptions of math as a subject primarily concerned with right answers, procedural fluency, and speed, Jennifer began starting

some of her lessons with an *agency warm-up*. Students look at a pattern, graph, or set of numbers with no question attached, then spend five minutes individually writing down questions, wonderings, ideas, and problems that come to mind followed by a five-minute student-facilitated whole class discussion. Jennifer was hoping that promoting open ended math exploration through the agency warm-ups would help students see math as an open, exploratory subject and encourage more participation in whole class discussions.

Jennifer relied on observational data to determine how the practice was impacting her student's engagement in math discussions. "I see students writing a lot more, but not sharing." When asked if she was tracking who was sharing, she said, "I'm not like formally tracking, but like informally I am. And most of the time it's the main four or five kids. So I've tried to make some of those kids the leaders so it takes them out of the discussion, and a lot of times the discussion just flops. One time I timed it and for a minute and seven seconds they didn't say a single thing."

Students also shared what they liked about the practice directly with Jennifer, "I've had a couple come up to me and say they really like the [individual] think time. And I never really thought that, because when you give an assignment they immediately jump to wanting to work with a partner or wanting to talk to someone. But I think it might be a defense mechanism of 'I don't really know what to do, so I'm going to talk to the next person.' So I think that's been a really big takeaway."

Informal classroom observations and student feedback formed the bulk of Jennifer's PDSA data collection. Prompted by student feedback and her observations Jennifer incorporated more independent think time into her classroom activities.

While Zara used observational data and student feedback to determine how her students' literacy skills were progressing, she didn't iterate on any one particular change idea. In response to MAP data that indicated her students struggled with informational texts she started having them read and analyze an article every week. Zara relied on a mix of MAP data and conversations with students to determine if they felt that their literacy skills were improving. "I love MAPs results and how we can use them. But not only that, with the ELs you have to be able to actually talk to them and see if they're confident in their improvement."

While Zara used observational data and student feedback to determine how her students' literacy skills were progressing, the other three focus teachers used a combination of anecdotal and observational data, student feedback, and student work or assessment data to inform whether to adopt, adapt, or abandon a particular change idea. Jessica relied on observation and conversations with students to evaluate her agency warm-ups, and Colby and Stephanie each iterated on and ultimately adopted word walls to support their bilingual students' vocabulary development. None of the focus teachers utilized a more formal method of data collection or analysis.

Structures that Support Improvement Science Implementation

The four focus teachers reported that four structures helped them engage with learning about and using the tools of improvement science: (1) using protocols, (2) the improvement teams, (3) regularly scheduled meeting times and (4) internal and external improvement capacity and support.

Protocols

Protocols were used to facilitate teachers engaging with three improvement science tools – fishbone diagrams, interrelationship digraphs, and driver diagrams (Appendices C, D & E).

Three of the four focus teachers discussed how using these tools in the improvement science professional development sessions helped them gain clarity about the root causes of their problem of practice, identify high leverage areas – or drivers – to focus improvement efforts, and to generate promising change ideas (see earlier sections of this chapter on fishbone diagrams, interrelationship digraphs, and driver diagrams).

Stephanie discussed how the protocols specifically helped guide her thinking, "I think having the printed out protocols really helped me and then being able to follow that protocol, allowed me to understand what it was that needed to be done so I could kind of tangle with it. It also really helped me when I can brainstorm on sticky notes and then collect them together."

Improvement Teams

All four focus teachers felt that belonging to an improvement group helped them focus on their improvement goals. Colby shared, "I don't think I ever made a concerted effort to do strategies for ELs compared to this year." He felt that the regular meeting times with his group contributed, "I think the fact that it's been our conversation piece every week has made me think about our ELs a lot more."

Stephanie also felt that the improvement group increased her focus on her bilingual students, specifically students who were identified by lower scores on the California English Level Development Test (CELDT):

> Not only do I have more ELs than I've ever had before in one class, but my ELs that I have are at a much lower CELDT level than the ELs that I've had in the past. The EL group has been a really good place to bounce ideas off each other, and it has forced me to be more mindful of who is an English language learner, and who does need that extra support.

Zara appreciated the opportunities to collaborate with other teachers and hear what they were testing and Jennifer appreciated the sense of community that working with her group

brought, “I love working with our group. Working with them has made me feel more community at [school] than anything.” She also reflected on a strong desire to contribute to overall school improvement efforts, “I felt like I had a sense of responsibility to get my piece to work, because it could impact other groups.”

Additional factors that affected the teacher experience with improvement science was the size of the improvement groups. Stephanie, one of the EL improvement group members shared, “Our group is a little bit large, there’s about seven or eight people, and so sometimes voices get lost in that.” In contrast, Jennifer, one of the Math improvement group members, shared, “because it’s just three of us, you kind of – you have a voice.”

Regular Meeting Times

Focus teachers participated in five biweekly two hour workshops, and also met with their improvement groups weekly for 30-minute check-in meetings for the first semester and biweekly 30-minute meetings over the second semester. A final one hour workshop at the end of the academic year served as the culminating celebration of their improvement work (see Table 3). Meeting times supported the spread of change ideas and the development of an improvement community.

All four focus teachers found the regular meeting times useful and felt they provided structure to the improvement work. Zara shared, “I guess, just the opportunity to hear ideas from all of my peers and then be able to practice those ideas and then come back and talk about them. That’s probably been the most beneficial thing so far.”

Stephanie also appreciated the regular time to discuss ideas and data with her improvement group, “just the time to sit and digest things with my group,” and felt the meeting

times were useful, “the PD times, the half day Wednesday PDs are really helpful. And we have the morning meeting times to work with our improvement science group as well.”

Internal and External Capacity and Support

In addition to the protocols, regular meeting times, and improvement groups, the four focus teachers mentioned several other structures supported improvement efforts over the year: (1) administrator support, (2) the School Improvement Coach program, (3) the Graduate School of Education resident program, and (4) the Induction program.

Administrator support. The school director was instrumental in allocating staff meeting time for the improvement science professional development, as well as for the weekly improvement group meetings. She also participated in the assessment improvement group and in the improvement science professional development planning sessions. In one interview, Stephanie reflected on which structures supported improvement work across the school, “I think it goes back to the meetings and the discussions that we’re able to have, and the support that we have from our admin.”

School Improvement Coach program. Two of the teachers from the school had attended a one day improvement science training as part of a School Improvement Coach initiative. During the training the teachers conducted empathy interviews and used protocols to construct a fishbone diagram, an interrelationship digraph, and a driver diagram around a problem of practice. The teachers who attended the School Improvement Coach training day became co-facilitators of the two hour improvement science professional development sessions and helped improvement groups work through the protocols.

Stephanie shared, “I think seeing the example and then having a knowledgeable person at our table who seemed to have been working with you ahead of time, that was really helpful.”

Graduate School of Education School Resident Program. A graduate student provided additional improvement science capacity. For her master's program she was shadowing the school director and using an improvement science framework for her master's thesis. She joined the EL improvement group with Stephanie and Colby. Stephanie shared that the resident "has been really helpful with her role in the GSE [graduate school of education], and she brings back a lot of the resources from her classes."

Induction program. Both Zara and Jennifer were also clearing their teaching credentials through an Induction program that also used an improvement science framework for a final project. Zara reflected, "Not only am I in the improvement science group here but both [Jennifer] and I are in an improvement group for induction. I think having [Jennifer] to bounce ideas off of and also to do the tasks together really helps me."

From these data it is clear that multiple support structures helped teachers engage with the tools and methods of improvement science. Protocols scaffolded how to use the tools, collaboration with improvement team members around a shared aim provided a sense of accountability and shared purpose, regular meetings provided time to look at data and decide next steps, and administrators and others in leadership positions provided support and additional improvement science capacity to facilitate the work.

Changes in Teacher Agency Over Time

Over the course of the intervention three categories related to teacher agency emerged. Related to the belief that they can meet the learning needs of all their students, the focus teachers (1) reported an increased intentionality around their improvement goals and (2) participated in adopting and spreading new teaching practices. The focus teachers also (3) reported feeling like

they were part of an improvement community and that they developed a deeper understanding of the problems they identified for improvement.

Intentionality, Inquiry, and Changes to Individual Practice

Stephanie and Colby reported that engaging in the inquiry process – the PDSA cycles – helped them focus on their bilingual students and both teachers adopted some of the SDAIE strategies they tested over the year. Stephanie recounted which strategies were becoming routine in her practice, "Word banks, some quizzes, sentence starters, which I've done before, and just, like, the physical layout of quizzes. Like making sure I put boxes and lines where I want the kiddos to fill in the blanks. I've also been intentional with building keywords into questions."

Colby also felt that new practices were now becoming routine, "I finally turned some of my project planning into a routine with vocab for EL students." Colby continued, "I have always, every single year, been telling myself to do word walls more often. Improvement science kind of kept that at the front of my mind. It was usually an afterthought, doing a word wall, but doing it throughout the year it's now become part of the beginning of planning for me." When asked what structures contributed to him iterating on and adopting the anticipatory word wall strategy he shared, "I think with the testing [PDSAs] and being in the EL group has just made me think about planning projects and assignments. To make sure I'm keeping the learning more at the front of the project, rather than just focusing on the product we're creating."

Stephanie also shared that the PDSA process was instrumental in helping her maintain a focus on her bilingual students, "the accountability of the PDSA surveys that we have to fill out every week. That has helped me say "oh, well what am I doing?" And it's forced me to look at the data a little bit more."

During second semester, the school assessment group helped the focus teachers analyze shifts in Youth Truth and MAP assessment data. Bilingual students reported an increased sense of belonging as compared to the previous year's Youth Truth data with 100% responding favorably to the question: *I really feel like a part of my school community*, and many teachers found that their bilingual students were showing growth on the MAP literacy and math assessments as well. When asked what he thought contributed to this growth, Colby said, "I think it's just the fact that we made a group, put it more in the forefront of people's minds to do something about our ELs and that we try to be a little more explicit in what we're doing in the classroom."

Colby's increased focus included other aspects of class in addition to the word walls. He also made attempts to connect around language with his bilingual students:

> That and just forcing myself personally to try to connect a little bit more with some of our ELs, especially in Spanish. I had a really fun day, maybe three or four weeks ago, in office hours when some of my ELs were here. We were speaking Spanish the whole time and I was teaching math in Spanish and just looking to them to help correct me. It made them try to pay attention more to understand. I mean it was office hours so we're talking about five kids, but it was really productive. I think it made the ELs feel really valued and they even begged me the next day to try to start class that way. So I did that too.

Stephanie also incorporated more student specific changes, when talking about one student in particular she shared, "she told me it really helps me when you come to me and you retell me what the directions are after you've told the whole class. And so I've been doing that and she's been doing a whole lot better."

Based on evidence from classroom observations, student interactions, and student quiz data, both Stephanie and Colby incorporated practices they had tested out over the course of year into their planning routine for new units/projects. Both believed that

through these practices they had the ability to improve learning outcomes for their bilingual students.

Sharing and Spreading Change Ideas

The regular meeting times and whole school focus on improvement science also helped the spread of change ideas from teachers in one improvement group to another. After one morning meeting when focus teachers and their improvement groups shared what they were trying with each other, Colby shared, "I think everybody agreed that it was great to connect with each other and just share good practices."

Jennifer also appreciated hearing about what change ideas other groups were testing. "I liked the share out today. We're all doing this improvement stuff, but this is the first time we've come together and really talked about what we're doing and how [it] can be interconnected." When asked what change idea from another group she wanted to test she said, "The sentence starters. The EL sentence starters. But I also like the Word Wall which is also from the EL group." Observational data from follow-up classroom visits confirmed that she implemented both change ideas.

Stephanie decided to adopt one of the change ideas from the belonging improvement group, "I appreciated what we did this morning with the speed dating, hearing other people's change ideas." She decided to create an advisory playlist of her students' favorite songs.

Later in the year, the EL improvement group also shared SBAC test taking strategies for supporting bilingual students. Zara who was planning to have her students practice with one of the released SBAC performance tasks adopted one of the shared strategies:

> It was suggested that I show them [her students] the instructions for the performance task and have them do a critique session on the instructions about what would help them better understand them. I think I'm actually going to

> tomorrow. I'll give them the actual instructions of a real performance task and have them break it down.

One change idea in particular became standard across the school. Over the year word walls became common practice in most classrooms.

Identification as an Improver in an Improvement Community

Over the course of the intervention, the focus teachers reported that the regular meeting times with their improvement groups and the three cross-group gatherings helped them feel like they were part of a larger improvement community.

Jennifer felt that the regular meeting times helped provide time to engage in improvement work. "I think having the PD, and then having that built in time to meet with your group after PD, and then a week or two weeks to implement something, that built in time, I think enables a lot more work to be done."

She also felt that engaging in improvement science had the potential to benefit more than just her classroom and that she felt like part of an improvement community. "I'm actually doing something that's going to benefit my classroom, that's going to benefit the school, the students." One of the communal morning meetings where the teachers gathered and improvement groups shared what change ideas they were testing, stood out to her in particular:

> One thing I really liked is in our meeting this morning, we all shared out as groups what we're doing so everybody else kind of knew. And I felt like that was one of the first times that we were all working on something to improve our school, but each of us were doing something different. So it gave that – I don't know – that sense that we were all trying to fix a piece to make it better. And I don't feel like I've ever really felt that before.

Colby also appreciated the fact that the whole school was engaged in improvement work. "It's something that we're doing as an entire school which makes it really helpful."

The whole school focus also supported teachers trying out other improvement group change ideas. Colby shared that members of his improvement group were "even participating in other people's groups."

Stephanie felt that engaging in the improvement science professional development was shifting the way she thought about issues in her classroom. "I do think, having gone through the improvement science process for a couple of months now, it's changing the way I'm thinking about other types of problems in the classroom." She said that data played an important role in that shift. "I've owned what's going on and I've owned the data and the way to improve the data. I think that ownership has allowed me to feel like a member of that [improvement] community."

Stephanie also shared that conversations about improvement work had started occurring outside of regularly scheduled times including in the teacher lounge or 'Hub'. "Now we check in a little bit more in the Hub, just casually."

The focus teachers valued the increased intentionality that improvement groups provided, and also valued the chance to work on a problem together. Stephanie shared, "I think if we can take what we've learned and for one semester or a whole year focus on one thing, then our whole school is going to make meaningful changes." Over the course of the year, strong ties developed between the EL support group and the belonging group since both were focused on supporting emerging bilingual students. Throughout the year strategies from the EL and the belonging groups were adopted and adapted by the majority of the teachers.

Stephanie and Colby also celebrated when the belonging improvement group reported that the yearly Youth Truth survey data indicated that they were very close to meeting the belonging group improvement aim: *By June 2017, 100% of students will feel an improved*

connection to the school. The survey indicated that 100% of emerging bilingual students responded positively to the survey measure: *I really feel like a part of my school community.*

A growing focus on supporting emerging bilinguals crossed improvement groups and informal conversations started about how to structure improvement work the following school year. By the end of the year, the teachers were coalescing around supporting their emerging bilingual students and launched improvement work the following school year around that singular focus.

Summary

These data describe how teachers engage with the tools and methods of improvement science and the impact of the improvement science professional development intervention on teacher agency. Quantitative pre- and post-survey data, while not statistically significant, indicated positive shifts in teacher agency over the course of the intervention in the following three ways, (1) an increased belief that they could meet the learning needs of all their students, (2) an increase in the perception of belonging to an improvement community, and (3) an increase in the degree to which they valued the professional development opportunities provided by their school.

Qualitative data from the four focus teachers, including interview data, documents related to the improvement science work, and classroom observation data, explored how they engaged with improvement science tools and methods. From these data four themes emerged, (1) using improvement tools to define the problem, (2) using data to inform practice, (3) structures that support improvement work, and (4) changes in teacher agency over time. Analysis of the themes and categories that emerged from the data provide insight into the changes that can occur in teachers' practice and a school culture through improvement science professional development.

CHAPTER 5: DISCUSSSION

Overview of the Study

This study explored how teachers experience improvement science professional development and what impact it had on their sense of agency. In this study, teacher agency was conceptualized as (1) a sense of belonging to an improvement community, (2) a confidence that they could meet the learning needs of all their students, and (3) a belief that the tools and methods explored in the improvement science professional development were useful and relevant for improving practice.

To determine how teacher agency was affected by engaging in improvement science professional development the following research questions were posed:

1. What is the impact of improvement science professional development on teacher agency?
    a. To what degree does the intervention improve teachers' sense of belonging to an improvement community?
    b. To what degree does the intervention improve teachers' sense of confidence in their ability to meet the learning needs of their students?
    c. To what degree does the intervention improve teachers' perceived value of professional development supports?
2. How do teachers engage with improvement science tools and methods?
    a. How do teachers use improvement science tools and methods?
    b. What structures support teachers in engaging in improvement work?
    c. In what ways do improvement science tools and methods impact teacher agency?

To explore these questions teachers participated in a series of improvement science professional development sessions that supported the creation and work of five improvement

groups. Both quantitative and qualitative data were collected, including pre- and post-intervention survey data and interview, classroom observation, and improvement science documentation data from four focus teachers. The quantitative pre- and post-intervention survey data was analyzed to determine the percentage of teachers responding 'agree' or 'strongly agree' on a five point Likert scale. Interview transcripts and observation notes were open coded in an iterative process to uncover significant themes (Esterberg, 2002; Saldaña, 2009). After an initial round of open coding, a second round of coding was completed using the research sub-questions as a lens to group existing codes and identify emerging themes.

Discussion of Findings

This study yielded important findings about how teachers engage with the tools and methods of improvement science and the impact of such engagement on their sense of agency. The following sections will discuss the affordances and challenges teachers experienced using improvement science tools and methods in their context and how these finding relate to the three identified areas of need for effective bottom-up reform: (1) the need to foster teacher inquiry, (2) the need to promote teacher collaboration, and (3) the need to build a shared professional knowledge base.

Structures that Promote Teacher Inquiry

From interview data as well as pre- and post-intervention survey data it was clear that teachers embraced improvement science tools and methods and were willing to engage in inquiry to guide improvement. Three key structures supported teacher inquiry, including learning about the tools and methods during the six improvement science professional development sessions, formation of improvement groups, and regular meeting times.

Teachers identified protocols as a key factor that supported them simultaneously learning and using the tools of improvement science – empathy interviews, fishbone diagrams, interrelationship digraphs, and driver diagrams – throughout the six improvement science professional development sessions. However, protocols are only useful if they are used as intended. To ensure that protocols served the purpose of moving work forward and having all group members equitably participate, a teacher-leader facilitator joined each improvement group who had used the protocol before. This was an intentional capacity building decision in order to foster teacher ownership of the improvement tools and methods.

One of the hopes for improvement science is that the tools and methods can be used by teachers to engage in inquiry to achieve more equitable outcomes for all their students. In order to achieve this goal, teachers need to first recognize that the tools are useful to their practice. Using the fishbone diagram, interrelationship digraph, and driver diagram tools in the initial professional development workshops provided teachers with a shared experience in which they deepened their own understanding of their identified problems of practice. All four focus teachers reported the tools were useful and helped them learn more about their problems of practice.

Additional structures that built capacity and ownership of the improvement work include the formation of improvement groups and protected time to meet together. The groups met regularly to discuss change ideas, debrief the PDSA process, and to look at data. Focus teachers each felt a sense of community and accountability to their groups and appreciated the time set aside for these purposes. The regular half hour check-ins reinforced the iterative inquiry cycle and provided teachers with a community to help with data analysis and planning next steps.

These structures were instrumental in supporting teachers as they engaged with learning and using the tools of improvement science over the course of the intervention.

Challenges Teachers Experienced Engaging in Inquiry

While the structures listed above are key takeaways for educators interested in developing an improvement culture at their school site, teachers also experienced challenges with using the tools and methods of improvement science. Teachers easily identified change ideas to try in their classrooms; however, they struggled with how to measure if what they were trying was resulting in improvement and what data to collect to engage in iterative cycles of inquiry.

For example, key insights that led to iteration in Colby's classroom came from anecdotal observations that provided insight into how his students were interacting with the word wall iterations. Colby was not alone, teachers overwhelmingly used perception data to determine if a change idea was working – 26 instances of collecting perception PDSA data compared to 10 instances of performance PDSA data, and of the different forms of perception data, informal observations were the most common – 15 instances as compared to four or fewer for the remaining types of perception data (Table 10). Exit slips and surveys assessing student or teacher opinions were also common. The preferential use of these types of data could be due to the fact that many teachers are already in the habit of using classroom observations and exit slip data to guide their instruction.

Focus teachers also struggled to engage in iterative inquiry. Over the course of the improvement science professional development Stephanie tracked student quiz scores as she tried out different vocabulary studying techniques with her students, however, she did not systematically track individual student scores or iterate on any one particular study technique.

While the class average quiz scores increased over the year, Stephanie found it difficult to determine which quiz study technique was most effective, if any.

In Colby's case, timing also impacted his ability to iterate based on previous data. He found it challenging to collect and analyze data on student vocabulary knowledge from end of unit tests before planning his next project. Instead, he relied on informal classroom observations to guide his word wall iterations. And while the final iteration – the anticipatory set – aligns with research on how predictions prime learning opportunities (den Ouden, Kok, & de Lange, 2012), he didn't collect any student performance evidence to bolster his claims that using an anticipatory set for vocabulary was a better strategy than having students contribute one word to a static word wall.

Zara who was new to project based learning felt that improvement science was an 'extra' thing to attend to, and while she appreciated the collaborative aspects of improvement science, such as sharing practices and looking at whole school data, she struggled to engage with the iterative PDSA process. These findings suggests that improvement science is not immune to the initiative fatigue felt by many teachers, and has the potential to feel like just another 'top-down' reform effort.

A concept that may have helped teachers address these challenges is the idea of *standard work.* In many professions, portions of the work are standardized to reduce variation. Doctors have checklists for routine procedures to ensure that a high level standard of care is met. A set routine in a teacher's classroom that supports an already identified student behavior may be a more accessible target for iteration compared to testing out a new teaching practice. The desired student behavior provides an easily identifiable measure that teachers can collect data on as they iterate on the already established routine.

In some cases, the nature of the problem also causes challenges to collecting useful data. Developing student literacy, mathematical agency, or vocabulary acquisition are all complex processes with a myriad of interconnected variables. In addition, developing neural connections – the foundation of learning – is a biological process that takes time to occur. When introducing a new practice, the layered nature of the biological learning process – it takes 10-15 instances of seeing a word in various contexts before students fully grasp it's meaning – can make it difficult to measure change over a short (weekly/biweekly) time-frame (Lemoine, Levy, & Hutchison, 1993; Francis, & Rivera, 2006). It may be the case that one or two targeted practices have only a small effect on student learning outcomes over the weekly/biweekly PDSA timeframes used by teachers in this study, but that the cumulative effect of repeated use of particular change ideas may lead to improved student learning outcomes. The increases in MAP data over the year and the steady increase in average quiz scores in Stephanie's classroom provide some evidence for this hypothesis.

Recognizing the significant cognitive load that a full day of teaching already demands, where does the cognitive work of data analysis fit in? What data do teachers find useful? And what data collection methods build on teachers' already powerful classroom observational skills? These findings suggest that there is additional work needed to support teachers in developing inquiry methods that they find useful and that fit with the day-to-day classroom workflow.

Despite these struggles with iterative inquiry, there was evidence of changing ideas about how to best collect and analyze data over the course of the intervention. After two months of testing SDAIE strategies individually, members of the EL improvement group decided to all try the same change idea in order to collect and analyze data collectively from across their different classroom contexts. Group members all moved word walls to the tabletops so they might be

more visible to students and decided to use test and quiz scores to determine whether this change was effective. This shift from thinking about data from their individual classrooms to thinking about data from across their classrooms in order to collectively learn from variation, suggests that teachers were beginning to think more systematically; moving beyond the silos of their classrooms and embracing the idea of learning from each other. This also suggests that aspects of teacher agency, such as the feeling of belonging to an improvement community, were strengthening too.

Finally, despite the challenges of effectively collecting robust PDSA level data to guide iterative inquiry, teachers overwhelmingly decided that the methods and tools were useful and decided to use improvement science to guide school wide improvement efforts again the following year.

Structures that Promote Collaboration

Regular meeting times and protocols also played a role in helping teachers develop a common vision of improvement over the course of the intervention. Using protocols to guide the work of improvement groups early in the semester supported them in developing a shared understanding of their problem and a theory of action to guide improvement efforts.

Teachers felt that the regular meeting times and the improvement group structure created a sense of accountability and shared purpose, and all of the focus teachers felt that being part of an improvement group was instrumental in them trying out teaching strategies that they may not have tried otherwise. In addition, Zara, Stephanie, Colby, and Jennifer all appreciated the opportunity to hear about other groups' improvement efforts, and felt that meetings where all teachers were present facilitated the spread of change ideas from one group to another. Teachers from different improvement groups incorporated versions of word walls into their projects and

by the end of the year word walls were present in almost every classroom. Further evidence of deepened collaboration included the whole school improvement focus on testing pedagogical strategies to support emerging bilingual students in the following year.

Challenges to Collaboration

Challenges teachers faced around collaborating included change ideas that required significant adjustments to the system such as the math improvement group lesson study effort. The school schedule included some flexibility due to a team teaching model – two core teachers shared 50 students – however, to take advantage of this built in flexibility required teachers to ask their teaching partners to cover their classes while they participated in the lesson observation and debrief session. Most teachers felt it was too large a request to impose on their teaching partners and so the math group struggled to find a time that would work for all group members to try their change idea. In this study smaller change ideas within a teachers locus of control were more likely to occur.

However, lesson study, with its emphasis on understanding student thinking and ability to make "various types of knowledge more visible, such as colleagues ideas about pedagogy" (Lewis, Perry, & Hurd, 2009) could be a powerful structure for increasing teacher collaboration and learning (Stigler & Hiebert, 1999; Rock & Wilson, 2005). In Japan, lesson study already operates as a vehicle for continuous improvement efforts (Lewis & Tsuchida; 1998) and for U.S. schools with existing lesson study structures, integrating improvement science tools and methods may increase opportunities for teachers to learn from each other.

Generating a Shared Knowledge Base

Generating a shared knowledge base and common vision of teaching and learning is crucial if teachers are to be able to clearly define the boundaries of their profession. In this study

the tools of improvement science provided opportunities for teachers to engage in dialog to more deeply understand the problem of practice, the root causes contributing to it, and each other's ideas about teaching and learning. In addition to the development of a school wide aim to better support emerging bilingual students, several teaching strategies were tested and ultimately adopted by multiple teachers, including word walls, sentence starters, and graphic organizers.

If teachers are to be the generators of a robust professional knowledge base, we must also support teacher agency. Teachers who feel that they are an integral part of an improvement community and that they can meet the learning needs of all their students are more likely to iterate on new and existing teaching strategies and be willing to collaborate and learn together. Both quantitative and qualitative data from this study showed increases in teacher agency over the course of the intervention. With respect to the belief that they could meet the learning needs of all their students, teachers reported increased intentionality with respect to their improvement focus and engaged in testing and adopting different teaching strategies over the course of the year. Teachers also reported feeling like part of an improvement community and felt that improvement science provided useful tools for developing a deeper understanding of persistent problems of practice.

The successful spread of change ideas to support emerging bilingual students, evidence of growth in student achievement from school wide MAP data, and Youth Truth data indicating that the belonging improvement group was close to meeting their aim, all contributed to teachers feeling like using improvement science tools and methods helped them develop teaching practices that resulted in better outcomes for their students.

Challenges to Generating a Shared Knowledge Base

Challenges teachers experienced in generating a shared knowledge base included limited generation of sharable evidence of change ideas working and the continued silo nature of classrooms.

Throughout the study teachers relied on sharing about change ideas verbally during their weekly meetings and did not have the opportunity to visit each other's classrooms to observe colleagues testing out change ideas. While teachers valued this form of information sharing, the benefits of seeing a practice in action is that the shared experience makes visible the nuances of practice that are often not included in a verbal retelling. In addition, the shared experience can serve as a foundation for building a common vision of teaching and learning. Without opportunities to see what change ideas looked like in practice across the classrooms, teachers may have missed opportunities to discuss the pedagogical 'moves' required to optimize a particular practice.

Teachers also struggled to collect sharable data, such as student work, in support of a particular change idea. Without evidence to support anecdotal teacher observations, some of the iterations – Colby's anticipatory set for example – did not spread as widely as the earlier word wall iterations.

These challenges underscore the complexity involved in developing a shared knowledge base for teaching and learning. In order to develop common practices, a shared vision for what good teaching and learning looks like must first be developed. This requires teachers to discuss common problems of teaching, their possible root causes, and what student learning looks and sounds like within a given context. Improvement science tools and methods can help facilitate

this process, and may be enhanced if paired with other structures that support additional shared experiences such as lesson study.

Implications for Improvement Science as a Professional Development Structure

Fostering an improvement culture has many benefits in an educational setting. Focusing on systematic inquiry and collaboration around teacher defined problems of practice has the potential to significantly impact teacher agency and ultimately the learning and wellbeing of their students. Findings from this study detail teacher experiences with improvement science tools and methods and provide insight into the structures that support improvement work as well as potential challenges to adopting an improvement framework.

Structures that supported teachers as they engaged in improvement science include administrative support, regular meeting times, protocols, and colleagues with improvement capacity to co-facilitate professional development sessions. Challenges to using improvement science tools and methods included deciding what data to collect to determine if a change idea was leading to improvement, and engaging in iterative inquiry to refine specific practices.

Supporting Improvement Science Professional Development

Complex systems are often resistant to change (Lorenzi, & Riley, 2000) and bringing about a cultural shift in a school requires attending to multiple aspects of the system. In this study transformational leadership was instrumental in creating a shared vision between the director and the teachers (for a review of transformational leadership see Anderson, 2017). Two teachers who were interested in co-facilitating the improvement science professional development sessions participated in a one-day workshop to build familiarity with the tools and protocols. By building capacity for improvement work, teachers were able to feel ownership over

the process from the beginning and were instrumental in conveying that all teachers were equally capable of learning and using the tools of improvement science.

In this study it is also important to consider the affordances of the pre-existing school culture in how teachers engaged with the tools and methods of improvement science and how these might be different or similar to other educational contexts. Evidence from the pre-intervention survey indicated that teachers felt a strong sense of agency to begin with, and a strong culture of teacher autonomy over project design likely contributed to teachers feeling comfortable trying new change ideas. In addition, the school did not focus on standardized test scores as the ultimate arbiter of teacher performance or student achievement, and instead used a mix of student perception data from the Youth Truth surveys and student exhibition of project work to gauge school success. However, the high teacher autonomy may have also limited the degree to which teachers were able to share specific iterations on change ideas. The variation in teaching structures present across the schools may have prevented more nuanced discussions of how best to iterate or implement a particular change idea. In schools where teachers and administrators share a specific common vision for classroom pedagogy, sharing and spreading change idea iterations may occur more rapidly.

Additional structures that supported teachers as they engaged with improvement science tools and methods included protocols, regular meeting times, and looking at student data. Protocols provided an equitable structure that ensured all voices were heard during the root cause analysis and the development of a working theory of action, increasing the likelihood that every teacher would develop a sense of ownership over the decisions made by their improvement group. The regular meeting times were used to build teacher capacity for improvement (the six professional development workshops) and to provide time to plan and debrief the process of

engaging in PDSA cycles (morning meeting times). The meeting times provided opportunities to look at student data, either whole school data or class specific data, and were reported to increase ownership over the improvement process. Teachers also reported that the school-wide focus on improvement and the improvement group structures helped provide an additional level of accountability and motivation for the work.

Challenges to Engaging in Improvement Science

At the heart of improvement science is iterative inquiry, gathering data on small change ideas to determine if improvement is occurring. Despite significant familiarity with scientific inquiry (three out of the four focus teachers taught science) teachers still struggled to collect data that they could reliably use to determine if a change idea was leading to improvement or to collect data in a systematic way to determine the next cycle of inquiry. Factors that may underpin these data collection and analysis challenges include the already time consuming reality of teaching, the complex nature of the problems teachers wish to work on, and teacher beliefs about what kinds of classroom structures and routines support student learning.

Teaching is an incredibly complex, time consuming, and emotionally intense profession. For teachers, practices that do not yield immediately useful information about their students can feel burdensome and irrelevant. Without adequate time allotted for planning PDSAs and analyzing the data to inform practice, improvement science risks becoming just one more thing that teachers are pressured to make time for in their already packed schedule. However, if they can see evidence that their change ideas are having an impact on students, improvement science has the potential to become a way of thinking. Colby's observations about how his students interacted with his word wall iterations eventually led him to adopt the anticipatory set into his routine. However, without careful observations and more systematic methods of collecting and

analyzing data it can be difficult to identify if change ideas are leading to improvement. In these cases inquiry efforts risk being reduced to teachers trying a series of one-off change ideas that never result in significant learning or refinement of a shared knowledge base.

Another possible challenge to engaging in systemic inquiry is teacher beliefs about what type of instruction best supports student learning. Determining which teaching practices are high leverage can be difficult without access to current research. It is possible that higher leverage change ideas might have yielded more compelling data and resulted in more sustained teacher inquiry around promising teaching practices. For example, if teachers decided to track how often focus students used specific vocabulary words during class discussions (or have students track themselves), or decided to increase the number of times they required students to use the vocabulary words across a series of lessons they may have detected more compelling results generating more excitement for the PDSA inquiry process.

Educators interested in adopting improvement science to guide improvement efforts at their school site may wish to provide additional support for determining which change ideas are truly high leverage, and what data to collect to determine if the changes are leading to improvement. Making data collection manageable by focusing on a few students or arranging for coaches or support staff to help collect data may also support teachers as they engage in iterative inquiry.

Limitations of the Study

The limitations of this study include the small sample size of twenty teachers and four focus teachers, the single school site, and the nine-month data collection time frame. The limited sample size reduces the probability of detecting a true effect and the qualitative data from the four focus teachers is limited by the fact that all four teachers share the same school environment

and culture. In addition, the limited duration of the study precludes any generalizations about potential long-term effects on teacher agency. The results of this study are suggestive, not conclusive.

Additional limitations include the researcher's position as co-facilitator of the improvement science teacher professional development for the participating teachers. As such, there is a risk of bias towards the intervention. The triangulation of multiple data sources was employed to mitigate this risk.

Finally, there is often a power asymmetry between the researcher and participants in research interviews (Kvale & Brinkmann, 2009). In addition to explicit requests for participants to speak their truths during interviews, observations and artifact data were used to check against this possibility.

Implications for Future Research

This study explored how teachers engaged with the tools and methods of improvement science over the course of an academic year and its impact on (1) their sense of belonging to an improvement community, (2) their belief that they could meet the learning needs of all their students, and (3) their belief that learning to use the tools of improvement science was useful to their practice now and in the future. Findings suggest that by participating in the professional development, teachers felt an increased sense of belonging to an improvement community and believed that improvement science was a useful framework to use to meet the learning needs of all their students. Teachers valued the opportunities for collaboration and tested change ideas across improvement groups. Teachers also used both whole school and PDSA level performance and perception data to inform their practice and guide improvement. Ultimately, teachers

adopted the tools and methods of improvement science and decided to unite around a singular focus: to improve instruction for their emerging bilingual students.

Due to the small sample size of this study and the singular school site, further research is needed to determine if these effects could be replicated in other school settings and which support structures were most effective. In similar contexts, where teachers report high levels of agency at the beginning of a study, using a retrospective pretest may ameliorate any response shift bias in pre- and post-intervention survey results.

Teachers also found the protocols to be a helpful scaffold for using the root cause analysis tools of improvement science, however, additional research about how teachers engaged with the protocols and the degree to which they supported equitable discussions and deeper thinking about the problem would be useful for others interested in using protocols to scaffold improvement work. In this study, teachers did not return to their fishbone or driver diagrams, despite that fact that many groups were developing a deeper understanding of their focus areas over the course of the year. The use of protocols to engage with improvement science tools may have formalized the activity, leaving teachers with a feeling that to revisit their fishbone, interrelationship digraph, or driver diagrams required a longer, more structured conversation than actually necessary. In practice, each of the tools can be used to dig further and further into narrower and narrower aspects of a problem. Perhaps having teachers use these tools to explore their rationales for specific change ideas they chose would have helped them see that the tools could be used for all levels of a system, from school wide issues to details of their classroom routines. In light of this, additional research on how to support improvement groups in returning to their working theory of action to make adjustments as they engage in improvement work would also be useful.

Another aspect of improvement science that teachers found challenging was the PDSA cycle. Further research is needed to elucidate what challenges to engaging in iterative cycles of inquiry exist across different contexts and possible structures to support teachers in collecting more informative and useful data to guide their improvement efforts.

Conclusion

Despite years of top-down reform initiatives, persistent disparities in opportunities exist for disadvantaged students. Over the past century decision making power has consolidated at the state and federal levels stripping teachers of control over their profession and situating power over educational decisions far away from the classrooms where teaching and learning occurs. There is a need for bottom-up educational reform that fosters teachers sense of professional agency and situates teachers as key decision makers and brokers of the boundaries of their profession. Improvement science, with its focus on teacher inquiry, teacher collaboration, and building a shared pedagogical knowledge base, is a bottom-up reform strategy that addresses these needs.

Using a mixed methods approach, this study found that engaging in improvement science professional development increased teachers' sense of agency in three ways. Teachers reported an increased sense of belonging to an improvement community, an increased belief that they could meet the learning needs of all their students, and that learning about improvement science tools and methods was useful to them now and in the future. Over the course of the study teachers developed a shared understanding of four teacher-identified problems of practice and generated four aim statements and working theories of action to guide their improvement efforts.

The teachers used various forms of data to inform their improvement work, including whole school performance data from state standardized tests and MAP assessments as well as

student perception data from the Youth Truth school climate survey. Teachers also collected performance and perception data from their students during the PDSA cycles to determine if classroom level change ideas were resulting in improvement.

Structures that supported the improvement work included administrator support, regular times to meet as an improvement group, having colleagues co-facilitate the professional development sessions, and the use of protocols to scaffold using the various tools of improvement, including fishbone diagrams, interrelationship digraphs, and driver diagrams. Challenges teachers experienced during the study included determining what data to collect to determine if a change idea was leading to improvement and using data to inform iterative cycles of inquiry during the PDSA cycles.

Despite these challenges, an improvement culture emerged at the middle school, with teachers coalescing around a singular improvement goal of supporting their emerging bilingual students. Two teachers from the EL support improvement group emerged as leaders for the new work, integrating improvement tools into their plans to launch the school-wide improvement efforts the next year. It is my hope that the results of this study will be useful for educators interested in implementing improvement science as a mechanism for bottom-up reform and for fostering teacher agency in order for teachers to exercise decision-making power over their own profession.

Works Cited

About Achieve. (2016). Retrieved September 18, 2016, from http://www.achieve.org/about-us

Anderson, M. (2017). Transformational Leadership in Education : A Review of Existing Literature Transformational Leadership in Education : A Review of Existing Literature. *International Social Science Review*, *93*(1).

Berube, Maurice R. (January 1, 1994). *American School Reform: Progressive, Equity, and Excellence Movements, 1883-1993*. Greenwood Publishing Group. p. 16. ISBN 978-0-275-95036-1.

Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching: Theory and Practice*, *21*(6), 624–640.

Bloomfield, D., & Cooper, B., (2003) *NCLB: A New Role for the Federal Government.* THE Journal: Transforming education through technology. Published 5/1/03 Retrieved from: https://thejournal.com/Articles/2003/05/01/NCLB-A-New-Role-for-the-Federal-Government.aspx?Page=1

Bodman, S., Taylor, S., & Morris, H. (2012). Politics, policy and professional identity. *English Teaching: Practice and Critique*, *11*(3), 14–25.

Brinkley, A. (2009). American History: A Survey Volume. II. New York, NY: McGraw Hill.

Brown v. Board of Education, 347 U.S. 483 (1954).

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015) *Learning to improve: How America's schools can get better at getting better.* Cambridge MA: Harvard Education Press.

Buchanan, R. (2015). Teacher Identity and Agency in an Era of Accountability. *Teachers and Teaching: Theory and Practice*, *21*(6), 700–719.

Carnegie Foundation for the Advancement of Teaching (2013). *Student Agency Improvement Community survey tool* [Google form].

Cochran-Smith, M., & Lytle, S. L. (2006). Troubling Images of Teaching in No Child Left Behind. *Harvard Educational Review*, *76*(4), 668–726.

Coleman, James S. (1966) *Equality of Educational Opportunity (COLEMAN) Study (EEOS)*, Retrieved from: http://files.eric.ed.gov/fulltext/ED012275.pdf

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. (4th ed.) Boston, MA: Pearson Education.

Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Cubberley, E. P., (1922) Public school administration: a statement of the fundamental principles underlying the organization and administration of public education. Houghton Mifflin Co. Boston Retrieved from: https://archive.org/details/publicschooladmi1922cubb on September 4, 2016.

Darling-hammond, L. (1994). Performance-Based Assessment and Educational Equity. *Harvard Education Review*, *64*(1), 5–30.

Day, C. (2002). School reform and transitions in teacher professionalism and identity. *International Journal of Educational Research*, *37*(8), 677–692.

Deci, E. and Ryan, R. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology*, *49*(3), 182-185.

Deming, W. E. (1993). *The new economics for industry, government, education*. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Doig, B., & Groves, S. (2011). Japanese lesson study: Teacher professional development through communities of inquiry. *Mathematics Teacher Education and Development*, *13*(1), 77–93.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Elmore, R. F., (2004) *School Reform from the Inside Out: Policy, Practice, and Performance*. Cambridge, MA: Harvard Education Press.

Elmore, R. F., & Elmore, F. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, *66*(1), 1–26.

Emirbayer, M., & Mische, A. (1998). What Is Agency ? *The American Journal of Sociology*, *103*(4), 962–1023.

Francis, D. J., & Rivera, M. (2006). *Practical Guidelines for the Education of English Language Learners: Research-based Recommendations for Instruction and Academic Interventions*. *Center on Instruction* (Vol. 1).

Gomez, K., Gomez, L. M., Rodela, K. C., Horton, E. S., Cunningham, J., & Ambrocio, R. (2015). Embedding Language Support in Developmental Mathematics Lessons: Exploring the Value of Design as Professional Development for Community College Mathematics Instructors. *Journal of Teacher Education*, *66*(5), 450–465.

Hallinger, P. (2003). Leading Educational Change: reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education*, *33*(3), 329–352. http://doi.org/10.1080/0305764032000122005

Hannan, M., Russell, J. L., Takahashi, S., & Park, S. (2015). Using Improvement Science to Better Support Beginning Teachers: The Case of the Building a Teaching Effectiveness Network. *Journal of Teacher Education*, *66*(5), 494–508.

Hattie, J. A. C. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, *37*(5), 449–481.

Howard, G. S., & Dailey, P. R. (1979). Response-shift bias: A source of contamination of self-report measures. Journal of Applied Psychology, 64(2), 144-150.

Iorio, S., & Yeager, M. (2011). *School Reform : Past, Present and Future*. Retrieved from: http://webs.wichita.edu/depttools/depttoolsmemberfiles/COEdDEAN/School%20Reform%20Past%20Present%20and%20Future.pdf

Kelley, T., & Kelley, D. (2013). *Creative confidence: Unleashing the creative potential within us all*. Crown Publishing, New York, NY

Kiser, E. (1996). The Revival of Narrative in Historical Sociology: What Rational Choice Theory Can Contribute. *Politics & Society*, *24*(3), 249–271.

Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and Teacher Education*, *21*(8), 899–916.

Lee, J., & Reeves, T. (2012). Revisiting the Impact of NCLB High-Stakes School Accountability, Capacity, and Resources: State NAEP 1990-2009 Reading and Math Achievement Gaps and Trends. *Educational Evaluation and Policy Analysis*, *34*(2), 209–231.

Lemoine, H. E., Levy, B. A., & Hutchinson, A. (1993). Increasing the naming speed of poor readers: Representations formed across repetitions. *Journal of Experimental Child Psychology*.

Lessinger, L. M. (1970). Every Kid a Winner: Accountability in Education. Simon & Schuster, New York.

Lewis, C. C., Perry, R. R., & Hurd, J. (2009). Improving mathematics instruction through lesson study: A theoretical model and North American case. *Journal of Mathematics Teacher Education,* 12, 285–304.

Lewis, C., & Tsuchida, I. (1998). A lesson is like a swiftly flowing river: Research lessons and the improvement of Japanese education. *American Educator*, 14–17 & 50–52.

Light, J. S. (2003). *From Warfare to Welfare:Defense Intellectuals an Urban problems in Cold War America.* Baltimore, MD: Johns Hopkins University Press.

Little, J. W. (1993). Teachers' Professional Development in a Climate of Educational Reform. *Educational Evaluation and Policy Analysis, 15*(2), 129-151.

Lofland, J., & Lofland, L. H. (1995). *Analyzing social settings*. Belmont, CA: Wadsworth Publishing Company.

Lorenzi, N. M., & Riley, R. T. (2000). Managing Change: An Overview. *Journal of the American Medical Informatics Association*, *7*(2), 116–124.

Louis, K. S., Leithwood, K., Wahlstrom, K., & Anderson, S. (2010). Investigating the links to improved student learning: Final report of research findings. Retrieved from Wallace Foundation website: www.wallacefoundation.org/knowledge-center/schoolleadership/key-research/Documents/Investigating-the-Links-toImproved-Student-Learning.pdf

Lukes, S., (1973) *Individualism.* Oxford, UK: Basil Blackwell.

MacConnell, K. & Caillier, S. (2016). Getting better together: One school's journey with improvement science. Manuscript accepted by *Phi Delta Kappan* for publication.

Martin, W. G., & Gobstein, H. (2015). Generating a Networked Improvement Community to Improve Secondary Mathematics Teacher Preparation: Network Leadership, Organization, and Operation. *Journal of Teacher Education*, *66*(5-), 482–493.

Mathis, W. (2010). The "'common core'" standards initiative: An effective reform tool? ... *Tempe: Education and the Public Interest Center & …*, (July). Retrieved from http://epicpolicy.org/publication/common-core

Mehta, J. (2013). *The allure of order: High hopes, dashed expectations, and the troubled quest to remake American schooling*. New York, NY: Oxford University Press.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: A sourcebook of new methods* (2nd ed.). Thousand Oaks, CA: Sage.

Nagaoka, J., Farrington, C. A., Ehrlich, S. B., & Heath, R. D. (2015). *Foundations for Young Adult Success: A Development Framework*. Chicago.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002)

Noland, T. K., & Taylor, B. L. (1986). The effects of ability grouping: A meta-analysis of research findings. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

den Ouden, H. E. M., Kok, P., & de Lange, F. P. (2012). How Prediction Errors Shape Perception, Attention, and Motivation. *Frontiers in Psychology*, *3*(December), 1–12.

Plano Clark, V., & Creswell, J. (2010). *Designing and conducting mixed method research.* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Rock, T., & Wilson, C. (2005). Improving teaching through lesson study. *Teacher Education Quarterly*, 77–92.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, *55*(1), 68–78.

Simmie, G. (2007). Teacher Design Teams (TDTs)—building capacity for innovation, learning and curriculum implementation in the continuing professional development of in-career teachers. *Irish Educational Studies*, *26*(2), 163–176.

Stigler, J. W., & Hiebert, J. (1999). The teaching gap: Best ideas from the world's teachers for improving education in the classroom. New York: Summit Books.

Taylor, J. (2010). *State and local implementation of the No Child Left Behind Act. final report: A report from the National Longitudinal Study of No Child Left Behind (NLS-NCLB) and the Study of State Implementation of Accountability and Teacher Quality Under No Child Left Behind (SSI-NCLB)*. Washington, D.C.: U.S. Dept. of Education.

United States. National Commission on Excellence in Education. (1983). *A Nation at Risk: The Imperative for Educational Reform.* A report to the Nation and the Secretary of Education, United States Department of Education. Washington, D.C.: The Commission.

United States, U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2015). *National Assessment of Educational Progress: NAEP*. Washington, DC: National Center for Education Statistics, U.S. Dept. of Education.

Valenzuela, D., & Shrivastava, P. (2007). *Interview as a method for qualitative research*. Presentation retrieved March 15, 2016 from http://www.public.asu.edu/~kroel/www500/Interview%20Fri.pdf

Vedantam, S. (Producer). (2015, October 12). *Hidden Brain: Students and Teachers* [Audio Podcast] Retrieved from http://www.stitcher.com/podcast/hidden-brain

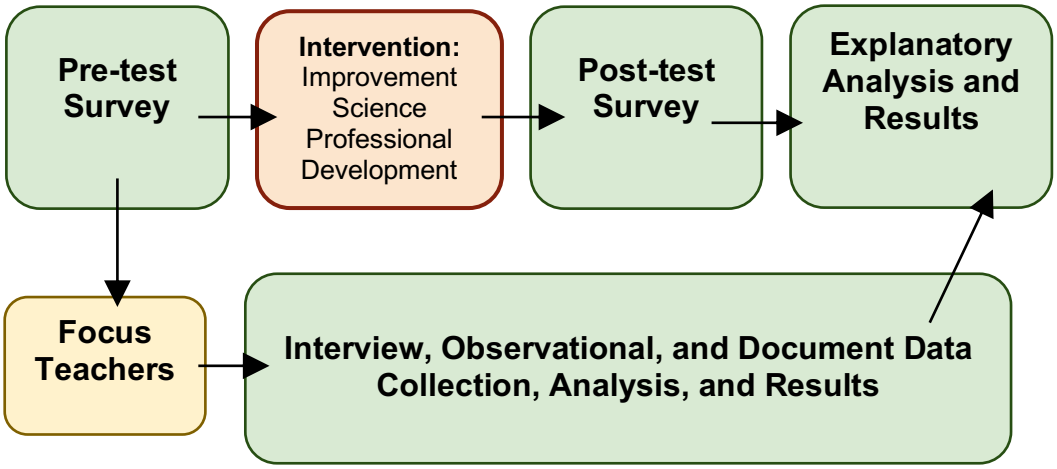
Wigfield, A., Eccles, J. (1992). The development of achievement task values: A theoretical Analysis. *Developmental Review, 12*(3), 265-310.

Wilson, P.F., Dell, L. D. & Anderson, G. F. (1993). *Root Cause Analysis: A Tool for Total Quality Management*. Milwaukee, Wisconsin: ASQ Quality Press. pp. 8–17.

Yin, R. K. (2009). *Case study research: Design and method* (4th ed.). Thousand Oaks, CA: Sage.

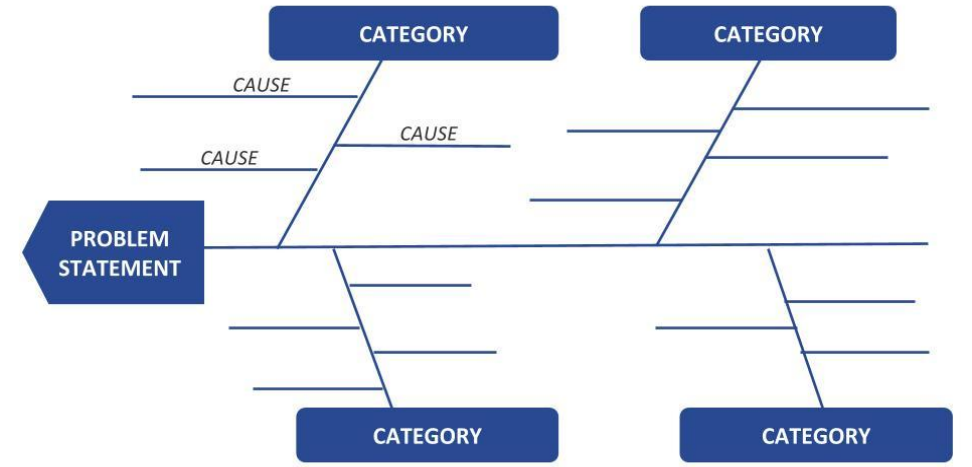
Appendix A: Design of the Improvement Science Professional Development Intervention

Model of the Intervention

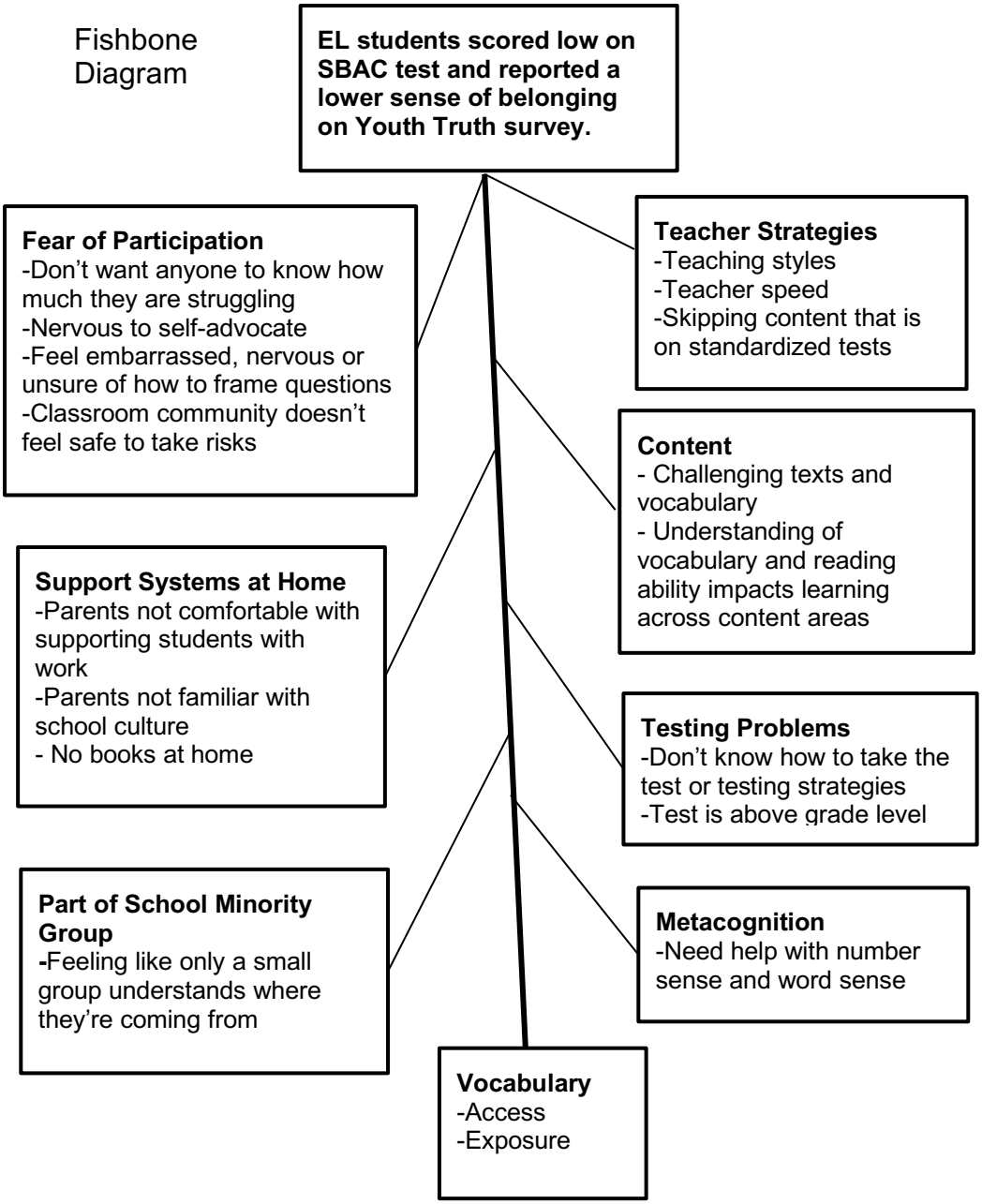


Appendix B: Improvement Science Tool Templates and Examples from EL Improvement Group

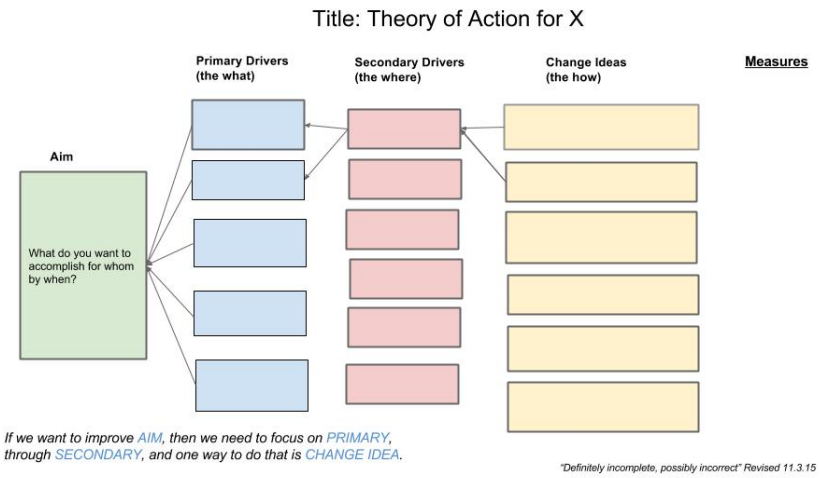
Fishbone Diagram Template



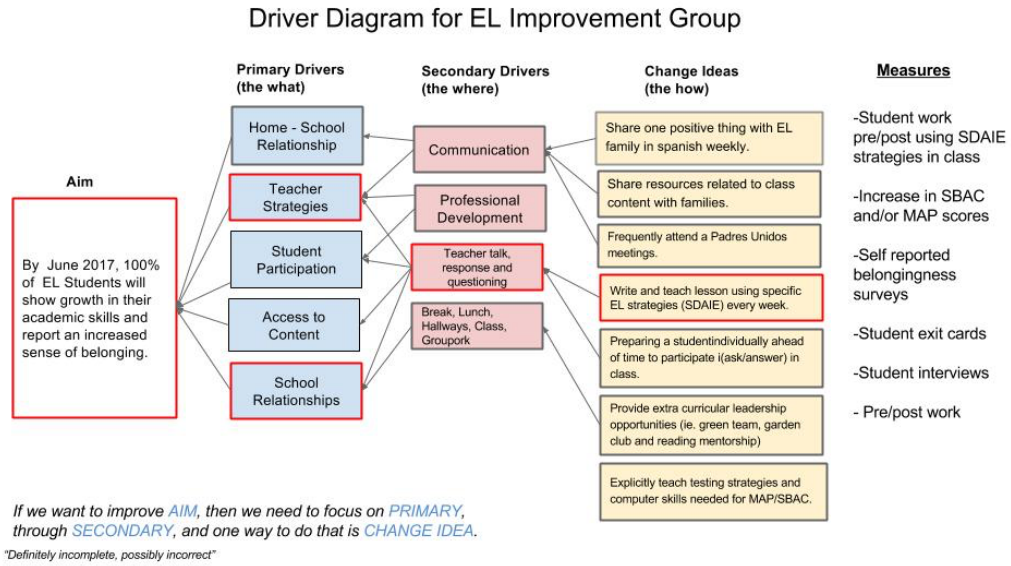
Fishbone Diagram



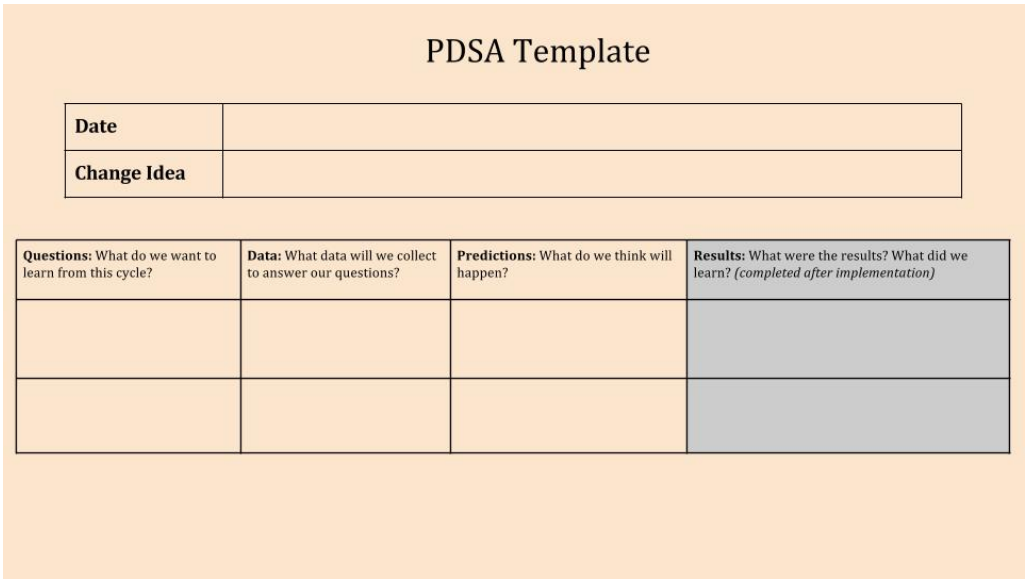
Driver Diagram Template



Driver Diagram Created by the EL Improvement Group



PDSA Form Template



PDSA Template

Date	
Change Idea	

Questions: What do we want to learn from this cycle?	**Data:** What data will we collect to answer our questions?	**Predictions:** What do we think will happen?	**Results:** What were the results? What did we learn? *(completed after implementation)*

Appendix C: Fishbone Diagram Protocol

The purpose of this protocol is to arrive at a deeper understanding of the problem we want to address (before jumping to solutions).

Norms:

- **Avoid Solutionitis**… the goal is to understand the issue, not solve it (yet)
- **"Yes and"...** the goal is to generate lots of ideas, and not fixate on one
- **Embrace "definitely incomplete; possibly incorrect"**
- **Share the air**

1. **Generating our Problem Statement** (5-7 minutes)
    - Individual: What is the problem we need to solve? See if you can express the problem in one sentence.
    - Share out problem statements and decide on one to pursue as a group
2. **Initial Brainstorm of Causes** (5 min.)
   Based on your work digging into the problem (i.e. empathy interviews, expert convenings, relevant data, research, etc.) and your own ideas/experiences, *individually brainstorm* as many causes as you can that might contribute to the problem/issue. Write each cause on a different post-it. *For meaty "big" topics, it can help to ask a chain of "why?".*
3. **Share & Categorize** (15-20 min)
    - *Share around:* Each person shares one cause contributing to the problem. If others have a similar cause, you can start to group those post-its together on your poster.
    - *Continue to share* your initial brainstorm, building on each other's ideas and adding new causes that may contribute to the problem.
    - **Cluster on your Poster:** Group related causes together, and give each category a title. (The stuff on the post-its are the details/bones on the fishbone).
4. **Post & Reflect** (5 min)
   Post your poster to the wall. Does your diagram capture the root causes you think are important? Anything missing? Then *each person* gets to vote with *one heart* and *one star*:
    - *High Leverage*: Put a **heart** by the factor, that if addressed, you think would have a significant impact on the problem.
    - *Practical*: Put a **star** by the factor that is within your control that your team could address with little effort.
5. **Debrief** (5 min)
   How did we do upholding the norms? How might we adjust this protocol in the future?

This protocol has been created by the High Tech High GSE Center for Research on Equity and Innovation.

Appendix D**:** Interrelationship Digraph Protocol

A convergent tool for dialogue that helps us determine which root causes are most important to address.

Norms:

- **Avoid Solutionitis**… the goal is to understand the issue, not solve it (yet)
- **"Yes and"...** the goal is to generate lots of ideas, and not fixate on one
- **Embrace "definitely incomplete; likely incorrect"**
- **Share the air**

Roles:

- **Facilitator** who monitors times, walks people through the process, and helps uphold the norms.

Protocol:

Step 1: Write the problem statement at the top of the poster.

Impact Group Problem Statements:

- Not all students experience quality internships.
- Not all students are engaged in quality presentations of learning.
- Too many students are chronically absent.
- Not enough students transition successfully to college.

Step 2: Identify 6-8 of the most significant causes contributing to this problem.

(Note: If you previously created a fishbone diagram, the categories from your fishbone can serve as the causes. In this case, skip to step 3.)

Initial Brainstorm of Causes (2 min.)

Based on your work digging into the problem (i.e. empathy interviews, expert convenings, relevant data, research, etc.) and your own ideas/experiences, *individually brainstorm* 3-4 of the most significant causes contributing to the problem (i.e. access, skills, parent knowledge, lack of training, etc.).

Quick Share & Categorize (10 min)

The goal here is to quickly identify the 6 - 8 big barriers that you are looking at en

- *Share around & Cluster:* Each person shares one cause contributing to the problem. If others have a similar cause, you can start to group those post-its together. Continue to share until everyone's ideas are out.
- *Select:* Identify the 6-8 most significant causes contributing to the problem and write them in a circle around the outer edge of the poster, like the numbers on a clock face.

Step 3: Each person makes a prediction. (2 min.) Which cause do you think is most important? (You can either share these now, or ask people to write their prediction on a post-it and share them later.)

Step 4: Construct the Digraph.

Starting with one cause, for each pair of causes ask yourself:

- Is there a relationship between these two?
- If yes, which causes the other, *the most*? Draw an arrow from one to the other to show directionality.
  - For example, the facilitator may ask the group "Does social anxiety about school cause a lack of motivation or visa versa?" If the group thinks social anxiety causes lack of motivation (more than lack of motivation causing social anxiety), draw the arrow going toward motivation (*from cause to effect*).
- You can decide there is not a causal relationship, but you must pick a direction if you do see a relationship.

Repeat until you have established a relationship (or not) between all the topics.

Tally (out, in) for each cause.

- ***The root causes with the most <u>outgoing lines</u> most impact the issue/problem. Star the top 1-2 root causes!***
- The root causes with the most ingoing lines tend to be the effects/symptoms of the causes

Step 5: Discuss Predictions & Implications

- Were our predictions correct? What root cause(s) have we learned are most at the heart of the problem?
- What have we tried to address the root causes - and what are the *successes* and *challenges* we've experienced? What will we do next?

This protocol has been created by the High Tech High GSE Center for Research on Equity and Innovation.

Appendix E: Driver Diagram Protocol

The purpose of this protocol is to generate a shared "theory of action" to drive a team's improvement efforts and ultimately achieve the aim.

Note: Driver diagrams are not intended to be set in stone. Your team's theory of action should evolve as you learn more about the problem/gap you want to address, the change ideas you are trying, and how best to achieve your aim. We encourage teams to revise your driver diagram to reflect your most current thinking/focus, and to keep track of versions 1.0, 2.0, etc. so you can reflect on the evolution of your learning.

Norms:

- **"Yes, and"**... the goal is to generate lots of ideas, not fixate on one
- **Embrace "definitely incomplete; possibly incorrect"**
- **Share the air**

Some helpful definitions:

- ***Aim*** = what you want to accomplish for whom by when (samples below)
- ***Primary Driver* (*the what*)** = what you need to focus on to achieve your aim (i.e. *financial literacy*)
- ***Secondary Driver* (*the where*)** = where in your system (the existing structures/processes) you should focus your energies to effect the primary driver (i.e. *parent workshops*)
- ***Change Idea* (*the how*)** = what you can try and test/refine (i.e. *Families complete the FAFSA together at the monthly parent meeting)*

Protocol:

1. **Clarifying/Refining our Aim (10 minutes)**

As a group, craft/refine your aim statement: What do you want to accomplish, for whom, by when?

It can help to begin by having each person, individually or with a partner, craft an aim statement. Follow this with a whip where each person/partnership shares their aim with

the group. Then the group can adopt/adapt from these to create an aim statement everyone feels good about. *Write your group's aim statement on the left side of your poster.*

Some things to consider:

- Is your aim ***measurable***? To help ensure your aim is measurable, it can be helpful to ask yourselves this question: "If X was the best it could be, what would it look like?" It can also help to get baseline data related to the issue/gap your team is working on.
- **Some sample aims:**
    - By Spring 2015, all HTHNC seniors will apply to a 4-year college.
    - By the end of 15-16, CAT HS will decrease our % of students who need remedial Math & English courses in college from 85% to 50%.
    - By May 2016, 60% of seniors at John Muir HS will score a 3 or 4 on the first reading of their senior research paper (compared to 20% last year).

2. Identifying Primary Drivers (20-30 minutes)

- ***Individual (no more than 5 minutes)*:** Each person identifies the top 4 drivers (i.e. *high leverage areas*) they think the team needs to focus on to impact the aim, and writes each driver on a separate index card.
    - **Facilitation Move**: It can help to think of drivers as X in the following statements: "If we figured out X, we could achieve our aim" or "If we don't figure out X, it is unlikely we would achieve our aim."

- ***Whip & Cluster*:** Each person shares their favorite driver with the group. If others wrote down a similar driver, group these cards together on the table.
    - **Facilitation Move**: As your group shares and clusters, it can be helpful to organize the "stacks" with the most cards to the top of the table, and those with the least to the bottom. This gives the group a visual indicator of which drivers might be most important.

- ***As a group, select 3-5 drivers*** that you think are essential for impacting your aim. Write those drivers on your driver diagram poster. This is your "theory of action" (i.e. if you could move these drivers, you could achieve your aim).
    - **Questions for the team to consider:**
        - Is this driver ***specific*** enough that we all understand what to focus on?
        - Is this driver ***impactful*** enough that it will move the work forward?

- Is this driver within our ***locus of control***, meaning we can do something about it? (i.e. Poverty is real, but may not be a helpful driver. However, Family Support could be an important driver and signals a way of working with families to reduce the effects of poverty.)
- Are these drivers ***necessary and sufficient*** for achieving our aim?
- Which driver do we think is our ***greatest lever*** for change? **(Star this one.)**

3. Identifying Secondary Drivers (if time permits)

If your team has time, you can add secondary drivers to your diagram. Secondary drivers articulate "where" in the system (i.e. what structures and processes) you can focus your efforts to impact the primary drivers. They can be helpful for focusing the team's efforts and coming up with more concrete change ideas. For example, helpful secondary drivers may be Advisory, Teacher PD, Number Talks, Classroom Meetings or Master Schedule.

4. Generating Change Ideas (15 minutes)

- ***Individual Brainstorm (3 minutes)*:** How might we impact the drivers we identified? What might we try? Where is this happening well already (bright spots), and what are they doing? *Write each idea on its own post-it.*

- ***Share & Align (10-12 min):*** Individuals share their ideas with the group, and post each change idea by the driver they feel it most impacts. *(helpful questions to consider on the next page)*
  - **Question for the team to consider:**
    - What do we notice about the ***alignment*** (or lack of) between our change ideas and drivers? (If you have lots of change ideas that don't align to your existing drivers, this could suggest a new driver is needed. If you have a driver without any change ideas, this driver may not be helpful to your current "theory of action".)

- ***Identify high leverage change ideas*:** As a group identify 3-5 change ideas that you think are most impactful, and that you could get moving on quickly. **Star these on your driver diagram.**

5. Debriefing the Process (5 minutes)

Group members discuss the following questions:

- How well did we do with upholding the norms and sticking to the protocol?
- Was this protocol helpful for identifying high leverage drivers and change ideas?
- What worked well about this protocol? What could be improved?

This protocol has been created by the High Tech High GSE Center for Research on Equity and Innovation.

Appendix F: Pre- and Post-intervention Survey

Teacher Survey
{This survey will be administered digitally and will not include the italicized subheadings}

Please rate each question on a scale of 1 - 5:

{Belonging to an Improvement Community}

1. My school feels like a community that is always trying to improve to get better at meeting the learning needs of our students

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

2. My colleagues actively test out new teaching practices to meet the learning needs of all their students

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

3. My colleagues use data such as student work, exit cards, or student surveys to determine if their teaching practices are working for all their students

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

4. I share new teaching strategies with my colleagues

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

5. I learn new teaching strategies from my colleagues

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

6. My director provides opportunities for me to learn new teaching practices

1	2	3	4	5

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

7. My director provides opportunities for me to create new teaching practices

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

8. My director provides opportunities for staff to learn from each other

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

9. My director supports staff efforts to improve their practice

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

{Mastery/Confidence}

1. I feel confident that I can meet the learning needs of all my students

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

2. I feel confident that I can learn new pedagogical tools to meet the learning needs of all my students

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

3. I feel confident that I can create new pedagogical tools to meet the learning needs of all my students

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

4. I actively test out new teaching practices to meet the learning needs of all of my students

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

5. I use data such as student work, exit cards, or student surveys to determine if my teaching practices are working for all my students

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

{Value}

1. The professional development I receive at my school helps me meet the learning needs of my students

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

2. The professional development I receive at my school helps me improve my teaching practice

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

3. The professional development I receive at my school will be useful to me in the future

1	2	3	4	5
Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

Appendix G: Interview Protocol for Focus Teachers

Exploring How Teachers Engage With Improvement Science Methods in Their Practice.

Date	
Time of Interview	
Place	
Interviewer	
Participant	
Title	
School	

Thank you for agreeing to participate in this interview. The purpose of this study is to better understand how teachers engage with improvement science methodologies.

Your interview data will be kept confidential. Only the researcher and a professional transcriptionist will listen to and transcribe the information you provide. The audiotapes will be destroyed following final analysis; no later than May, 2018.

Your participation is entirely voluntary and may be withdrawn at any time. If the length of the interview becomes inconvenient, you may stop at any time. There are no consequences if you decide not to participate.

Semi-structured Interview Questions
Questions such as these will be asked during biweekly interviews with focus teachers.

How do teachers engage in improvement science methods?

Knowledge of improvement science methods - challenges/supports in the environment

1. What improvement science tools do you find most useful? Why?
2. What structures or school systems helped you learn about these tools?
3. What structures or school systems made it difficult to learn about these tools?

Using of improvement science methods - challenges/supports in the environment

1. Can you describe a time when you used improvement science methodologies?
2. What was hard about using improvement science? Why?
3. What was easy about using improvement science? Why?
4. What structures or school systems helped you use improvement science tools?
5. What structures or school systems made it difficult to use these tools?

Discussing/Sharing of improvement science methods - challenges/supports in the environment

1. Can you describe a time when you discussed or shared ideas or results from using improvement science methods with a colleague?
2. What structures or school systems helped you discuss or share these ideas or results?

3. What structures or school systems made it difficult to share or discuss these ideas or results?

How do improvement science methodologies affect teacher agency?
What improvement science structures support/impede teacher agency?

Value of improvement science methods - challenges/supports in the environment

1. How confident do you feel that improvement science methods or tools will help you meet the learning needs of all your students? Why?
2. What structures or school systems helped you feel that improvement science methods and tools could help you meet the learning needs of all your students?
3. What structures or school systems made it challenging to feel that improvement science methods and tools could help you meet the learning needs of all your students?

Confidence/mastery of improvement science methods - challenges/supports in the environment

1. How confident do you feel using improvement science methods or tools? Why?
2. Which improvement science tools or methods have been easier to use? Why?
3. Which improvement science tools or methods have been more challenging to use? Why?
4. What structures or school systems helped you feel confident to use improvement science methods or tools?
5. What structures or school systems made it difficult for you to feel confident in using improvement science methods or tools?

Sense of belonging to an improvement community - challenges/supports in the environment

1. Do you feel like you are part of an improvement community? Why?
2. What structures or school systems helped you feel that you are part of an improvement community?
3. What structures or school systems made it challenging to feel that you are part of an improvement community?

Appendix H: Observation Protocol

Exploring How Teachers Engage With Improvement Science Methods in Their Practice

Date	
Time of Observation	
Teacher	
Specific Places Observed (use a separate protocol for each classroom observed)	

The purpose of these observations is to find evidence of teacher engagement with improvement science methods and what factors support or hinder their engagement. Specific examples are listed below each category.

These are provided to guide the researcher and should not be considered the only possible manner in which engagement could be evidenced.

Knowledge of improvement science methods
Teachers understand the PDSA process. Teachers use improvement science tools such as empathy interviews, fishbone diagrams, or system analyses to understand problems related to practice. Teachers design useful data collection tools to collect evidence to inform decisions to adopt, adapt or abandon a particular teaching practice.

Using of improvement science methods
Teachers are actively engaged in an improvement project in their classroom or at their school site. Teachers collect data (student work samples, exit cards, survey data, etc…) to inform next steps.

Discussing/Sharing of improvement science methods
Teachers are actively trying out another teacher's classroom practices to see if they work in new contexts.

Appendix I: Digging into Data Protocol

Purpose: To help a group engage in productive dialogue about data, and to build collective capacity to make sense of data relevant to teaching and learning. You can use this protocol with multiple small groups, each unpacking a different piece of data, and then sharing out to the group. Or you can use this protocol to have everyone unpack the same piece of data (in this case, modify Stage 3).

Roles:

- *Facilitator* who guides the group through the process and ensures that the protocol and norms are upheld.
- *Notetaker* who captures notes from the discussion and is prepared to share out for the group.

Norms: Data conversations can make people feel vulnerable. A structured dialogue helps maintain safety and focus.

- Share the air: step up, step back, invite others in.
- Hard on the content, soft on the people.
- Focus on learning, not judging.
- Stick to the protocol.

Phase 1: Getting Oriented (5 minutes)
Participants take 2-3 minutes to individually review the data. The facilitator then leads a quick check-in: Does everyone understand what is being presented?

Phase 2: Discussion Rounds (25 minutes)
Round 1: Observations & Celebrations (whip, people may pass or say "ditto")

- Each person shares one thing they noticed that they want to celebrate and/or call attention to. During this portion it is important to just describe what you see as objectively as possible. Resist the urge to interpret or pose questions.
- Helpful sentence frame: **I notice….**

Round 2: Questions (whip other direction, start with a different person, people may pass or say "ditto")

- Each person shares a question that emerges for them from the data.
- Helpful sentence frame: **I wonder...**

Round 3: Hypotheses (facilitator facilitates a discussion)

- Participants share possible hypotheses or explanations for what they see, trying to identify multiple alternative explanations.
- Helpful sentence frames: **This could be because... Or it could be because...**

Round 4: Next Steps (facilitator facilitates a discussion)

- Participants share what they might do next given their understanding of the data.
- Helpful sentence frames: **One thing we could do next is...**

Phase 3: Share out (optional, time varies)

The notetaker from each group has **1 minute** to share highlights from their group's discussion of the data. We recommend sharing the following:

- One celebration/noticing
- One question that emerged
- One possible hypothesis
- One next step

(Skip/modify this step if all participants have looked at the same data or if you only have one group.)

Phase 4: Debrief (5-7 minutes)

The facilitator leads the group in reflecting on this process.

Helpful Guiding Questions:

- What was this process like for you?
- What adjustments would you make?
- How did looking at data influence your understanding of the issue?
- What are we learning about how to package data so that we can engage in productive conversations?
- How might you use this protocol in your work, with your team?
- What data are we craving?

This protocol has been created by the High Tech High GSE Center for Research on Equity and Innovation, and adapted from a similar protocol designed by the Carnegie Foundation for the Advancement of Teaching and Learning.

Appendix J: Word Wall Iterations

Stephanie's word wall template (taped to desks):

Cell Vocabulary

Research your assigned Cell Vocabulary word and include the following in the template below.

Your Cell Vocabulary Card must include:

- ❑ Cell Vocabulary Word written large
- ❑ The definition of your vocabulary word (in kid-friendly language)
- ❑ An illustration/image to help illustrate the meaning/significance of your Cell Vocabulary word

Plant cell Animal cell Cell wall Cell membrane Vacuole Cell Cell Theory	Prokaryote Organelle Cytoplasm Nuclear envelope Nuclear membrane Chromatin Chromosome Nucleolus	Rough Endoplasmic reticulum Smooth Endoplasmic reticulum Golgi apparatus Lysosome Mitochondrion Chloroplast Cytoskeleton Ribosome	Concentration Diffusion Equilibrium Osmosis Eukaryote Y-Chromosome X-Chromosome RNA Nucleus	Exocytosis Cell specialization Tissue Organ Organ system Lipid bilayer Centriole DNA

Vocabulary Word: _____ *Definition:* Student Name:_____	*Illustration:*

Colby's final anticipatory planner document:

Project: Pre-Vocab

As a class we are going to do our best to understand everything we can about our upcoming project. Write what you already know about each word or words that pop in your mind. Draw a quick illustration/ image for each word.

Vocab Word	Words that come to mind and my best guess...	Illustration	Now I know this means...