

Understanding the Role of Equity in School Improvement Science: A Tale of Two Districts

By

Kimberly Amidzich

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Educational Leadership and Policy Analysis)

at the

UNIVERSITY OF WISCONSIN-MADISON

2021

Date of final oral examination: May 26, 2021

The dissertation is approved by the following members of the Final Oral Committee:

Richard Halverson, Professor, Educational Leadership and Policy Analysis

Julie Mead, Professor, Educational Leadership and Policy Analysis

Carolyn Kelley, Professor, Educational Leadership and Policy Analysis

M. Bruce King, Professor, Educational Leadership and Policy Analysis

Walter C. Stern, Assistant Professor, Educational Policy Studies and History

Table of Contents

Abstract	vi
Chapter 1.....	1
Introduction.....	1
Chapter 2.....	6
Literature Review.....	6
Table 1	7
Review of the Research: Findings	8
What is Improvement Science?	9
Defining the Common Processes of Improvement Science	9
Problem-Specific and User-Centered.....	10
Recognize Variation.....	11
Embrace Measurement and See the System that Produces the Current Outcomes.....	12
Standardized test results measure student achievement of common standards.....	15
Disciplined Inquiry Through Plan, Do, Study, Act (PDSA) Cycles.....	16
Evolution and History of Improvement Science	17
TQM and Continuous Improvement from Business Concept to Healthcare and Education	18
Research Studies Focusing on School Use of the Tools of Improvement Science	20
Addressing Problems of Inequity Using a Continuous Improvement Approach.....	22
Predictable variation in performance based on demographics.	23

An Equity Critique of Improvement Science	24
Voices of traditionally marginalized and questions of power.	25
Summary of Literature Review.....	28
Conceptual Framework	30
Research Question	33
Chapter 3.....	34
Methods.....	34
Design.....	34
Participants	36
School and District Administrators.	37
Teachers.	37
Data Collection Procedures	38
Instruments.....	39
Analysis.....	41
Trustworthiness and Validity.....	44
Ethical Considerations.....	45
Pilot of Administrator Interview Protocol.....	46
Limitations and Potential Implications	48
Contributions.....	49
Chapter 4.....	50
Northwestern District (Case #1).....	50
School District Context.....	50
Student Diversity.....	51

High School Achievement Measures	52
High School Engagement Opportunities.	53
Implementation of Improvement Science	54
Measurement.....	55
Perception Data.....	56
Academic Data.	57
Disaggregating Data.	58
Problem Specific Focus	59
Processes to Understand the System that Produces the Outcomes.	60
Disciplined Inquiry to Improve	62
Scorecard.	62
Plan Do Study Act (PDSA) Cycles for Improvement.....	63
Summary of Improvement Science Implementation.....	64
Equity Work.....	65
Perceptions of Facets of Diversity	66
Representation of Voices.....	67
Specific Improvement Focus Examples	69
Mental Health	70
Relationship Focused Programs.....	72
Case Summary	73
Chapter 5.....	76
Southeastern District (Case #2).....	76
School District Context.....	76

Student Diversity.....	77
Student Achievement Measures.....	77
Student Engagement Measures.....	78
Implementation of Improvement Science	78
Measurement.....	79
Academic Data.....	80
Perception Data.....	80
Behavior Data.....	80
Problem Specific Focus	81
Understanding the System that Produces the Outcome.....	82
Positive Variants.....	82
Disaggregating Data.....	84
Disciplined Inquiry to Improve	85
Plan, Do, Study, Act Cycles (PDSA).....	85
Equity Work.....	87
Mascot Change.....	89
Student Advocacy.....	90
Case Summary	92
Chapter 6.....	94
Findings	94
Implementation of Improvement Science.....	95
Ideas to Improve Implementation of Improvement Science.....	97
Focusing on Problems of Equity.....	98
Champion for Equity	99

Lived Experiences.....	100
Perspective Taking.....	102
Catalyzing Event.....	102
National Ignition.....	102
Local Spark.....	104
Measurement of Discrepancies.....	104
Disaggregating the Data.....	104
Considerations for Improvement Science Work.....	105
Implications for School Leaders.....	108
Opportunities for Future Research.....	110
Conclusion	111
References	113
Appendix A	117
Appendix B.....	120
Appendix C.....	123
Appendix D	124
Appendix E.....	125

Abstract

Schools across the United States employ Improvement Science principles as a strategy to make schools better. Despite these efforts, disparities in outcomes by race, social class, disability, and English language proficiency persist. The review of literature highlights implementation of continuous improvement tools in elementary and secondary education over the last three decades; however, consideration of demographic markers are limited in these studies. In recent years, the literature documents concerns about the treatment of equity and students from traditionally marginalized groups within continuous improvement and the extensive use of standardized test measures as a means of measuring improvement efforts. Drawing on the principles of Improvement Science (Bryk, Gomez, Grunow, & LeMahieu, 2015), this study seeks to understand how these principles, and their implementation in two districts, consider equity and disparities in outcomes to achieve improved outcomes for all learners through a lens of Critical Theory (Capper, 2019). Influenced by Bryk's (2017) hypothesis that Improvement Science is the key to making "real progress on the educational inequities we care about," (p. 4), the comparative case study seeks to articulate the convergence of Improvement Science and narrowing achievement and opportunity gaps in schools.

Chapter 1

Introduction

As a child, school was designed for me, and I thrived in a suburban public school setting. I idolized my teachers. I engaged in school projects and activities. I completed my work, and I wanted to be a teacher and a leader in schools just like the ones that had worked so well for me. I went to college, became a teacher, and continued the system that was effective for me, returning to the public school system from which I graduated.

When I moved from the classroom into administration, the country was buzzing about standardized tests and measurement of achievement. Nationally, Congress had reauthorized the Elementary and Secondary Education Act coined “No Child Left Behind.” The data around standardized test scores and a national focus on accountability influenced my thinking and decision making as a school and district leader. As an administrator in Wisconsin for the last eighteen years, I have participated in the implementation of increasing frequency of standardized tests to measure student achievement, and I have examined standardized test scores at great length. Aggregate test performance in my suburban district has been strong, and the schools continue to be ranked highly by test scores in the state. Disaggregated results painted a different picture, but given the relatively small numbers of students from many of the low performing demographic groups, that data was frequently given little thought.

For as long as there have been schools, educators have sought strategies to improve student outcomes. Improvement Science has emerged over the last three decades and purports to provide schools with a set of tools to systematically focus on growth over time. Bryk, Gomez, Grunow, and LeMahieu (2015) outline a detailed process to help schools by *Learning to Improve: How America’s Schools Can Get Better at Getting Better*. Improvement Science

focuses on recognizing problems through data; engaging teachers, students and parents in addressing problems within the school setting; developing strategies to make improvements; and measuring growth throughout the process. Improvement Science borrows tools, ideas, and principles from business under the name Total Quality Management and healthcare in which the focus has been called Continuous Improvement. The current use of Improvement Science in education has roots in these ideas, with some schools adopting Total Quality Management or Continuous Improvement in educational settings prior to the practice of Improvement Science.

Three years ago, a new superintendent entered the school district in which I work. As a leader, he had employed the principles of Improvement Science and believed in the value of the tools and strategies to continuously improve schools for the community. He engaged the leadership team in learning the principles and implementing tools and strategies. We began by embracing measurement and analyzing data to not only celebrate the bright spots, but also to identify problems in the system. We learned how to examine perceived problems to see the system that led to the current outcome. We set goals and reported the data in short cycles to see if changes made led to improvements.

As the leadership team worked to implement the principles of Improvement Science, I began to notice some tendencies in the process. During the early stages of implementation, the principle of measurement drove thinking towards standardized test scores. Test scores were numbers and perceived measurement of learning with precision. While I have since questioned the value of these standardized tests, I relied on test scores to provide information on the achievement of students in the system. Our district increased the frequency of standardized testing tools and focused on these scores to measure growth. Student scores remained consistent, with little statistically significant improvement. Nationally, the Improvement Science process

shows positive impact on outcomes in many cases, and “achievement and attainment trends on U.S. measures reflect a clear pattern of growth” (O’Day & Smith, 2016, p. 305). O’Day and Smith (2016) have studied Improvement Science with and through the Carnegie Foundation for the Advancement of Teaching. However, the same data set that shows overall growth also indicates some achievement gaps were narrowing only slightly while others remained unchanged (O’Day & Smith, 2016). School report cards across the state of Wisconsin document achievement gaps by demographic factors, including race, economic status, disability, and English language proficiency. The question becomes, for whom are schools working and improving? In my district, disparities in test scores for traditionally marginalized demographic groups became apparent as shifts in national accountability measures focused on disaggregated data.

These gaps became the focus of improvement work under the new superintendent. As a white, middle class, female who was raised speaking English in a Christian home, I had a narrow purview through which to examine the system that led to the current outcomes. My peers, all white leaders, raised in middle and upper class homes speaking English, had similar perspectives. Therefore, proposed strategies to improve test results focused on deficits of students. We didn’t question the basic structure of our school. We didn’t question the curriculum selections. We didn’t question the process of remediation. The system was working for the majority. While we sought input from students and teachers, the questions and ideas were all applied through the perspective of white, English speaking educators with financial means for whom the system of school had been effective. When school leaders identify achievement gaps in this data set, leaders may misinterpret and misuse the data to group and separate students (Jimerson, 2016). Furthermore, this practice disproportionately impacts groups of historically-

marginalized students negatively by isolating them for remediation (Jimerson, 2016; Park, Daly, & Guerra, 2012). My experiences mirror the evidence in the research as we developed solutions of intervention groups and pulled students out of class with frequency and consistency. This strategy did not improve learning for students of color, for students from low income homes, for students with disabilities, nor for students building English proficiency.

Our schools maintain strong performance in most indicators when considered on the aggregate, but the principles and practices of Improvement Science were not addressing the disparities in outcomes. Beginning in 2016, school and district leaders have read professional literature and engaged in professional conferences focused on equity and diversity. As white leaders, we engaged in discussions about race and oppression to develop strategies that may impact the disparities in outcomes for our students traditionally marginalized in the educational system. This learning has begun to change how we collectively develop strategies to improve and how we hear the voices of the people we serve. Our district had employed the principles of Improvement Science, but we hadn't considered the impact of race, economic status, disability, and English proficiency in our processes.

After decades of improvement work and reform in elementary and secondary education, inequality is still pervasive across many measures. I recognize these disparities in my own district, despite the use of Improvement Science principles. Bryk (2017) identifies disparities that persist in our schools resulting in predictable failures “based on the family a child might be born into or the community in which that child might live” (p. 3). Increasing our competency around the history of marginalization, appreciating the imbalance of power when listening to the voices of the people we serve, and understanding diverse perspectives improves how leaders in my district apply the tools of Improvement Science.

Experience with implementation of Improvement Science and expanding my understanding of how my identity impacts my leadership led me to consider the intersection of Improvement Science and equity research. As such, I sought to examine the questions, how has existing literature documented the employment of Improvement Science and what have leaders tried to improve? To what extent and in what ways is equity addressed within current literature about Improvement Science? In chapter 2, I will detail the principles of Improvement Science and explore what improvements schools have made as documented in the literature. Then, I will consider the research on equity in school improvement efforts and critique the treatment of equity in Improvement Science.

Chapter 2

Literature Review

I conducted a literature review to explore the implementation of Improvement Science in schools in the United States and the treatment of equity in Improvement Science. To do so, I searched ERIC and Education Research Complete relying on the search terms Total Quality Management, TQM, Continuous Improvement, and Improvement Science in combination with education, schools, equity, diversity, social justice, race, and disability. I confined my search to peer-reviewed scholarly journals. The results returned a total of 200 articles.

These studies included schools and colleges in the United States and around the world. I eliminated articles focusing on countries outside of the United States and articles in a college setting or another industry such as business or healthcare. Given that the populations of many countries around the world have greater homogeneity than the United States and social issues of equity are quite different, limiting the literature to the United States elementary and secondary education supports one of the questions I am investigating. To remain focused on strategies for school-wide improvement, I further eliminated articles that focused on a specific instructional strategy for teachers in specific subjects such as literacy or science. I determined that thirty of the articles focused on the use of continuous Improvement Science tools and leadership strategies in elementary or secondary schools within the United States. The search criteria and results are shown on Table 1.

Table 1

<i>Search Parameters and Results</i>		
Search criteria	Number of Articles	Focused on school-wide implementation in K – 12 education in the United States
“Improvement Science” and education and schools and “United States”	4	1
“Total Quality Management” and education and schools and “United States”	55	11
TQM and education and elementary and secondary	39	15 with 2 overlapping another search
“Improvement Science” and education and equity and “United States”	0	0
“Continuous Improvement” and education and schools and “United States”	81	11 with 6 overlapping a another search
“Continuous Improvement” and “United States” and equity	2	2 with 2 overlapping a another search
“Continuous Improvement” and “United States” and diversity	4	0
“Continuous Improvement” and “United States” and “social justice”	0	0
“Continuous Improvement” and “United States” and race	2	0
“Continuous Improvement” and “United States” and disability	9	0
“Total Quality Management” and “United States” and equity	1	0
“Total Quality Management” and “United States” and diversity	3	0
“Total Quality Management” and “United States” and “social justice”	0	0
“Total Quality Management” and “United States” and race	1	0
“Total Quality Management” and “United States” and disability	1	0
“Improvement Science” and “United States” and equity	0	0
“Improvement Science” and “United States” and diversity	0	0
“Improvement Science” and “United States” and “social justice”	0	0
“Improvement Science” and “United States” and race	0	0
“Improvement Science” and “United States” and disability	0	0

As expected, the Carnegie Foundation for the Advancement of Teaching surfaced within the articles. I then searched www.carnegiefoundation.org for the terms ‘equity,’ ‘diversity,’ ‘social justice,’ ‘race,’ and ‘disability.’ I filtered the search results by ‘publication’ and eliminated those that focused on college settings, resulting in the inclusion of two additional peer-reviewed publications in my analysis.

Researchers referenced several books and authors repeatedly. Most researchers identified Deming’s *The New Economics: For Industry, Government, and Education* (1993) as the foundation of definitions for TQM and Continuous Improvement. Several researchers relied on Bryk, Gomez, Grunow, and LeMahieu’s book *Learning to Improve: How America’s Schools Can Get Better at Getting Better* (2015) to define Continuous Improvement. I included these books in the literature review, given the historical references made in several of the articles focused on Improvement Science.

I analyzed these 32 peer-reviewed publications and 2 books focusing on three questions: 1) How does the author document and define the improvement process and improvements implemented within the study? 2) On what improvements have leaders focused in schools? 3) To what extent and in what ways is equity addressed within the study? Has existing literature documented ways schools use Improvement Science in ways likely to address equity and narrow disparities in outcomes?

Review of the Research: Findings

Drawing on the literature, I will first define Improvement Science through common principles and processes in all of the studies, considering the historical iterations of improvement processes to help define the current efforts. Then, I will highlight where leaders have focused improvement efforts within the studies. Finally, I will share the limited body of research on

equity within improvement efforts and critique the ways in which schools are not likely to address equity.

What is Improvement Science?

First, Lewis (2015) makes an important distinction between Improvement Science and experimental science, citing the difference of implementation. “Experimental science assumes scale-up occurs through faithful implementation of a proven program in new settings” (p. 55). Improvement Science, on the other hand, “assumes scale-up occurs through integration of basic knowledge ... to detect and learn from variations in practice, to build and share knowledge among practitioners, to motivate frontline innovators, and so forth” (p. 55). Improvement Science leans on the context of the specific, local environment to improve outcomes for the users of the system.

Lemire, Christie, and Inkelas (2017) believe practitioners lack consensus on a definition of improvement science “despite its growing popularity” (p. 23). They further articulate a level of entanglement with the concepts of “continuous improvement” and “systems improvement” (p. 23). Improvement Science borrows from common tools of similar ideas within continuous improvement, the work of Deming (Bryk, 2017; Lewis, 2015; Lemire et al, 2017), and pillars of the Baldrige Award for Quality Excellence (LeMahieu et al, 2017), but the practice of Improvement Science has evolved a unique process. Therefore, I define Improvement Science using shared processes and characteristics as well as a common foundation as cited by Bryk, Gomez, Grunow, and Lemahieu (2015), Bryk (2017), Lewis (2015), and Lemire et al. (2017).

Defining the Common Processes of Improvement Science

Bryk et al. (2015) outlined six principles to guide the disciplined inquiry process for continuous improvement, which is labeled Improvement Science:

- make the work problem-specific and user-centered
- focus on variation in performance
- see the system that produces the current outcome
- measure improvement
- anchor practice in disciplined inquiry
- accelerate learning through networked communities

I will focus on this iteration of Improvement Science to evaluate the shared characteristics, practices, and processes employed throughout the research.

All of the studies I examined documented some defined components, processes, and/ or tools used in Improvement Science. I will focus on the shared characteristics of user and problem focus, understanding variation in performance, seeing the school as a system, embracing measurement, and rapid cycles of disciplined inquiry for improvement (Bryk et al, 2015).

Problem-Specific and User-Centered. Bryk (2017) believes that the key to improvement is to shift from a program focus to a problem solving focus, which requires listening to and understanding the needs of the “user” or customer (p. 5). Listening to the voice of the customer is critical in generating solutions to problems in schools. Unlike experimental science, the problems are considered in context and students’, parents’, and teachers’ experiences with the curriculum, instruction, and other school processes guide improvement.

Bryk et al. (2015) turn to business practices for examples of how product manufacturers design new products by observing people doing work. More recently, Bryk (2017) explains that being user-centered is about listening to the people experiencing the problem in schools, citing a team of non-profit partners who spent many hours listening to people experiencing homelessness to improve the lives of homeless. Throughout the literature I reviewed, researchers rarely

described the process of listening and practice of collecting voices of the users and education examples are limited. When considering an improvement that focuses on equity and disparities in outcomes, this principle requires representation of the users most impacted by the gap in outcomes to understand “this disparity through the eyes, mindsets, and emotions of all involved” (Bryk, 2017, p. 6). Because many of the examples in the literature are silent on this principle, I do not know which user voices were considered in the improvement process. This silence speaks to the potential lack of focus on equity in the current body of literature.

Recognize Variation. By focusing on variation in performance, school leaders identify positive outliers for “high-leverage processes that can be detailed as ‘standard work’” (Bryk et al., 2015, p.46), meaning processes or steps that can be replicated to get similar positive results across the organization. Margiotta and McCannon, who were highlighted by Bryk (2017), explained what they learned from hours of listening to people who are homeless. Adaptation is more important than fidelity to a specific program. Successful improvement relies on understanding what is adaptable or variable in the local context.

Research-based programs promote a standard effect size for success based on an experimental science approach. When schools focus on programs, teachers and leaders expect fidelity to the implementation to result in the standard effect size. Fidelity measures the accuracy of key design elements for the program including training, materials, and specific processes (Bryk, 2017). However, every study on a program includes a range of results, even when there is fidelity to the program design elements. For example, Reading Recovery reports a standard effect size of 0.7 (Bryk, 2017).

To their credit, Reading Recovery actually looked at variability in program effects across school sites. Even with a very well detailed and supported program, as is Reading

Recovery, wide variability was found. In some places, Reading Recovery worked exceptionally well, but in other places, it did not work at all. (Bryk, 2017, p. 7)

The variability of outcomes could not be explained by the variability in implementation; therefore, context matters. Improvement Science recognizes the variability in results with a given program and seeks to understand the context that contributes to the variability.

Hattie (2012) synthesizes the research on instructional strategies and ranks the effect sizes of popular, research-based strategies and programs. The use of a standard effect size allows for comparison of a variety of strategies and programs, but the effect size does not guarantee the results (Bryk, 2017; Lewis, 2015). In a school employing Improvement Science, teachers and leaders recognize the variability in a research-based intervention based on context, and measure to determine local effect size and the variability of the impact (Lewis, 2015). Understanding which students and classrooms exceed the standard effect size and which perform below the standard effect size creates opportunity to learn about which specific strategies are most effective in each local context.

Embrace Measurement and See the System that Produces the Current Outcomes. In order to understand how the current system and processes contribute to the current outcome, Bryk et al. (2015) believe all staff must engage in building an understanding of the problem and identifying root causes. “Systems theory directly informs the science of improvement—an appreciation that all change takes place in the context of a dynamic and adaptive system, why understanding the system’s composition is a fundamental condition for improvement” (Lemire et al., 2017, p. 26). To build this understanding, Bryk et al. (2015) detail specific tools for a “Causal System Analysis” (p. 66). These tools include clearly defining the problem identified within the data, asking why questions, and considering a variety of factors that are contributing

the current outcomes perceived to be a problem. Bryk et al. (2015) suggests using “The Fishbone Diagram” as an option for illustrating primary factors and secondary factors contributing to the problem represented in Figure 1.

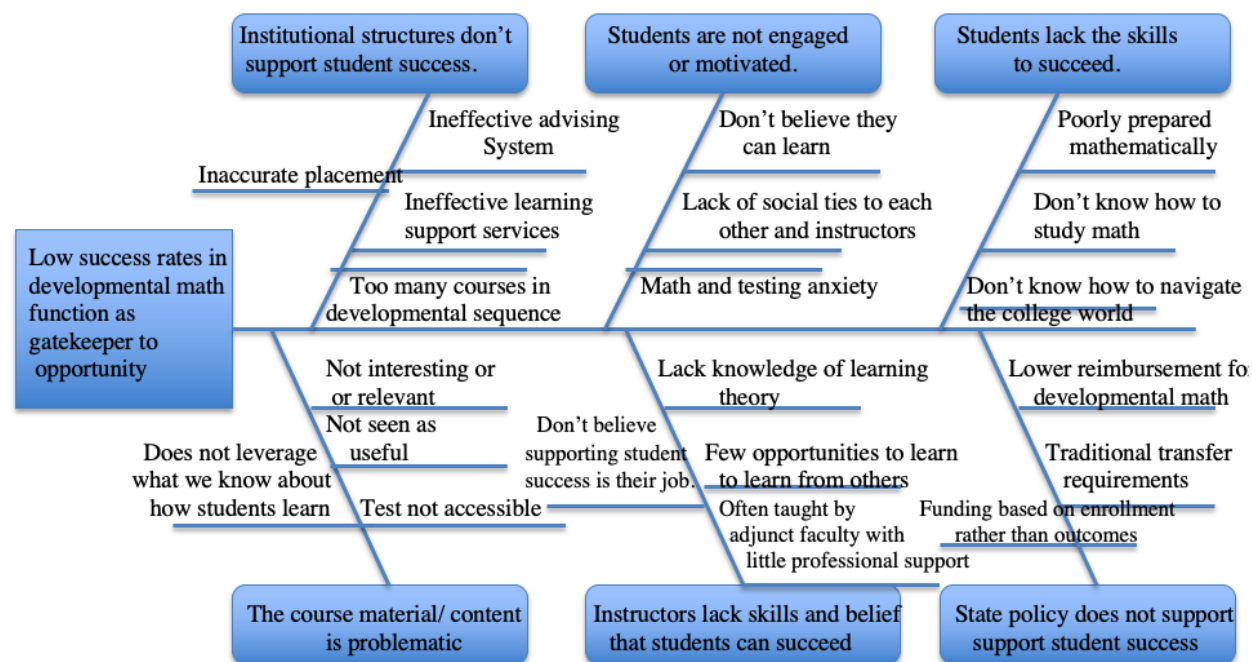


Figure 1. Fishbone diagram (Bryk et al., 2015, p. 68)

In several studies that authors highlighted as effective implementation of continuous improvement, leaders engaged stakeholders in attempting to understand the “root causes” of variations perceived to be a problem within the system as identified by the data (Blankstein, 1996; Bryk et al., 2015; Shargel, 1996). At George Westinghouse Vocational and Technical High School, in New York, New York, the principal included staff in identifying concerns and challenges in the school community. Among the 23 areas of concern, staff identified cutting class as a significant problem (Rappaport, 1996; Schargel, 1996). A student leadership class considered the concerns and worked with staff to identify root cause for students missing class and implement changes (Schargel, 1996). At Kate Sullivan Elementary School in Tallahassee, Florida, the principal worked with volunteers from Florida State University to develop a system

to collect, analyze, and widely share data so that any member of the school could identify a concern to define a potential problem (Blankstein, 1996, p. 27). In each case, the researcher remains silent on the demographic composition of the school, disparities in outcomes based on demographic markers, and on the demographic composition of the voices collected within the process of root cause analysis.

In this process, leaders investigate the differences from classroom to classroom and from school to school to understand how context and different factors impact the implementation of a practice or performance of students (Bryk et al., 2015). For example, the team may examine instructional methods, class placements, student schedules, or other processes that contribute to student performance on test scores (Bryk et al., 2015). Each of the studies that focused on a specific problem described measurement tools that focused on a narrow aspect of the problem, such as writing scores (Schmoker & Wilson, 1995), feedback cycles for teachers (Bryk et al., 2015) and daily attendance (Rappaport, 1996; Schargel, 1996, Schmoker & Wilson, 1995). While some studies highlighted “diverse population of students in terms of income, culture, and national heritage” (Steyn, 2006, p. 3), this diversity was not defined nor measured. In this study, improvements on the aggregate were assumed to be positive for all of the diverse learners. None of these studies addressed problems of inequity.

Schools and districts that have adopted a formal continuous improvement process require outcome data because “we cannot improve at scale what we cannot measure” (Bryk et al., 2015, p. 87). The organization focuses on “system outcomes for a defined population of beneficiaries and the processes that lead to those results” (O’Day & Smith, 2016, p. 315). Effective principals and school leaders consider multiple sources of data and then use tools to identify the causes to plan for improvement.

Standardized test results measure student achievement of common standards. Many schools have been forced to focus on standardized test results based on accountability requirements and have applied tools of continuous improvement to impact these measures. Sheboygan Area School District, Wisconsin routinely examined standardized test data, highlighting how the Continuous Improvement positively impacted test results (Westfall, Peltier, & Sheehan, 2005). Leaders at Fort Pitt Elementary in Pittsburgh, Pennsylvania, Centennial Elementary in Evans, Colorado, and Frederick County Schools in Maryland also focused improvement efforts on standardized test performance using tools of continuous improvement (Schmoker & Wilson, 1995). Winder-Barrow High School in Georgia influenced the process for improving test scores by attending to standardized test results using root cause analysis to focus on local context (Weller & Weller, 1998). In these studies, test scores represent quality measures for production and staff engaged in the use of causal analysis tools to understand the system leading to the existing results. Standardized test data used in each of these cases are disaggregated, but each study focused on problems in the aggregate. These studies highlight how this principle of Improvement Science is applied, but none addressed problems of equity within the data.

Which data leaders choose as the focus of improvement matters. Leaders who focus narrowly on standardized tests alone tend to lead improvement focused on easily-measured outcomes rather than on the problems that are most pressing in the organization (Park et al., 2012). These choices empower some data over other data and may silence problems of equity. Furthermore, Park et al. (2012) assert that compliance with accountability requirements empower test score data over other measures.

While many researchers identify test scores as a common data point measured among schools to determine problems as a focus of improvement work (Datnow & Park, 2018; Park et al, 2012; O'Day & Smith, 2016; Schmoker & Wilson, 1995; Weller & Weller, 1998; Westfall, Peltier, & Sheehan, 2005); some principals and school leaders routinely reviewed other data sources as suggested in the core principles of Improvement Science (Arif & Smiley, 2003; Bryk et al., 2015). Some of the studies I reviewed focused on teacher retention data for improvement efforts (Bryk et al, 2015). In Montgomery County, Maryland and Fresno, California, school leaders measured outcomes by participation in Advanced Placement and rigorous college prep coursework (O'Day & Smith, 2016). Stakeholders at Westinghouse Vocational and Technical School focused on attendance as a measure (Rappaport, 1996; Schargel, 1996; Schmoker & Wilson, 1995). Weller and McElwee (1997) addressed dropout rates and attendance to improve student achievement scores. In every study, schools examined data to measure positive outcomes or potential problems.

Disciplined Inquiry Through Plan, Do, Study, Act (PDSA) Cycles. Plan, Do, Study, Act describes the protocol implemented in schools and districts following Improvement Science. According to Bryk et al. (2015) educators plan for a change, implement the change, and collect data within a short time to be analyzed to determine if the change is an improvement, and then educators act on the analysis by making adjustments to the plan, beginning the cycle again. Lewis (2015) identifies as the core framework of improvement science the plan-do-study-act (PDSA) cycle, a process for rapid cycles of learning from practice, coupled with three fundamental questions that drive improvement work: 1) What are we trying to accomplish?; 2) How will we know that a change is an improvement?; 3) What change can we make that will result in improvement? (p. 23)

Seven studies specifically identified a Plan, Do, Study, Act (PDSA) cycle as a tool for improvement (Bryk et al., 2015; Cohen-Vogel et al., 2016; Datnow & Park, 2018; LeMahieu, Nordstrum, & Greco, 2017; O’Day & Smith, 2016; Schmoker & Wilson, 1995; Weller & McElwee, 1997). Several of the examples pre-date literature on Improvement Science, exploring a previous iteration labeled Total Quality Management. Many of the examples stem from research related to the Carnegie Foundation for the Advancement of Teaching. In the Austin (Texas) Independent School District, the improvement goal focused on increased teacher retention. In the PDSA cycle, the improvement focused on increasing and improving feedback for new teachers. The “short cycle” measurement was a three-question survey about the quantity of feedback and efficacy of the feedback (Bryk et al., 2015, p. 137). In Broward County, Florida and Fort Worth, Texas, district leaders encouraged an innovation process to document practices that were adding value to test scores through PDSA cycles (Cohen-Vogel et al., 2016). In Menomonee Falls, Wisconsin, teachers engaged in PDSA cycles for curriculum improvements (LeMahieu, Nordstrum, & Greco, 2017).

Evolution and History of Improvement Science

Addressing an audience at the Carnegie Foundation’s fourth Summit on Improvement in Education, Bryk (2017) recognized W. Edwards Deming as “the improvement guru whose ideas have inspired productive change across many different industries and sectors” (p. 2). Lewis (2015) and Lemire et al. (2017) acknowledge Deming’s influence in the development of Improvement Science, citing Langley et al. (2009). Deming introduced quality measures for management, which he applied as Total Quality Management in manufacturing to Japan (Arif & Smiley, 2003; Arif, Smiley, & Kulonda, 2005; Capper & Jamison, 1993; Kelso, 2012; LeMahieu, Nordstrum, & Greco, 2017; Weller & McElwee, 1997). In 1986, Deming

documented fourteen principles which guided his leadership in Japanese auto manufacturing (Arif & Smiley, 2003; Arif, Smiley, & Kulonda, 2005; Aumiller & Hackmann, 2009; Capper & Jamison, 1993; Kelso, 2012; LeMahieu, Nordstrum, & Greco, 2017; Weller & McElwee, 1997). These fourteen principles drove improvement in manufacturing and business for the past thirty years, inspired the Malcolm Baldrige Award for Quality Excellence (Arif & Smiley, 2003; Aumiller & Hackmann, 2009), and evolved in healthcare and education under the label continuous improvement (Bryk, Gomez, Grunow, & LeMahieu, 2015). To provide greater context for and further define Improvement Science, I describe Total Quality Management in manufacturing and the subsequent influence in educational accountability and quality improvement practices and processes.

TQM and Continuous Improvement from Business Concept to Healthcare and Education

Deming, influenced by his professor Walter Shewhart, refined a process for quality measurement (Kelso, 2012). Deming believed that “manufacturing would be improved through a focus on identifying and correcting problems during the manufacturing process” (Kelso, 2012, p.111). Deming applied his ideas to auto manufacturing in Japan following World War II, eventually labeling the idea “Total Quality Management” (Arif & Smiley, 2003; Kelso, 2012).

Manufacturers valued Deming’s success in improving quality and reducing costs and began to study the process (Arif & Smiley, 2003; Kelso, 2012). Deming’s fourteen documented principles guided the practice of Total Quality Management.

Langley et al. (2009) identify Deming’s ‘system of profound knowledge’ as the intellectual foundation for improvement science. Following Deming, a system of profound knowledge is structured around four types of knowledge: 1) Knowledge of systems; 2)

Knowledge of psychology; 3) Knowledge of variation; 4) Knowledge of how knowledge grows. (Lemire et al, 2017, p. 24)

U.S. businesses embraced these principles, launching the Total Quality movement in the 1980's (Arif & Smiley, 2003; Gehlberg, 2007; Jimerson & Reames, 2015) By 1987, Congress initiated the Baldrige Award to promote "quality awareness and recognize quality achievements of business organization and provide vehicles for sharing success" (Arif & Smiley, 2003, p. 741). The criteria for the Baldrige Award built upon the business values rooted in Deming's Total Quality Management system (Arif & Smiley, 2003).

As Total Quality Management was sweeping the business community, "there came a growing crescendo of criticism towards public schools from media reports, fundamentalist religious leaders, and public office seekers" leading to a series of commissioned reports on public education (Arif & Smiley, 2003, p. 750) "The influence of these reports had become the efficient and scientific approaches to K – 12 education programs of 'Goals 2000' legislation" (Arif & Smiley, 2003, p. 750). Citing Deming from 1991, Bryk (2017) characterized education as a field with "miracle goals without methods" (p. 2). Deming's influence led the U.S. government to extend the Baldrige awards to schools in 1999 to encourage focus on quality as a method to achieve the proposed Goals 2000 (Arif & Smiley, 2003; Aumiller & Hackmann, 2009). The Baldrige Award for Quality Excellence in Education championed "five pillars of 'Total Quality Schools,' including: Customer Focus, Total Involvement and Staff Development, Quality in Operational Results, Problem Prevention and Resolution, and Continuous Improvement" (Arif & Smiley, 2003, p. 742). It is common for schools practicing Improvement Science to apply for the Baldrige Award, and the Carnegie Foundation for the Advancement of Teaching highlights several recipients on the website.

Recognizing the complexity of schools and the challenges facing children and teachers every day (Gelberg, 2007) and seeing value in this process of improvement, the Carnegie Foundation for the Advancement of Teaching invested in researching and documenting a continuous improvement model that was influenced by the practices of Total Quality Management in schools and extended the practices into networks. In 2015, Bryk, Gomez, Grunow, and LeMahieu documented characteristics of continuous improvement in education in the book *Learning to Improve: How America's Schools Can Get Better at Getting Better*. They argued that “Educators across the country were confronting common challenges and trying to solve complex problems. Many learned valuable lessons that could have propelled” others forward toward success (p. 3). Motivated by a desire to improve education, the authors considered how the common process could support success across schools. According to Cohen-Vogel et al. (2016),

Improvement science has emerged to address this failure [to understand the conditions under which teaching and learning take place], focusing attention not only on the identification and development of effective practices, but also on the processes of implementing and adapting practices to particular contexts so they can be scaled within a system. (p. 2)

Bryk et al. (2015) define Improvement Science more broadly, referring to it as a “methodology that disciplines inquiries to improve practice” (p. 10). Improvement Science encompasses and extends the common processes frequently referenced as continuous improvement or quality improvement (Bryk et al., 2015).

Research Studies Focusing on School Use of the Tools of Improvement Science

Each study referenced a model of Improvement Science that employed at least one of the tools, and some highlighted specific aspects of the Improvement Science principles. No study articulated how each of the six principles unfolded in the context of the study.

Bryk et al. (2015) emphasize a shift from improvement efforts focused on program implementation to improvement that is problem-focused and user-centered. This principle engages the voice of the users by listening and understanding how they experience the problem. The studies in the literature provide little insight into the implementation of this practice. Only Westinghouse Vocational School documented consideration of customer needs through surveys and engaged the voice of the customer (Rappaport, 1996; Schargel, 1996). In other school districts and schools, many leaders assumed the needs of the users as improved test scores without seeking input. In Sheboygan, Wisconsin (Westfall, Pelitier, & Sheehan, 2005), Barrow, Georgia (Weller & Weller, 1998), several schools in New Mexico and Scott County, Kentucky (Steyn, 2006), Brooklyn, New York and Montgomery County, Maryland (Schmoker & Wilson, 1995), schools focused on test scores without seeking input from students or parents on the value of this improvement or the strategies for improvement. If leaders did seek input from users for strategies to improve test scores, the literature is silent as to how. This assumption of what the user wants does not align with the first core principle in Improvement Science, even in studies focused on the implementation of an iteration of continuous improvement.

Bryk (2017) believes Improvement Science has the potential to eliminate current inequities in education. These gaps in outcomes have been documented in test scores, graduation rates, and college attendance rates for specific demographic groups since *A Nation at Risk* in 1983 (Arif & Smiley, 2003; Steyn, 2006) through current school report cards mandated by the *Every Student Succeeds Act* (Datnow & Park, 2018; Park et al, 2012). These “disparities are

often associated with race and poverty... and include any place where our educational institutions systematically foreclose opportunities to some children where predictable failures occur year after year after year” (Bryk, 2017, p. 3). Bryk (2017) believes that employing the processes and tools of Improvement Science can reduce these disparities in outcomes. Because problems are identified through review of data, leaders who attend to disaggregated data and focus on problems of inequitable outcomes will impact disparities through Improvement Science.

Addressing Problems of Inequity Using a Continuous Improvement Approach. The studies focused on schools deeply engaged in formal processes are relatively silent on issues of equity. While federal accountability requires disaggregation of data by demographic groups for test scores, attendance, and graduation rates, most schools in studies focused on Improvement Science using test scores considered aggregated data. Only a couple examples in one study stemming from the Carnegie Foundation addressed inequitable outcomes for students based on race and poverty (O’Day & Smith, 2016). Furthermore, several researchers raise concerns that issues of equity have been ignored or mistreated in continuous improvement work generally (Datnow & Park, 2018; Gil & Kim, 2018; Jimerson, 2015; Park et al., 2012). Specifically, students not achieving proficiency on standardized tests are removed from the classroom for isolated interventions, and these practices disproportionately impact students of color (Datnow & Park, 2018; Jimerson, 2015). In addition, some researchers focus on the demographic perspective and potential bias of the adults in a school, which impacts what data and solutions are considered and which are not (Capper & Jamison, 1993; Gil & Kim, 2018). To that end, I will first articulate the researchers who addressed inequity through Improvement Science and then describe concerns shared by researchers with a critical theory perspective on capturing the

voices of traditionally marginalized groups in schools and the impact of the perspective and potential bias of the leaders on improvement efforts.

Predictable variation in performance based on demographics. O’Day and Smith (2016) acknowledge “residential segregation, poverty, low levels of parental education, and limited access to social supports” as impacting “students’ educational achievement” (p. 301). These long-standing disparities contribute to “attitudes toward the students” resulting in “low expectations” that “have been well documented” and leading to discriminatory application of policies (p. 304). O’Day and Smith (2016) highlight several school districts focusing on these inequities in achievement by employing a continuous improvement approach. In several cases, schools improved academic outcomes for students of color and students from low-income homes, focusing on the measurement of demographic proportionality of students enrolled in Advanced Placement (AP) and college prep courses.

In Montgomery County, Maryland, the superintendent implemented a continuous improvement approach to address gaps “between White students and their African-American and Hispanic counterparts” (O’Day & Smith, 2016, p. 318). Using a process to identify root causes, school leaders, in discussions at the district level, identified structural practices that impacted students of color at disproportionate rates, including prerequisites for advanced courses and disproportionate distribution of resources leading to disparate class sizes (O’Day & Smith, 2016). Similarly, schools in Fresno, California implemented a continuous improvement approach to address problems uncovered in “course-taking patterns and other opportunities for underserved students” (O’Day & Smith, 2016, p. 319). These school districts focused on equitable outcomes through Improvement Science; however, the study did not articulate how students or staff engaged in the improvement process. The causal analysis is described at a

leadership level. Bryk (2017) advises as you focus in on trying to address an educational inequity, bring the voices of the people who are most directly impacted into the conversation. Seek to understand the dynamics of this disparity through the eyes, mindsets, and emotions of all involved. (p. 6)

Both Fresno and Montgomery County made improvements, narrowing inequitable outcomes. O'Day and Smith (2016) remain silent on how leaders engaged and understood the voices of the users impacted most by the problem in the generation of solutions and the perceptions of the users impacted by the improvement.

An Equity Critique of Improvement Science

Lewis (2015) notes that the value of Improvement Science over experimental science lies in the context and the “profound knowledge” (p. 54) drawn from social sciences and understanding people. In school systems dominated by white educators and influenced by middle-class values, this profound knowledge may lack understanding of the experiences of students and parents of color, students and parents from low income homes, as well as other marginalized groups. According to the National Center for Education Statistics, white students represented forty-nine percent of students in public schools in the United States in 2015. At the same time, white teachers represented eighty percent of the faculty across the United States in the same year. Lewis and Diamond (2015) explain that history shapes the “parameters within which” students and families of color operate, highlighting how “long histories of formal and informal racism contribute to persistent racial stratification and play a role in what kinds of resources families have” (p. 166). In turn, these ideas influence how teachers, counselors, and school leaders “interact with and respond to these [black or Latina/o] youth and their families” (p. 166). Even when school leaders recognize the problems, political pressure from middle-class

white parents maintain white privilege in the system (Lewis & Diamond, 2015) and perpetuate inequity. While Bryk (2017) believes that using the process defined in Improvement Science closes the gap, improvement practices may still be biased and ignore conflicting messages from students, parents, and educators.

Voices of traditionally marginalized and questions of power. In the studies I analyzed, none focused on a problem of educational inequity in which the leaders brought in the voices of those impacted most. Capper and Jamison (1993) focus on potential problems of inequity with TQM by highlighting an anecdote from the use of a solution seeking process for “unproductive experiences” in early elementary school for students with disabilities (p. 29). In this case, the authors highlight that the process did not engage families or students with disabilities to surmise their experience with the perceived problem. Studies on Westinghouse Vocational School in New York identified the demographic composition of the school, but the impact of the improvement was not disaggregated by racial or income groups. In addition, inclusion of voices of the users did not focus on problems of equity (Rappaport, 1996; Shargel, 1996). According to O’Day and Smith (2016), Improvement Science will only impact disparate outcomes for students from traditionally marginalized groups when inequity is the focus and conditions in the system allow for input from those directly impacted by the equity problem.

As Capper and Jamison (1993) cautioned in their study of Total Quality Management applied to schools, failure to listen to and understand the perspective of members of traditionally marginalized groups, students, and/ or families could result in a biased identification of root causes and potential solutions that further marginalize some of the students targeted with the improvement process. Because many of the studies highlighted the use of test scores and other data tied to state and federal accountability as the focus of improvement efforts in schools

employing Improvement Science, I concluded that some schools are motivated to address equitable outcomes based on standardized test results and public reporting. Gil and Kim (2018) affirm, “Equity of voice and equally valuing all stakeholders is not a reality everywhere ... thereby reducing the likelihood that the norms, values, and contexts of underrepresented communities are actually embedded in the development of what is counted in accountability” (p. 3) and subsequent improvement efforts.

Datnow and Park (2018) further argue that schools marginalize some demographic groups not only in the solution-generating process but also in the manner in “which data are prioritized” or interpreted by being “narrowly focused on highlighting student achievement gaps” (p. 132) when applying a continuous improvement approach. In defining the problem, students from low-income homes and students of color are typically labeled “as defective and lacking” (p. 133), which impacts the perceptions of students, teachers, and parents. Capper and Frattura (2007) suggest this practice sends the message that students who are isolated for instruction are inferior and less capable people. In turn, this practice reduces expectations and forecloses opportunities for rigorous course work leading to college and employment opportunities. Jimerson (2016) highlights schools that are “analyzing test scores and identifying ‘bubble students’ or assigning children to educational triage in preparation for the next exam” (p.62). Not only does this practice remove students from the regular classroom instruction with social interactions for targeted interventions for a portion of the school day, it also contributes to the message that rigorous course work, such as Advanced Placement (AP), is too hard for students from traditionally marginalized racial groups (Lewis & Diamond, 2015). In these studies, exclusionary practices alienate students and separate students of color and students living in poverty at disproportionate rates (Datnow & Park, 2018; Jimerson, 2016).

Related to issues of equity, when analyzing a problem illuminated by data, a team of stakeholders may develop solutions that are rooted with bias. Rose and Roschester (2008) concur that “teachers and administrators may have conflicts of interest that produce a bias in how new policies are evaluated” (p.446). Furthermore, new innovations are limited by the current knowledge and understanding of the teachers within the system. Capper and Jamison (1993) describe a problem-solving meeting in which one of the authors was involved. In reflection, the teachers’ potential solutions may not question or consider certain causes:

For example, without students and parents on the teams, and without the explicit inclusion of those who represent students who typically struggle in school, we did not question the basic structure of our programs, or our training or professional knowledge. (p. 29)

As a result, the team failed to fully “see the system that produces the current outcomes” (Bryk et al., 2015, p. 57), which is a critical component of Improvement Science. Therefore, improvements may perpetuate the status quo because the voices that are elevated in the generation of root causes and proposed improvements are generated through a lens that does not make explicit the perspectives of those making these decisions and the perspectives that are missing in these decisions (Capper, 2019). Apple’s (2012) work in critical theory in education reminds us that any educational improvement effort needs to be questioned from the perspective of: from whom did these efforts come; in whose interest are they; and who will benefit? These questions are essential to the study of Improvement Science.

Dominant power structures may limit parents and students from traditionally marginalized groups from expressing their understanding of the problems (Capper & Jamison, 1993). Within a school, the application of an Improvement Science approach may elevate the

voices of leaders and teachers while minimizing voices of students and parents, even when input is sought in the problem-solving process. Capper and Jamison (1993) explain how historical school management structures influenced by power structures based on race potentially inhibits participation of traditionally marginalized parents, students, and teachers. Datnow and Park (2018) concur, offering an understanding of how “power issues and other social justice concerns are features of the data use process” (p. 132).

Capper and Jamison (1993), from a critical theory perspective, question if continuous improvement is focused on “social justice and greater educational equity for those who have typically been disempowered” (p. 27). The concerns expressed by Datnow and Park (2018), Gil and Kim (2018), and Jimerson (2016) suggest that Improvement Science principles do not inherently address equity in meaningful ways. O’Day and Smith (2016) argue that Improvement Science provides an avenue for leaders focused on problems of equity in the data to address gaps in performance for racial groups. The studies offer little evidence that school leaders consider Improvement Science principles to address equity problems, and those who do may minimize the first principle.

Summary of Literature Review

I identified three themes within the literature I reviewed. First, all of the studies focused on implementation of at least one principle of Improvement Science practice. However, no single study outlines the full implementation of each principle and may omit one, two or three of the principles. Therefore, specific application of tools and processes for one or more of the principles are assumed by naming Total Quality Management, Continuous Improvement, or Improvement Science. Therefore, school leaders may employ the components of Improvement Science inconsistently or in convenient bits and pieces. Researchers documented limited aspects

of the process in the studies. O'Day and Smith (2016) clearly documented five features of the implementation of a continuous improvement approach which loosely aligned to the six principles of Improvement Science. Numerous studies reference Deming's principles or TQM (Arif & Smiley, 2003; Aumiller & Hackmann, 2009; Capper & Jamison, 1993; Gelberg, 2007; Greenbaum, 1996; Hirumi & Kaufman, 1992; Jimerson, 2016; Kelso, 2012; LeMahieu et al, 2017; Rappaport, 1996; Schargel, 1996; Schmoker & Wilson, 1995; Steyn, 2006; Weller & McElwee, 1997; Weller & Weller, 1998), and imply adherence to the practices. Only one study considers the lens of the leaders in implementing the practices (Capper & Jamison, 1993). In a future study, it will be helpful to focus on the different interpretations of Improvement Science based on a leader's epistemological lens and how it impacts implementation.

Second, the empirical studies on specific cases of implementation of Improvement Science were muted, if not silent, on issues of equity. Demographics play an important role on improvement efforts, specifically those focused on equity. In the analysis of Westinghouse, Shargel (1996) identified demographic composition by race; however, he failed to analyze the implications of the demographic composition on the improvement process. Only the researchers in the study of Mesa High School (Park et al., 2012) consider the impact of equity or demographics on the continuous improvement process. O'Day and Smith (2016) highlighted several school districts that improved problems identified in equity; however, despite outlining the shared features of Improvement Science, they omitted a full description of the methodologies employed to engage the voices of the students and families in these specific cases and how leaders engaged the voices of the impacted groups of students.

Finally, bias impacts the improvement effort most significantly in the definition of problems and root cause analysis of a problem, resulting in an improvement plan that ignores the

demographic composition, needs of the school community, as well as the values of underrepresented voices (Capper & Jamison, 1993; Datnow & Park, 2018; Jimerson, 2016). The studies not only omit the demographic composition of the schools, in most cases, but also ignore naming the stakeholders included in the improvement efforts, exploring how traditionally marginalized voices are heard in the continuous improvement process. Silence on demographic composition does not imply inequity; however, it indicates equity was not a focus of improvement. Ignoring marginalized voices does not necessarily lead to inequitable outcomes, but it could point to further empowering the dominant race and class interests in the improvement efforts. I will focus a study on two districts' application of Improvement Science when explicitly addressing issues of equity to build an understanding of how voices of traditionally marginalized groups are included and how bias has been addressed.

Conceptual Framework

The findings in the literature review informed the conceptual framework for the proposed study. First, I focused on the framework for Improvement Science. Schools implementing Improvement Science share common processes. Bryk et al. (2015) documented six core principles that mirror components of other improvement frameworks including Total Quality Management and the criteria for Baldrige Award for Quality in Education. I have employed the first five of the principles, excluding the last principle around Networked Improvement Communities, to form a conceptual framework of the implementation of Improvement Science and represent the components of Improvement Science as defined by Bryk et al. (2015) in the rectangular boxes in Figure 2.

Next, I considered the conclusions from the literature review focused on problems of inequity. Capper and Jamison (1993) and Gil and Kim (2018) expressed concerns about how

voices of traditionally marginalized groups were heard and valued in the process of identifying needs and addressing problems. O'Day and Smith (2016) noted that equity is impacted only when it is the focus of the improvement process. Datnow and Park (2018) and Jimerson (2016) identified concerns about the impact of labeling and isolating students, and separated students for remediation when schools address variation in performance. These concerns point to a need to consider Improvement Science through the lens of critical theory (Capper, 2019).

Critical theory “pivots upon relationships of power—who has power, who does not— and assumes the presence of suffering and oppression in organizations” (Capper, 2019, p. 69). Oppression stems from demographic variations, which may include gender, race, class, language, and disability. As Datnow and Park (2018) and Jimerson (2016) identified, oppression in schools may manifest in isolation and exclusion. Lewis and Diamond (2015) document bias that continues to illustrate academic oppression by race impacting course taking patterns and grades in a racially diverse high school. One form of oppression stemmed from the power or perceived power of parents of white students and teacher recommendations for challenging college level courses (Lewis & Diamond, 2015). Capper and Frattura (2007) address leadership strategies by which leaders recognize and acknowledge oppression in the system, employing specific tools. Leaders for social justice disaggregate data to uncover patterns in outcomes such as achievement, discipline rates, course participation, student engagement and family engagement (Capper & Frattura, 2007). Additionally, social justice leaders who subscribe to critical theory leadership address discrepancies by developing an understanding and educating staff in practices that minimize oppression, isolation, and exclusion (Capper & Frattura, 2007). I represented each of these social justice ideas about equity and the power differential for historically marginalized and

oppressed groups in schools. I connected each of these ideas about equity as arrows matched to principles of Improvement Science within the conceptual framework represented in Figure 2.

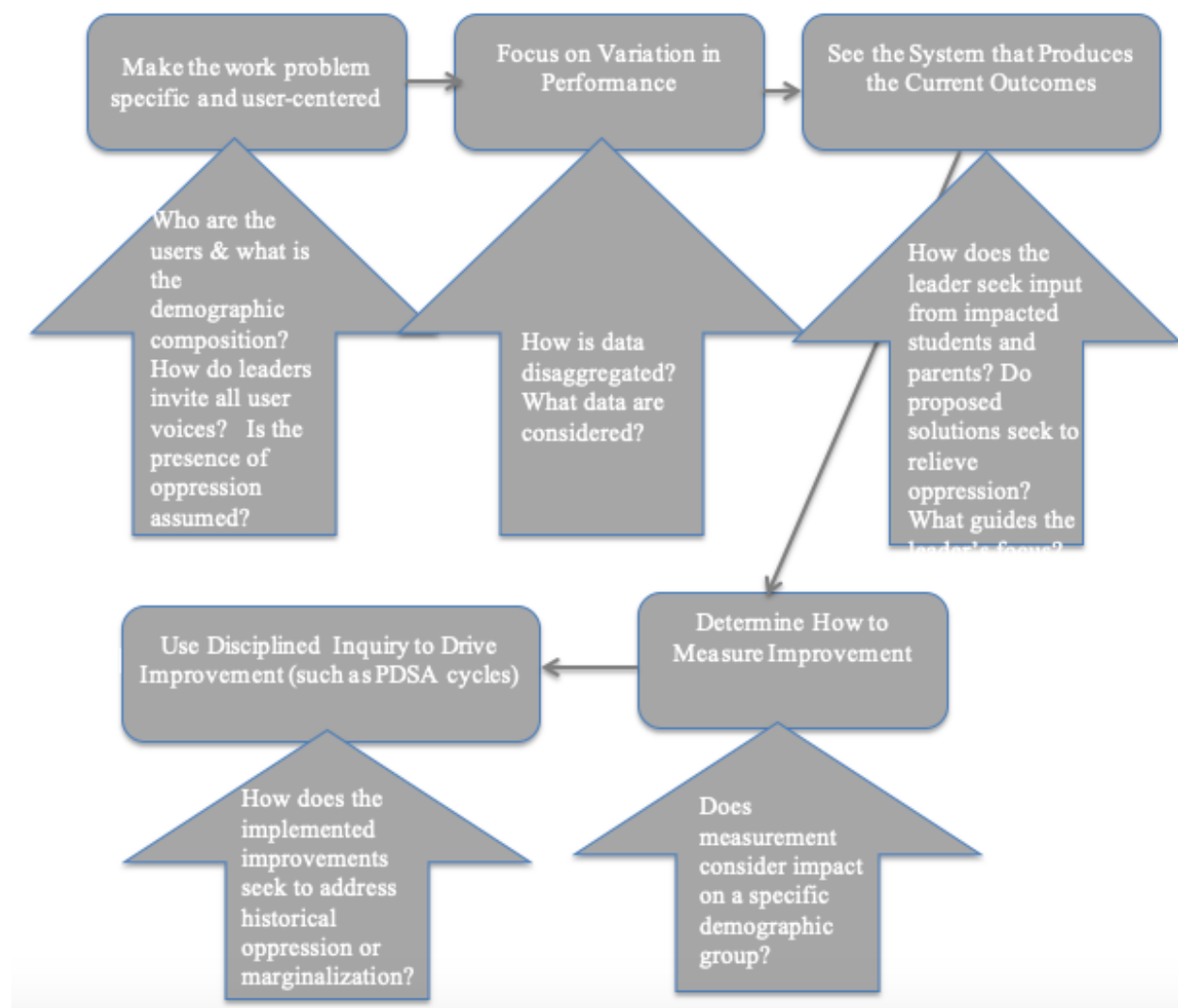


Figure 2. Conceptual Framework Improvement Science Informed through an Equity Lens

Research Question

To address the gaps in the literature, the following research question will guide the study: In schools using Improvement Science, how and to what extent do school leaders attend to diverse populations and equity when implementing the improvement process? Under what conditions do problems of inequity rise to a level of problem identification in which schools focus the principles of Improvement Science?

Chapter 3

Methods

In this section, I will detail the design of the study to answer the research questions. I will document the participants as well as the procedures. Then, I will provide an overview of the instrumentation followed by the methods of analysis.

Design

In addressing my research questions, I used a qualitative methods approach. Cresswell (2013) explains that

Qualitative researchers try to develop a complex picture of the problem or issue under study. This involves reporting multiple perspectives, identifying the many factors involved in a situation, and generally sketching the larger picture that emerges.

Researchers are bound not by the tight cause-and-effect relationships among factors, but rather by identifying the complex interactions of factors in any situation. (p. 47)

I concluded from the literature that the application of Improvement Science involves complexity based on the context and the identity of the users in education. I believe I was best able to address this complexity with a qualitative approach.

According to Marshall and Rossman (2016), a qualitative approach provides researchers the most valuable strategy to:

- “elicit multiple constructed realities, studied holistically” (p. 152);
- elicit “tacit knowledge and subjective understandings and interpretations” (p. 152);
- “delve in depth into complexities and processes” (p. 152);
- and “explore novel, ignored, or often marginalized populations” (p. 152).

A qualitative approach allowed me to describe and explore “context, setting, and participant frames of reference” (Marshall & Rossman, 2016, p. 153). Given the nature of my research question, these features provide the most appropriate method.

Marshall and Rossman (2016) identify six major genres for qualitative research that include ethnographies, phenomenological, sociolinguistic, grounded theory, case study, and arts-informed/ multimodal inquiry. Cresswell (2013) considers five specific methods of qualitative research, which include narrative research, phenomenology, grounded theory, ethnographic research, and case study. I used a case study approach which “explores an issue or problem, and a detailed understanding emerges from examining a case or several cases” (Cresswell, 2013, p. 123). Identifying a case or cases employing Improvement Science provided an “in-depth picture of it” (Cresswell, 2013, p. 123). I capitalized on the ability to “stress the importance of context, setting, and participants’ frames of reference” (Marshall & Rossman, 2016, p. 153) in implementing Improvement Science.

A case study approach offers several options: an intrinsic case, an instrumental case, a multiple case study or comparative case study (Marshall & Rossman, 2016; Cresswell, 2013). I selected a comparative case study in which I focused on the idea of equity in Improvement Science. A comparative case study “selects multiple case studies to illustrate the issue” (Cresswell, 2013, p. 99). This approach offered “the logic of replication” (Cresswell, 2013, p. 99), which afforded better opportunities to generalize findings. I capitalized on the opportunity to explore two instances of the phenomenon of an equity focus in the implementation of Improvement Science.

Participants

I used purposeful sampling as described by Cresswell (2013) to identify two public school districts deeply engaged in Improvement Science. I selected districts that have applied for the Malcolm Baldrige Award or comparable state level award and/or presented experiences with Improvement Science at a state, regional, or national level conference or in a professional publication, relying on the reputation of the district as an exemplar of Improvement Science implementation within the regional and national network of Improvement Science. Studying two districts noted for strong Improvement Science implementation provided the opportunity to understand these districts' equity efforts in relatively optimal conditions. Then, I searched publicly available enrollment data on the website of the state department of education for demographic distributions of race, students receiving free or reduced lunch, students labeled English learners, and students with disabilities for each of the identified districts. I identified two districts from the group with a diverse population as defined by at least 20 percent of students of color, students receiving free or reduced lunch, students with disabilities, and/or students labeled English learners; that is, 20 percent or more for any one or any combination of these. By selecting districts with at least 20 percent of students from traditionally marginalized groups, I ensured that there are sufficient numbers of students to examine publicly reported data for these disaggregated groups as part of the study on impact for improvement efforts. From this group of schools, I called a leader in identified districts and asked three questions: 1) Tell me a story about a problem you are trying to solve; 2) How did you identify that problem; and 3) Who did you ask about the problem? As such, I was able to identify a district with a documented and explicit goal focused on equity which has engaged the voices of a historically marginalized group to identify and address the problem and one district with a problem of equity who has not

asked the impacted group about the problem to provide a comparison. I requested consent from each superintendent to participate in the study.

In each school district, I asked the superintendent to identify one school that meets three criteria. First, the identified schools are engaged in “Disciplined Inquiry to Drive Improvement” (Bryk et al, 2015), and as such, the principal should have documentation of Plan, Do, Study, Act (PDSA) cycles (Bryk et al, 2015; LeMahieu et al, 2017). Second, the school engaged in a continuous improvement approach for at least three years to ensure the school has fully implemented the practices. Finally, the student population included at least 20% of one or more of the following populations: students of color, students receiving free or reduced lunch, students with disabilities, and/ or students labeled English Learners. Given the research questions address the role of equity, I needed a reportable number of students from one of these groups to consider potential impact on equitable achievement outcomes.

School and District Administrators. Within each of the schools, I interviewed multiple school and district administrators. At the district level, I included the Superintendent, the district administrator who oversees curriculum, the district administrator responsible for pupil services, the Business Manager, and the district administrator who takes the lead for Human Resources. At the school level, I interviewed the building principal and assistant principals. For each site, I interviewed between five and eight administrators.

Teachers. I requested names of professional staff from each administrator, whom have led a goal cycle focused that considered equity. I interviewed some staff individually and spoke with a focus group to provide opportunities for collective reflection on the experience. When requesting names of teachers from the leaders, I sought nominations of teachers who have been reluctant or critical of the implementation. I also invited participants in the focus group to

contact me for an individual interview if he/ she has ideas to offer that he/ she felt needed clarification. None of the focus group participants elected to follow up with an individual interview.

Data Collection Procedures

To address the research questions, I interviewed the participants, beginning with the superintendent at each site, followed by other district office administrators, principals, assistant principals, and then teacher leaders. I scheduled the interviews using email, identifying a time, location, and communication platform. I conducted some interviews in person, and due to COVID restrictions, some interviews were conducted via video call. To ensure an accurate accounting of the interview, I recorded the interview for later transcription.

I followed all protocols of the Internal Review Board. Prior to interviewing each participant, I emailed a consent form for review and request and receive signed consent in person or via email. I requested contact information from each participant and determined each participant's preferred method of communication for follow up. Following the interviews, I sent my emerging interpretations and themes to select participants (i.e., school or district leader, dissenting teacher, members of traditionally marginalized groups, etc.) by email. To ensure accuracy, I asked participants to email me clarifications each believes is necessary to best represent their perspective and to ensure trustworthiness as a member check. In one of the districts, I interviewed the recently departed former superintendent and shared my emerging interpretations at the end of the interview in a video conversation as a member check.

Anticipating that leaders will cite specific communications, professional development events, data collection tools, and documentation for Plan, Do, Study, Act cycles, I requested documents referenced by the participant during the interview. Some participants shared

documents for review in person and some shared via email following the interview. Focusing on the conceptual framework, I reviewed the documents, looking for relevant insights and correlating the ideas in the interview with documents within the school and district. The documents illuminated some changes in focus and attitude towards equity and provided a source of validity when analyzing interview responses in conjunction with work product. At one site, I was also able to request an opportunity to observe a student meeting in which equity was being discussed.

Finally, I reviewed and analyzed publicly reported student achievement and participation data using the school report cards published by the state, in addition to other engagement data reported and disaggregated in the state reporting system. I analyzed trends in the data comparing race, economic status, disability status, and language proficiency status.

Instruments. For the interviews, I followed the Administrator Interview Protocol (Appendix A) and the Teacher Focus Group Protocol (Appendix B) as dictated by the role of the participant in the school. I conducted semi-structured interviews as described by Merriam (2009). This protocol provides consistency with common scripted questions and also allows opportunities to clarify questions and ask follow up questions of the participant.

To guide the interview with each administrator, I developed the Administrator Interview Protocol (Appendix A). This protocol begins with an introduction to the research and several background questions about the school and the administrator's history in the district. Then, I developed questions focused on the implementation and components of Improvement Science included in the conceptual framework. In the next section of the interview protocol, I focused on equity training in the district, if any, and the impact of an equity focus on the use of Improvement Science in the school. Specifically, I designed questions to understand the extent

to which data is disaggregated and how disparities in any disaggregated data were addressed and measured. Finally, I developed questions focusing on the impact of equity focus through an Improvement Science lens.

For the teacher focus group at each school, I have developed the Teacher Focus Group Protocol (Appendix B). Similar to the principal interview, I began with an overview of the focus of the research and will ask several background questions about the school and the teachers in the school. I asked similar questions about the implementation and components of Improvement Science. Following this section, I shifted focus to equity, asking questions about the impact of learning about equity and diversity on the implementation of Improvement Science and the impact of equity focus on groups of students when subjects of an improvement process.

In reviewing documents referenced in the interviews and provided by the participants, I used the Document Analysis Protocol (Appendix C). I identified two lenses through which I reviewed each document. First, I considered evidence of components of the Improvement Science framework, including: Who are the users? What data is being considered? How did the author “see the system” (Bryk et al, 2015, p. 57)? What was the impact of the improvement effort? Next, I examined the document for evidence of an equity focus. In this review, I sought to find evidence of focus on disaggregated demographic groups and the nature of the treatment of students within the improvement process. Most importantly, is there evidence of isolation or inclusion of the students who are experiencing the problem?

After completing the interviews, focus group and document analysis at a site, I searched the state department of education for publicly available outcome data. I reviewed achievement data including disparities in proficient performance on the state assessment by race, economic status, disability status, and English language proficiency status. I located and documented

disparities in attendance, graduation rates, and discipline rates by race and economic status. I also reviewed opportunity data, including disparities in specific college and career ready programs, such as Advanced Placement participation and success rates or participation in school activities by the same demographic groups. Finally, I located data that relates to specific improvement efforts cited in interviews, focus group, and/ or a document analysis. I charted the data on the State Data Documentation Table (Appendix D).

Analysis

I recorded and transcribed each individual interview and focus group interview. I actively listened and reserved time after each interview to make notes and reflect on what I heard. During the interview, I made field notes on computer word processing software to document my emerging ideas and to serve as back up for the digital recording. I documented the role of the participant, the location of the participant, the date of the interview, and the experience level of the participant. In addition, I developed a pseudonym for each participant to maintain his or her anonymity. I used each protocol as a template to draft my field notes. I asked follow up questions to clarify my understanding of the participants' responses and maintain focus within the scope of the study (Merriam, 2009). Then, I named and stored my field notes in a folder by site. I employed online qualitative data analysis software to assist in the retrieval, coding, and determination of emerging themes (Marshall & Rossman, 2016; Merriam, 2009).

I began analysis after each interview, focus group and document analysis. Immediately following each interview, I reserved at least an hour to review my field notes and document ideas that I missed during the interview. I focused on the development of categories that capture themes in the data, including notes on methods and interpretations (Merriam, 2009). I developed

a coding system for the categories (Cresswell, 2013; Marshall & Rossman, 2016; Merriam, 2009) and cataloged interviews, focus groups, document observations, and research notes. Based on the literature review, I anticipated coding for fidelity to the Improvement Science framework, demographic considerations within the improvement process, and empowerment of historically marginalized voices within the improvement process. Specific subcategories were used from the conceptual framework as shown on Table 2. I was able to maintain organization by employing the coding themes in the margins within the online qualitative data analysis software. I also made notes on adjustments to my methods in future interviews as well as notes regarding emerging theories as I made comparisons.

Table 2

Coding for Ideas and Themes

Conceptual Framework Idea	Letter Code
User-Centered (identification of perceived user)	PU
Problem-Specific focus	PS
Focus on Variation in Outcomes	V
See the System (Root Cause Analysis)	RCA
Measurement (Standardized Test)	MST
Measurement (Aggregate)	Mag
Measurement (Disaggregated)	MD
Measurement (Data beyond testing)	DDM
Disciplined Inquiry to Improve	PDSA
Assumed Oppression	AO
Representative sample of race when listening to users	RR

Representative sample of economic status when listening to users	REcon
Representative sample of languages when listening to users	RLang
Root cause focuses on perspective of group impacted most by the problem	GP
Improvements impacted opportunity gap	OGap
Improvements impacted achievement gap	AGap
Improvements impacted satisfaction ratings/ sense of belonging survey ratings	Rate

I maintained a section of my field notes within the software to index the interviews and note the coding contained within each interview, focus group and document observation. I evolved the categories throughout the study to ensure all new ideas were considered and encompassed within the categories, which I continued to narrow and combine through the constant comparative method (Glaser, 1965; Marshall & Rossman, 2016; Merriam, 2009). Glaser (1965) described four stages in this constant comparative method of analysis, which include: “(1) comparing incidents applicable to each category, (2) integrating categories and their properties, (3) delimiting the theory, and (4) writing the theory” (p. 439). As I coded my field notes following each interview, focus group, and document observation, I followed the first and the second stage.

As I completed data collection at a site, I again reviewed all my field notes from the site with the assistance of the qualitative data analysis software. At this point, I had begun to develop theories based on the integration of categories and the saturation of incidents and examples within a category (Glaser, 1965). I used these emerging theories to develop major themes and write section headings (Glaser, 1965) before repeating the process at the next site.

Trustworthiness and Validity

I approached this study through the lens of a district-level administrator in a district in the third year of implementation of an Improvement Science framework. This positionality may have influenced my perception of the data. As the implementation began, I had concerns that focusing on short cycles of improvement or PDSA cycles may lead teachers to shift focus from quality instruction, relationships, or engaging experiences to test scores because they are easier to measure. My perspective influenced how I interpreted data and made meaning of the participants' responses. Reflection on this positionality (Merriam, 2009) and other tools for validity and trustworthiness helped to manage these potential biases.

Citing Lincoln and Guba (1985), Marshall and Rossman (2016) identify four strategies to establish validity. Lincoln and Guba (1985)

...urged qualitative researchers to be in the setting for a long period of time (prolonged engagement); share data and interpretations with participants (member checks); triangulate by gathering data from multiple sources, through multiple methods, and using multiple theoretical lenses; and discuss emergent findings with critical friends to ensure that analyses are grounded in the data (peer debriefing) (p. 84).

I used three of these strategies to ensure validity of the data and interpretations.

I sought to triangulate the findings by using multiple sources. I selected to use two cases to include multiple voices in the data. I had four different sources within each site, including interviews with administrators, focus groups with teachers, document analysis identified by the participants, and state-reported attendance, discipline, and performance disaggregated data. I

compared the ideas shared within each source to each other to reflect on consistency (Merriam, 2009), and I compared ideas between participants and across data sources to triangulate data.

I conducted member checks (Marshall & Rossman, 2016; Merriam, 2009) as another strategy to ensure trustworthiness and validity. Following the transcription of the interview, I provided a brief summary and my interpretation and categorization of ideas that are emerging with select participants at that time. According to Merriam (2009), participants will affirm the ideas as they recognize them. By sharing the data with the participant, I offered the participant the opportunity to affirm, adjust, or refute the interpretation of the interview. Marshall and Rossman (2016) refer to this as an opportunity to “ask the participants whether [I] ‘got it right’” (p. 328). In this process, I sought disconfirming ideas and follow up with interviewees who questioned the process or implementation of Improvement Science.

As I completed data collection at each site, I engaged in peer debriefing (Marshall & Rossman, 2016). To do so, I discussed my findings and ideas with my advisor at the University of Wisconsin – Madison. This process provided an opportunity to talk “through logic and clarity of [my] interpretations” and to ensure my understanding is grounded in the data (Marshall & Rossman, 2016, p. 328). At one point, I found myself filling in the gaps and assuming meaning when administrators used specific vocabulary associated with Improvement Science. Following a peer debriefing, I focused on what was said and what was not said in the interview to adjust my interpretations.

Ethical Considerations

Marshall and Rossman (2016) believe that a researcher “must think beyond careful procedural matters and documentation for the protection of human subjects” (p. 93). To this end, researchers should build trust with the participant, gain informed consent, respect the privacy of

the participant, and ensure that “participants are not harmed by participating in the study” (p. 95). Prior to beginning the study, I requested and received approval from the Internal Review Board (IRB) at the University of Wisconsin – Madison. I also requested and received site permission from the superintendent.

Next, I provided each participant with written information on the purpose of the study and the potential risks to the participant and gain their informed consent (Marshall & Rossman, 2016). I respected and understood that some participants have a power differential that may present a risk to the participant (Marshall & Rossman, 2016). The Internal Review Board specifically cited another potential vulnerability for perspective subjects in this study. Teachers and building administrators may have Institutional Vulnerability, causing them to be coerced or unduly influenced to participate given the power of the district administrator based on the employment and evaluation relationship. I maintained confidentiality of the identities of the participants by assigning pseudonyms for the people and the site. Within the focus groups, I established ground rules around confidentiality.

Pilot of Administrator Interview Protocol

In June 2019, I identified a superintendent in a school district actively implementing an Improvement Science approach with identified goals around equity. He implemented Improvement Science in a previous district, initiated implementation in his current district, and has spoken at state and regional conferences on the implementation of Improvement Science. In addition, he has engaged in equity work as part of his current district over the last three years.

I began the interview by providing a brief overview of the purpose of my study, attempting to limit details that may influence his responses and thus reducing the impact of

socially desirable bias. I gained his consent to participate in a pilot of the Administrator Interview Protocol (Appendix A).

Many of the responses aligned with my expectations based on literature and experience with implementation of Improvement Science. The subject provided examples of PDSA cycles and Improvement Science work in the district. He highlighted data sources and experiences consistent with the conceptual framework. We were engaged with the interview questions for approximately 50 minutes, which aligned with the expected time frame.

Following the interview, I engaged in a discussion about what worked well and suggestions for adjustments to the questions. The superintendent suggested two changes, which I made, to the protocol. First, in the questions regarding the implementation of Improvement Science, he suggested that I adjust “How was it decided that your school would employ an improvement science method?” to ask, “What does Improvement Science look like in your school/ district?” This change focuses more on describing the context and allows for a more robust understanding. Secondly, rather than asking about “users” as defined in the research, I could shift the question to ask, “who do you serve?” This targets the same concept but will elicit a better understanding of the perceived users. Finally, the superintendent suggested shifting questions around data collection to read “How are data collected and shared in your school/ district?” This additional question helps to understand the use of data within Improvement Science in the case.

In the second section of the interview protocol focused on Diverse Populations and Equity Focus, he suggested eliminating questions around demographic composition that could be collected using the statewide dashboard. As such, I brought these numbers to the table and asked the questions around which demographic markers have been a focus of improvement work and

markers beyond state reported numbers. In this way, I did not spend valuable interview time on data that I already have available. In addition, in question six, I shifted the term “customer” to “people you serve” and eliminated the leading questions around language, technology, and transportation barriers.

Limitations and Potential Implications

As with any research study, I recognize that this study has limitations. First, my positionality must be taken into account. I am a white woman raised in a middle to upper class home, and I am attempting to study educational equity issues, many of which I have not personally experienced. As a result, I have an outsider understanding of the experiences documented in research around racial and economic equity concerns. I depend on my learning and the voices of people from traditionally marginalized groups to understand these ideas.

I am a district level administrator in the midst of implementing an Improvement Science approach. As such, I believe in the power of the tools of Improvement Science, and this influenced how I heard and interpreted ideas from the participants in the study. I have employed several methods to counteract this potential bias and ensure validity of my conclusions. Nevertheless, my perspective served as a limitation in this study.

Examining only two schools limits the breadth of the research study. A sample size of two may uncover some phenomenon and ideas, but more would reveal greater depth of understanding. Limited time and resources prevent me from expanding the study to more participants and/ or a longitudinal examination.

Given the limited representation of teachers and administrators of color in schools as a whole, the majority of my participants had similar backgrounds to me. This limited the perspective on equity to one aligned with racial and social groups perceived to have power in the

system. In addition, I elected not to include students in the study, and this exclusion limits perspective to those of adults. Again, time and resources prevent me from expanding the scope of the study, and I recognize this limitation.

Contributions

Current studies either examine Improvement Science without regard for issues of equity or question the use of Improvement Science highlighting potential issues with equity. This study seeks to examine how equity is addressed in Improvement Science. The concept of addressing inequity through Improvement Science will support efforts to understand local context when implementing change that positively effects students from traditionally marginalized groups and narrows opportunity and achievement gaps. I am influenced by and seek to support Byrk's (2017) hypothesis that Improvement Science is the key to making "real progress on the educational inequities we care about" (p. 4). To date, other studies have skirted the edges of this idea, but there is limited understanding of the convergence of Improvement Science and narrowing achievement and opportunity gaps. This study recognizes the intersection of equity and Improvement Science in which schools have sought to impact equity outcomes and understanding the nuances of implementation that achieved these goals.

Chapter 4

In this section, I will describe the cases of Improvement Science work and the ways each district considered problems related to equity. The superintendent in each of the identified cases has presented at regional and national conferences on the topic of Improvement Science, sharing details of the improvement process in action within their schools consistent with the purposeful sampling selection of the cases. I selected to conduct a case study at the school level, and asked the superintendent to identify which school was using Improvement Science to fidelity and had engaged in goals that considered equity. Both districts identified the lone high school in the district as the site for deeper analysis. I will begin by describing each district and school in the context of the local community and the state. This will include demographic information and publicly reported student achievement and engagement (opportunity) data. Then, I will outline the district's and school's implementation and use of Improvement Science. Next, I will attend to the participants experiences with equity work. Finally, I will highlight how equity did and did not play a role in the problem identification, measurement, or disciplined inquiry in seeking to make an improvement.

Northwestern District (Case #1)

School District Context

Located approximately an hour and a half drive from a major metropolitan area, the school district serves approximately 5,000 students. The district includes seven elementary schools, one middle school (grades 6 – 8), and one high school (grades 9 – 12). The district and school leadership articulated that the community is not very diverse. This is true if the only facet of diversity is identified as racial/ethnic; approximately 10% of the student population in the district identifies as a person of color. However, a larger percentage of the school population

lives in poverty with 40% of students qualifying for free and reduced lunch. I selected the Northwestern Schools based on the economic diversity within the student population.

Northwestern School District is recognized for their work with Improvement Science and the positive impact the work has had on publicly reported academic measures. As I sought to document the facets of Improvement Science, there were gaps in implementation and documentation of all principles of Improvement Science. I selected the district based on their national and regional reputation and the sizeable population of students qualifying for free and reduced lunch. In the following narrative, I will begin by documenting academic outcomes and opportunities. I will provide evidence of the components of Improvement Science observed and the lack of intentionality of focus on the historically marginalized population of students qualifying for free and reduced lunch. Finally, I will turn attention to the emerging considerations of equity.

Student Diversity. In the Northwestern School District, the rate of poverty has been consistent over the last five years with a slight increase in poverty in the 2019-2020 school year based on enrollment data on the state department of education website. Consistent with national trends, Northwestern School District achievement and opportunity gaps are evident on the high school achievement measures. The achievement gap between students living in low income homes and students with access to income above poverty level widens as students progress through K12 education.

Students living in poverty have higher rates of food insecurity and lack of proper nutrition can impact brain development. A faculty member shared that some students at the high school depend on the school food market, an extension of the local food pantry, for their groceries and nutrition. According to the local food pantry newsletter, the food bank saw a 35%

increase in requests for food during the fall of 2020. One student waited to pick up his family's food box until the end of each Friday to access additional food if available to make sure he had enough food after making sure his younger siblings were properly fed. According to the local food pantry newsletter, the non-profit commits to serving as a consistent source of food because "the lack of consistent, healthy nutrition cannot be the reason that they do not succeed."

Families living in poverty have less access to preventative medical care and less access to resources such as pre-school and technology. At Northwestern School District, students and families are supported in accessing care through an onsite health clinic.

High School Achievement Measures. At the high school, approximately 1,500 students attend school each day. Overall, achievement on statewide assessments have been consistent with the state average, and graduation rates have climbed from below state average in 2016 to just over the state average in 2019, according to the state department of education website. Just under sixty percent of graduates enroll in college after graduation with one third of those students enrolling in a 2-year institution and two thirds in a 4-year institution. Post-secondary enrollments mirror the state averages over the last five years based on data accessed on the state department of education website. The principal reports that only forty percent of students earn a degree from a post-secondary institution, which concerns her.

Taking and passing an Advanced Placement (AP) course in high school contributes to a student's readiness for college (Domenach, 2017, p. 4). At Northwestern High School, students have access to fifteen AP courses according to the Course Selection Guide for 2020-2021 on the school's website. In addition, students may elect to take one or more of fifteen dual enrollment courses transcribed through the local technical college. The state reporting dashboard provides information on the percentage of students taking AP courses with success rates disaggregated by

economic status. Since 2014, participation in AP exams has increased from approximately 14% to 24% with success rates increasing from 62% to 75%. During the same period, statewide participation in AP exams has gone from 14% to 16% and success rates have remained consistent at approximately 66%. Northwestern High School is increasing participation and success on AP Exams at a faster rate than the state.

According to Burris and Wellner (2005), “Achievement follows from opportunities” and assert “the persistent practice of tracking denies a range of opportunities to large numbers of students” (p. 594). AP, or “advanced” courses in general, can serve to track students into stratified learning opportunities. Narrowing the gap in academic outcomes can be accomplished by expanding access to high-track curriculum, such as Advance Placement, to all students (Burris & Wellner, 2005, p. 594). In Northwestern High School, not all students take Advanced Placement courses; therefore, I measured the impact of limited opportunities by disaggregating the participation rates. When disaggregated by economic status, the state reporting dashboard shows a substantial discrepancy in participation and success rates on AP exams, with students from low income homes increasing participation and success at a slower rate than their peers. At Northwestern High School, ten percent of students receiving free or reduced lunch took an AP exam in the spring of 2019 (which is the most recent data reported publicly on the state department of education website) compared to thirty percent of students who do not qualify for free or reduced lunch. Similar to the aggregate, Northwestern students are improving participation and success on AP exams at a rate faster than the state, but the discrepancy between economic classes is growing.

High School Engagement Opportunities. *Redefining Ready* identifies participation in after school activities as an important predictor of career readiness (Domenach, 2017, p. 4). At

Northwestern High School, students at the high school level have the opportunity to engage in one or more of twenty-five competitive activities and sports. The school offers a number of music programs including a marching band, a show choir, and an orchestra performance group. Arriving at the school, I watched the marching band rehearse on the football field. Students participate in music programming at a rate of twenty-five percent; athletic programs engage approximately forty percent of students; and academic after school activities include just under fifteen percent of the students as of the 2016-2017 school year based on the state department of education website. Participation in co-curricular activities was not disaggregated by race/ethnicity, economic status, or other distinguishing demographic characteristics at the school level or on the state department of education website.

Implementation of Improvement Science

In 2016, a new superintendent took the lead in the Northwestern District. The superintendent shared that the Board charged her with the task of improving student achievement based on a historical record of below average to average performance on state reported test measures. Using the tools learned with a national consulting firm on the use of Improvement Science, the superintendent embarked on the development of a strategic plan with input from community stakeholders. A coach from the national consulting firm continues to work with the leadership team on data collection, data analysis, goal setting, and “hardwiring” “always” practices for the system. For Northwestern, this meant annual student and family perception surveys, implementing Plan, Do, Study, Act cycles at the building and classroom level, and administrators completing rounding conversations with staff on a regular cycle. I will outline the practice of rounding in a description of how the administration sees the system that produces the current outcomes. Since 2017, school and district leaders have presented their continuous

improvement journey and shared their stories at state and regional conferences on Data Driven Decision Making and Continuous Improvement.

Drawing on the interviews with district administrators, building administrators, and high school teachers, I will highlight evidence of the use of Improvement Science within the district and the school. This will include four of the six principles identified in the research, of which I observed evidence throughout the case study. I will document the following principles:

- Measure Improvement, focusing on the data considered in identifying a problem and the data to determine if a change results in improvement
- Make the work problem-specific and user-centered, understanding which problem-focus led to an improvement
- See the system that produces the current outcome, which will highlight processes used to understand the current practice when generating solutions to the problem
- Anchor practice in disciplined inquiry, reviewing the cycles of improvement documented on the identified problem.

Evidence of the principles of focus on variation in performance and accelerate learning through networked communities did not surface throughout the data collection process. Then, I will focus on considerations of equity or lack thereof in the Northwestern School District within a recent problem or set of problems identified for improvement.

Measurement

Bryk et al. (2015) assert, “We cannot improve what we cannot measure” (p. 87). In a focus group which included the Director of Instruction, the Director of Pupil Services, the Director of Human Resources, the Director of Technology Services, and the Director of Business

Services, the team identified data that is collected at regular intervals. In interviews with the high school principal, the high school assistant principals, and the school counselors, each identified a similar set of data markers considered annually in the process of problem identification and seeing the system that produces the current outcomes. Unlike studies I considered in my review of existing literature, the data collected reaches beyond test score data. When asked about sources of data, nearly every participant listed at least four non-academic indicators before remembering to identify a student achievement test score. Considering the district documented student test scores as the focus of the district-wide problem, which was referred to as the “mega-result” in conversation, I was surprised test scores were not the first measurement identified.

Each participant in interviews identified at least five sources of data that he/she reviewed at set intervals throughout the school year. The data is reviewed, but not every review results in problem identification. Most data are monitored to ensure continued consistency for outcomes. During the routine data review, a problem may surface in the data. During interviews with leaders at the high school, I learned that school leaders targeted a problem identified in ACT scores for improvement over the last three years. Because students only take the ACT once per year, this measurement was insufficient for disciplined inquiry for improvement. Following the principles of see the system that created the current outcomes and disciplined inquiry, school leaders identified potential root causes. For each hypothesis for a root cause, the team identified specific measures. The team considered extended absences resulting in missed instruction as a root cause and monitored attendance data as a measurement for improvement.

Perception Data. Among the data collected at least annually, surveys were a key tool. Students and parents participate in satisfaction surveys that have been calibrated and validated. Students take an annual survey developed by the state agency focused on youth risk behaviors.

Staff provide input through an employee satisfaction survey. Students and families are asked to take an engagement survey developed by the company coaching leadership through Improvement Science each year. This perception survey includes question sets focused on engagement, belonging, teacher effectiveness, administrator effectiveness, and quality of facilities.

Survey data may be used to identify a problem or examine the system that is producing the current outcomes or may be used to measure improvement. When low ACT test scores were identified as a problem at the high school level, the perception data on the survey showed low scores for the questions around engagement in learning. The school leaders considered this correlation of data to consider root causes of the main problem, which was low ACT test performance.

Academic Data. College and Career Readiness Indicators provided another source of data. Administrators at the high school indicated that they examine attendance records and patterns at least monthly. Teachers and school administrators discuss student outcomes measured by grades, course failures, and performance on local benchmark assessments routinely. School leadership tracks student participation in extracurricular activities and service hours in the community using a student portfolio software. All of these measures were documented as an aggregate for consideration. In addition, the principal reported that administration considers actual college attendance and completion data.

During the interviews, district office administrators mentioned that state assessment scores were improving, and they are. However, little attention or discussion focused on this source of data or focused on test scores. “Student Achievement” is identified as one of the five major areas of focus on the District Scorecard for the 2018-2019, 2019-2020, and the 2020-2021

school years. All three measurements identify a statewide standardized assessment. At the high school level, the goal on the 2019-2020 District Scorecard states, “We will increase student combined proficiency resulting in the growth of the ACT Composite Score average.” In speaking with the high school principal, this data was easily accessible and cited during the interview. She highlighted the improvements in performance since she began her tenure as high school principal in 2018. She explained how students had opportunities to become familiar with the format of the test through practice opportunities she put in place when she started through interim assessments and an optional ACT prep academy. Based on subsequent responses and other participant interviews, this initial explanation of improvement did not appear to be the true focus of improvement efforts that actually targeted instructional focus on standards and student attendance to access instruction.

Disaggregating Data. In order to identify a problem in the data, one must have access to data that tells the story of not only overall academic outcomes, behaviors, attendance, and perceptions but also the story of potential disparities in outcome measures. In this district, most data are reviewed as an aggregate. Only two pieces of data that the administrators identified as routinely reviewed data are disaggregated by race, economic status, gender, home language, or other demographic identification—the Youth Risk Behavior Survey and standardized test data. When asked about discrepancies in the data for these measures, district and building administrators acknowledged that they have not thoroughly examined the disaggregated portion of the data.

One administrator at the district office level shared that the administrative team was concerned about the aggregate data for achievement and identified an overall instructional problem. The problem was analyzed using a couple of tools, which I will describe in the problem

specific focus. The leaders felt it was unnecessary to analyze the data by demographic markers because the system wasn't working to achieve the desired outcomes for most students. Then, upon reflection, the same district office administrator suggested that examining the data on a disaggregated basis may help with developing strategies to improve outcomes for the aggregate. Another district office administrator recognized the need to consider disaggregated data, saying "the school teams when they're looking at all of the students and that data, what are all the contributing factors to this student being successful or not successful and then really honing in on what those things are and providing the supports and services they need to continue to close the gap or to keep them on track or to move them forward." In this context, she was addressing the question of reviewing assessment data in disaggregated demographic groups, and then transitioned to using data to support an individual student. I heard a pervasive tone of acknowledgement that the team should consider disaggregated data for all of the metrics. Some of the participants believed school teams were looking at measures in a disaggregated way, and one specifically stated it was "happening in pockets." In all the interviews, participants pointed to someone else or teams they believed were considering disaggregated test scores, but no participants identified a time when they used disaggregated data to identify a problem or seek to understand the root of a problem. The lack of evidence that individuals or teams are analyzing metrics in a disaggregated format indicates it is not happening.

Problem Specific Focus

The district office administrators identified the problem focus as "the mega result," which they identified as student success. The Northwestern School District uses the term mega result as the overall outcome that is impacted by many factors. This is the main problem focus for improvement efforts. The district and the school outline the problem as below state average

standardized test results. The school-level focus of the case study is the high school, and therefore, I will focus on the measurement used at Northwestern High School: ACT test performance. During the focus group and the interviews, each member of the team had a different perspective of what in the system was contributing to ACT scores below state average. I will share the process and ideas considered in examining the system that produced the outcome of below state average test results.

Processes to Understand the System that Produces the Outcomes. The Northwestern School District has several tools on which they depend to analyze the system that is producing the current results. I will share two tools discussed in the interviews: rounding and fishbone analysis (which was presented in chapter 2). Then I will connect these tools to the specific problem focus of below average state test scores.

During the district office administrator interviews, several administrators identified a tool called rounding. An administrator described rounding as a process of engaging in individual conversations between administrators and students or staff to collect qualitative insights into lived experiences. After analyzing a data set of district and school level results on the student perception survey, administrators “roll out” results. A district office administrator highlighted a roll out of data from a student safety survey several years earlier. Following the roll out there was concern about response to a question about kids being nice to each other at school. Focusing on this data, the administrator had observed behaviors in the building and talked to kids about how the survey results didn’t match her experiences. By rounding with students in the lunchroom, she learned that many students reflected on negative social media interactions outside of school with classmates. These experiences affected their perceptions of how “nice” kids were to each other. From this, the school leaders set a goal to improve student responses to

the question “kids are nice to each other in this school.” The district office administrator, reflecting on her previous experience as a principal in the same district, shared how she used these ideas and formed focus groups of students to advise on strategies for improvement. She said, “I worked with Student Council, and we collectively worked on how we could bring a positive culture.” None of the participants at the district office level related this use of rounding to the identified problem focus of below state average standardized test results.

The high school principal shared an example of rounding about student absences which was a perceived root cause of low standardized test performance on the ACT. This process provided insights from the users for ways to improve outcomes, offering perspective on the system that produces the current result. One of the themes that emerged focused on health concerns, including fatigue and anxiety. Student rounding responses were consistent with data on the Youth Risk Behavior Survey and led to the introduction of on-site care for mental health and partnership with the local food pantry. She perceived that the change resulted in an improvement in attendance. Chronic absenteeism dropped from twenty percent to thirteen percent over the last three years. While there was reduced rates of chronic absenteeism, it is difficult to link the improvement to specific changes.

At the high school level, the staff described the use of a fishbone process to identify factors contributing to the current ACT results, focusing on the question, why are we seeing these results and what factors are contributing. Administration focused on attendance as a primary contributing factor. In order to impact ACT scores, the team focused on the root cause which they theorized was a lack of engagement in academic learning. Short cycle goals last year and this year focused on increasing attendance and addressing mental health concerns of anxiety

and depression that were interfering in engagement and clear indications of a lack of engagement in curriculum and instruction within the classroom.

Disciplined Inquiry to Improve

Multiple participants described a process of disciplined inquiry. As I reflected on their descriptions, I noted that the explanations lacked detailed examples. At times, the participants blurred multiple cycles of inquiry into one explanation. This made it difficult to draw direct connections between cycles of improvement and the direct improvement perceived by the participants collectively. Participants provided limited documentation of the referenced processes. In documenting disciplined inquiry, I drew the conclusion that staff loosely followed the concept of disciplined inquiry but there was not accountability in documenting or reporting efforts.

Scorecard. Every three to five years, the community engages with the Board of Education and administration to align vision and values. This process begins with a three-day community convening in which well over 100 stakeholders participate. District Office administrators describe a strategy to invite representative stakeholders to the table to make sure multiple perspectives are included. The multiple perspectives included business leaders, religious leaders, parent leaders, senior citizens, students, staff, healthcare leaders, and leaders in other community organizations. I asked about representation of different racial, ethnic, disability status, gender and gender identity, economic status, or home language. After a pause, one district office administrator indicated none of these was a facet of representation considered in the process, stating, “We don’t target them separately.” She suggested that the community is not very diverse. This theme of perceived homogeneity persisted throughout the interviews. While I selected the district as diverse based on the percentage of students qualifying for free or reduced

lunch, the participants did not recognize, see, or hear the factors impacting this facet of diversity. It appears these voices are underrepresented in the conversation.

Once the district focus is established, each school, district division, and teacher team analyzes and considers changes that are influencing the “mega result,” which Northwestern School District defines as student outcome measures on standardized tests. The scorecard sets specific measurement targets, and each school employs the Plan, Do, Study, Act (PDSA) cycles to make a change and measure the short-term impact.

Plan Do Study Act (PDSA) Cycles for Improvement. Plan, Do, Study, Act (PDSA) describes the process in which teams develop and test solutions to a specific problem focus identified in the data. In the Northwestern School District, administrators and teachers are responsible for a PDSA goal every year, with analysis of impact of a change in the system every 90 days.

The main focus of improvement identified in the community strategic planning process and documented on the District Scorecard focuses on ensuring “that every student achieves or exceeds academic growth targets” on statewide standardized assessments. At the high school, the administration focused a series of PDSA cycles on attendance as low attendance was identified as a contributing factor to low ACT performance. In an interview, the principal shared a number of changes made over the last several PDSA cycles to garner improvement in attendance with the belief that improved attendance would support improved achievement on the ACT. Over the last two years, absences have been reduced. High school administrators believe PDSA changes including a campaign to tell students and families the importance of attendance, eliminating challenges to attending including access to free food and school supplies in a school store and providing onsite health and mental health services contributed to this improvement. The

administrators were not able to provide documentation of the PDSA cycles leaving this conclusion open for interpretation.

While the administrators each referred to a PDSA process at the classroom level for curriculum focus and at the school level for ACT focus, none provided documentation of the process. When specifically asked for the PDSA report or document used to track the disciplined inquiry at the building level, no such documentation was retained. This lack of documentation indicates that the school leaders relied on memory for the disciplined inquiry process. In the interviews, they loosely described the steps and referenced conversations and meetings. Without documentation, I am unable to verify the fidelity of the process. I am limited to the memory of the participants, and I question the representation of Improvement Science, specifically disciplined inquiry, within the Northwestern School District.

Summary of Improvement Science Implementation

After five years of development in Improvement Science implementation, the Northwestern School District and Northwestern High School have engaged in multiple facets of Improvement Science principles. However, there remains challenges of coherence and consistency. Leadership has engaged with stakeholders to make many changes, not all of which are improvements and not all of which have emerged from Improvement Science work. Throughout the data collection process, participant responses wander to share stories of many different changes. I interpret these diversions and lack of clear focus in the interviews to be indicative of distractions from problem-focus and inconsistent attention to user focus.

I observed evidence of a commitment to collecting data and measuring outcomes valued in their schools and communities. The analysis of the data has some clarity of focus around student test results as a problem specific focus. In the focus on outcomes for standardized test,

the district and building leaders have employed additional principles of Improvement Science by working together to see the system that produces the current outcomes and employing disciplined inquiry in isolated improvement efforts that contribute to the performance on standardized test measures.

Some of the tools such as rounding and fishbone diagram have been used but inconsistently. Simultaneously, teachers and school leaders are distracted by other problems, flirting with tools aligned to the principles of Improvement Science. In these cases, they insert additional solutions that may dance the edges of the main problem focus and create confusion among the stakeholders what the true problem-focus is. The school goal focused on test scores below state average. The targeted change during the 2019-2020 school year was decreasing absenteeism. During interviews with school leaders and professional staff, individual teams focused on other changes outside of the specific tools and focus. One leader talked about increasing clubs and activities to help students find an adult with whom they can connect. One professional discussed the structure of delivering mental health services onsite. Participants lacked a clear connection during the interviews.

The district has presented their implementation of Improvement Science at regional conferences, focusing on the components in place. There is evidence of effective improvement. At the same time, the consistency of the work appeared limited.

Equity Work

I will provide context for the local environment related to understanding and connecting the impact of factors of race, poverty, and language on problems identified within the school. First, I will articulate participant perceptions of identity and equity. Then I will connect their perceptions with their interpretations of data and analysis of the system that produces the current

outcomes as well as improvement efforts. Finally, I will link an equity lens to the improvement efforts the participants perceived impacted outcomes and lack of significant improvement.

District office, school administrators, and teachers with whom I spoke all identify as white. When asked about equity work and their professional development in this space, all felt their learning was in its infancy, and most reflected exclusively on their development around racial identity. Almost every administrator echoed a building administrator who said, “Our school isn’t very diverse.” When prompted to reflect on other facets of diversity markers such as economic status, a new narrative emerged. Participants acknowledged that a significant percentage of students qualify for free or reduced lunch. Some participants viewed this facet of diversity as a deficit. Some participants considered the impact of these lived experiences and valued the perspectives of students and families.

Capper and Jamison (1993) caution that lack of attention to understanding the perspectives of traditionally marginalized groups could result in biased identification of root causes and potential solutions that further marginalize students. I recognized a white, middle-class normative culture during interviews with participants. Some participants revealed underlying and unconscious bias in their discussion of students and families living in poverty.

Perceptions of Facets of Diversity

During the summer of 2020, national protests aligned with Black Lives Matter following the murder of George Floyd in Minneapolis impacted conversations in Northwestern School District. A staff member retweeted President Trump, which read “when they start looting, we start shooting.” Social media reactions prompted a recent graduate of the high school who identified as a person of color to meet with District Office administrators and shared their experiences with racism in the schools. The student had been successful based on all metrics the

school typically uses to measure college and career readiness, and the student planned to attend an elite, four-year state school in the fall. The revelations and subsequent discussions were disruptive and uncomfortable for the predominantly white staff. I attempted to identify the former student to interview; however, no school personnel would disclose her identity. One administrator agreed to share my contact information with the former student, but the former student did not respond to requests to contact me.

Given this context, it was not surprising that staff in the schools had racial diversity as the only diversity marker on their minds. Administrators and teachers did not immediately see or reflect on the economic diversity of the student population in the school district and the high school. During interviews, I prompted further reflection with questions specifically about students from low income homes. Many administrators and teachers would then dive into several programs in place to support all students, but especially students who are struggling financially and with mental health based on trauma. In both administrator and teacher interviews, representatives of the school discussed programs for providing food, free mental health visits, and free school supplies. Two of the school administrators took a tone of benevolence and pity for circumstances. One of the administrators shared a very different perspective. I will share these insights in the context of elevating the voices of those most impacted by the problems typically associate with students living in low income homes.

Representation of Voices

Two of the high school administrators began discussing strategies that help everyone “must be helping the kids who struggle, but we don’t have data on that.” As the discussion continued with one of the administrators, he referenced “those kids get left behind.” He talked about clubs to help connect with “like minded individuals” were important to give kids a push to

get involved. These administrators chose language that indicated a perception of students from lower economic classes as other or having deficits.

While I did not interview students, voices from research including Diamond and Lewis (2016) indicate that some students experience an undercurrent of lower expectations for students from demographic groups perceived to have deficits. One parent interviewed by Lewis and Diamond (2016) explained “I think expectations are less. I think you have to prove yourself” (p. 95). I perceive the language used by the administrators in the interview are consistent with this student’s interpretation from the high school in the case study by Lewis and Diamond (2016). “These kinds of assumptions” by adults and peers “often operate at subconscious levels and involve presumptions about more than socioeconomic status, including presumptions about such things as likely family investments in education and student academic trajectory” (Lewis & Diamond, 2016, p. 91). Hearing the statements from these two administrators, I wondered if students in Northwestern High School felt these same low expectations.

In each of the interviews in which administrators or teachers mentioned listening to student voices, I asked how they ensure all of those voices in the school community are empowered and heard. Two of the administrators talked about how relationships encouraged students to speak up. One high school administrator said, “There’s somebody in this building that a kid can connect with and feel comfortable connecting with intentionally.” He went on to add that, “We’ve asked our freshman...to identify one or two people that they relate to.” His perception was that all students had an adult with whom they could connect and as a result, their voice would be heard.

In the interview with the high school principal, the same question prompted a different reflection. She discussed ways in which they could better elevate voices of some groups of

students. She referred to a “sub-culture” of those “whose voice is muffled by the culture we see.” During this interview, the administrator shared personal experiences living in poverty. As she shared, she reflected on how it has impacted her choices as an adult and her values in relationships. Living in a much smaller home when her income would allow for more was a source of pride. She discussed how those values of family as a priority were adopted by her children who attend the high school. The tone of her responses to questions about students reflected this perception and value for positive aspects of diverse perspectives.

These differing perspectives of the value and worth of students living in poverty may have led to subtle differences in expectations for students living in poverty. This administrator in the second year at the school likely has not lifted stereotype threat for students living in poverty; however, the value and expectations implied in the interview of this participant is likely seen and heard by students within the school. In addressing the main problem of below state average test scores, the leadership of this administrator in identifying potential root causes reflected high expectations and focused on changes to systems rather than students. For example, with absenteeism considered a root cause, this administrator expected that students would attend if they could and addressed external barriers to attendance rather than assuming the student did not want to come to school.

Specific Improvement Focus Examples

In all the interviews, a mental health program and grant sought by the district surfaced as an improvement effort to increase attendance in order to increase access to instruction which would result in improved academic performance. In addition to monitoring test scores, the district and school teams routinely reviewed leading indicators for the targeted problem contributing to test scores with the surveys on youth risk behaviors, perception surveys, engaging

in the community conversation, and talking with focus groups, it became clear that mental health and student trauma needed attention. This section highlights how the loose attempt at disciplined inquiry into targeted interventions focused on addressing factors contributing to absenteeism may have impacted test performance and may have contributed to addressing inequitable access to instruction.

Mental Health

Three years ago, the district identified problems with anxiety, depression, and trauma response leading to chronic absenteeism. All administrators and teachers discussed elevated scores on a survey for suicide and noted the frequency of death by suicide in the district with approximately one to two students each year. In focus groups and discussions with community stakeholders, leaders learned that students and families didn't have sufficient access to mental health services such as therapy. Coupled with high rates of anxiety and overall student performance below state averages, leaders identified mental health as a focus to target improvement. Around the same time, the state department of education offered a competitive grant for mental health supports.

With funding from the competitive grant, the district launched mental health services hosted at the high school. A school social worker was added to oversee mental health services located within the guidance offices at the high school. Every family in the district is offered access to "Northwestern Therapy" which includes a monthly visit with a licensed therapist for whomever in the family is in need of services.

The high school principal discussed the perceived stigma associated with mental health was a roadblock for many at first, particularly families living in poverty. As such, the superintendent required every administrator to participate in two mental health therapy sessions

per year. The goal was to help shed the stigma and encourage participation. By situating the services within the guidance office, students and families were not singled out or identified as attending therapy when they arrived thus protecting their privacy and confidentiality. In addition, the high school principal recognized how poverty contributed to mental health needs. She described “they’re living in cycles of abuse and addiction, and they never had a helping hand to get out.” She understood this to be part of the system keeping academic performance from improving, based on rounding conversations with students and families. By developing a resource for mental health services at school, she believed “breaking that cycle and sharing that therapy and asking the right questions and getting them the right resources, kids are able to see other options for themselves.” Leaders believed the option of pursuing post-secondary education became attainable and motivated students to attend, see the value of the ACT, and engage in instruction.

To ensure that no students were slipping through the cracks and all students were referred for services, all students were trained in QPR which stands for Question, Persuade, Refer. This research-based strategy is designed and used as mental health first aid. Teachers were also trained to recognize the signs of suicidal ideation and practice QPR.

There is evidence that the program is working. Suicides are down with no death by suicide in the last two years. Students and families are using therapy sessions. All those interviewed perceived a positive impact from access to mental health services. Subsequent to the implementation of mental health services at the high school, absentee rates went down and academic test performance increased from below state average to above state average. Participants in the study did not draw this direct correlation between mental health services and student achievement in school. I found evidence that the change contributed to the improvement

in the identified problem focus; however, this specific change does not appear to have been made through the use of the tools of Improvement Science.

Relationship Focused Programs

When the high school principal took the lead two years ago, she was frustrated that the programming efforts for course offerings were not responsive either to the needs of the students or to the data. She described efforts by her predecessor to increase college level course work that was focused exclusively on four-year college pathways. The data showed that fifty-seven percent of students signed up for college and “only” forty percent completed a four-year college degree. To her, there was a disconnect between program development and the data. She believed that it is important to prepare students for college, but the message the school was sending was that you are not valued if you don’t drive towards four-year college. She believed this was contributing to the below average test scores and high rates of depression and anxiety. Given her perspectives on poverty, her interpretation of the data and how she heard the voices of the students was different than that of her predecessor or others on the school staff. She was seeing the data and hearing the voices through a different lens.

Considering the data and talking with students and families, she initiated multiple programs that focused on valuing multiple paths to college and/ or career, as well as relationships for students to find their connection to adults in the school and access mental health services. She identified and assigned a champion to lead each program. These programs included LINK Crew, a school-based market for free food and school supplies for all students that benefited students in need most, and student mentor programs. All of these programs were considered in the context of valuing students and proactively managing student self-image to prevent anxiety and reduce rates of depression. She worked to increase access to credit bearing

dual enrollment college courses with the technical colleges to build viable pathways to the trades and technical careers.

While college going rates have remained the same, student achievement scores on the ACT and participation in college level courses have risen dramatically on the aggregate. The school leaders perceive that these improvements have been in response to the changes implemented. School administrators indicate that the focus on providing mental health supports led to improved school attendance which in turn provided higher engagement in academic instruction, including college level course work.

Case Summary

Northwestern School District and Northwestern High School applied Improvement Science principles to focus on a problem of below state average ACT scores. Using tools such as rounding and fishbone diagramming, root causes became the focus of change efforts. PDSA cycles reportedly measured if the change resulted in improvement, which it did. Improved attendance and student self-perceptions and mental wellness were also believed to contribute to improved outcomes on the ACT and numbers of students taking college level courses. The high school principal's lived experiences influenced how she understood students most impacted by the problem leading to current outcomes. While there remain disparate outcomes, students living in poverty are seeing improvements in mental health and academic outcomes.

The collective commitment to Improvement Science offered opportunities to refine practice. Participants discussed four of the principles in interviews; however, I was only able to access written documentation for two principles. I interpreted the lack of shared documentation and retention of documents for fishbone diagrams, rounding conversations, and disciplined

inquiry cycles as an indication that the attention to these principles were unclear to the stakeholders.

Administrators were collecting and reviewing measures of student achievement and identifying metrics for improvement. Administrators are focused on a problem documented on the District Scorecard. Administrators perceive that the District Scorecard also documents disciplined inquiry, but the full implementation of PDSA cycles are not documented. Through the processes of rounding and fishbone analysis, administrators are attempting to identify the facets of the system that are significant in the outcomes of the organization. The lack of documentation provides evidence that processes are fragmented, and all staff are not engaged in the process of improvement. While the Northwestern School District has been recognized for their commitment to Improvement Science and presented the story at regional conferences, evidence indicates the implementation lacks consistency and fidelity to the processes. Leaders could increase collective focus on specific problems through a set of tools to document, store, and share publicly themes and demographic composition of rounding conversations, fishbone analyses, and an A3 form for documenting PDSA cycles (Appendix F).

My research question asks: In schools implementing Improvement Science, how and to what extent do school leaders attend to diverse populations and equity when implementing the improvement process? At Northwestern High School, individuals are trained in and employing tools and principles of Improvement Science, but the documentation of these principles is inconsistent. When engaging in a practice associated with Improvement Science, administrators acknowledge that they have not considered equity directly in the process. Changes have been made, and there has been significant improvement in standardized test performance, reduced absenteeism, and expanded access to high-track curriculum. At the same time, students

qualifying for free and reduced lunch did not experience improvement at the same rate. This demographic group experienced improvement in standardized test performance and increased numbers accessed high-track courses at a slower rate, resulting in the gap widening. Within the study, I found it difficult to connect the improvements to the use of the principles of Improvement Science. I did not see evidence that equity played any role in efforts to improve, and improvements for the students qualifying for free and reduced lunch appear unplanned.

Chapter 5

Southeastern District (Case #2)

School District Context

Situated in a suburban community of a medium-sized city, the school district contains four elementary schools, a middle school, and a high school. The district served a student population of approximately 4,000 students at the time of my study.

The Southeastern School District is regionally and nationally recognized for excellence in Improvement Science. In documenting the implementation of the principles of Improvement Science, evidence I collected confirmed the reputation of the district's use of Improvement Science. The suburban district is racially and ethnically diverse, and this facet of diversity went relatively unnoticed or ignored until 2018. The community and the school have recently confronted disparate experiences for students of color in the schools, and the school system has just begun to focus Improvement Science principles on problems of inequity. The superintendent said Improvement Science is the tool to focus on "our most wicked problems" in schools. Inequalities certainly meets the definition of "wicked" as a pervasive and challenging problem. The district is on the precipice of applying the principles of Improvement Science to problems of equity, and the following narrative will highlight how the district is poised to address this "wicked" problem.

I will begin by describing the student population with details on the data used to measure system outcomes for students. Then I will articulate the principles of Improvement Science in action within the Southeastern Schools with a focus on the high school experiences. Finally, I will surface Southeastern High School's recognition of disparate experiences and outcomes for students based on race.

Student Diversity. The Southeastern District includes just over 25% students of color with 8% identifying as Black, 5% identifying as Hispanic, 8% identifying as Asian, and 6% identifying as multiracial. The district has had a similar distribution of racial and ethnic identity over the last five years as reported on the state department of education website.

Nationally, there is a documented discrepancy in achievement and opportunity for students of color, particularly Black and Hispanic students. In addition, Black and Hispanic students are more likely to be suspended from school. According to the state department of education website, in the Southeastern schools, Black students scored proficient on statewide assessments at half the rate of their white peers. Black students were suspended four times more than their white peers, and Hispanic students were suspended five times more frequently during the most recent publicly reported data in 2018-2019.

Student Achievement Measures. The percentage of students identified as proficient or advanced on the statewide standardized assessment has consistently exceeded the statewide performance by a sizable margin for the last five years as reported on the state department of education website. However, assessment scores appear stagnant with a slight decrease in the percentage of students achieving proficiency in each of the last two years for which data is reported, which was the spring of 2019. Between 2012 and 2016, the Southeastern Schools saw growth in standardized test performance moving from state average proficiency scores to exceeding state averages as reported by district leadership in presentations on continuous improvement.

Students of color are consistently scoring below state average. According to the state department of education website, only 30% - 40% of students who identify Black consistently scored proficient on the statewide standardized assessment measure compared to nearly 60% of

their white peers. Performance has been declining for students who identify Hispanic with proficiency levels of 38% in 2018.

At the high school, students have numerous opportunities to engage in high-track coursework. The Course Guide on the high school website includes 25 AP class offerings, 5 courses transcribed with four-year institutions across the state, and 9 courses transcribed with a local technical college. Based on the five-year trend data on the state department of education website, over 400 students take an AP test associated with the AP class offerings with a pass rate of 75 percent. The number of students taking AP courses and passing AP exams has remained consistent over the last five years. Notably, white and Asian students are three times more likely to take an AP course than their Black and Hispanic peers.

Student Engagement Measures. Southeastern High School engages almost 1300 students in academic learning and a variety of after school activities. The school website lists over 35 academic and social clubs and activities. Among the clubs listed, there are several focused on facets of identity, including Awareness of Racial Inequality (ARI), Gender Sexuality Alliance (GSA) and Voice. During interviews with school leaders, these clubs were highlighted as active groups advocating for improvement in equitable access and outcomes for students. Attendance at school is over 95 percent for all students and consistent across races over the last five years based on data on the state department of education website.

Implementation of Improvement Science

The Southeastern District began its Improvement Science journey with a new superintendent a decade ago. At the time, student achievement measures were below state averages, and the local board of education charged the new superintendent with improving academic achievement and opportunities for students. Based on the new superintendent's past

experiences and learning from two improvement models from business (Total Quality Management and Lean Practices), she focused on implementing data metrics for anything that could be measured and engaged with two national groups applying Continuous Improvement tools and Improvement Science research in schools. Within the first five years, student achievement overall as measured by standardized test scores had climbed dramatically and surpassed state averages by a wide margin. These standardized tests included the state assessment for grades three through eight, the ACT, and AP exams.

In this section, I will use data from interviews with district level administrators, building administrators at the high school, and teachers to outline the application of five of the six principles of Improvement Science. These principles include measure improvement, make the work problem-specific and user-centered, see the system that produces the current outcome, focus on variation in performance, and anchor practice in disciplined inquiry. The sixth principle of accelerate progress through networked communities was not documented in the interviews or shared documents for the case. Then, I will highlight how the Southeastern High School has or has not considered equity in their improvement work.

Measurement

Similar to many districts identified in Chapter 2 in the Improvement Science research, the Southeastern District considered academic outcome and achievement data first. Participants specifically cited curriculum-based measures and unit assessments in addition to state standardized test results. Beyond academic data, other metrics considered included perception data as measured by an annual survey given to students, families, and staff as well as attendance and behavior data markers. Several years ago, managing all of the data became cumbersome, and according to the former superintendent who initiated the efforts to practice Improvement Science

in the Southeastern Schools, the district was “data rich but information poor.” This meant that there were many data points, but the storage and management made it difficult to use the data to draw conclusions and follow trends. Every building administrator created their own tools to monitor data points. Recently, the District added an administrator position focused on data analytics. This administrator worked with the Director of Technology to develop a database and dashboard for all staff to use. Several administrators at the district and school level referenced the new data dashboard throughout the interviews.

Academic Data. Based on interviews with the Superintendent, the Director of Instruction, the Director of Pupil Services, and the Director of Data Analytics, standardized test performance was the first measurement identified for annual monitoring. Some district office administrators focused on unit assessments in reading and math as important leading indicators. Some district office administrators focused on state assessment data and reading and math computer adaptive screening assessments as important metrics. At the high school level, course grades and national college admissions tests were the most important marker of academic achievement.

Perception Data. Some administrators discussed survey data as an important measurement for monitoring; however, this source of data received limited consideration in the interviews. Annually, students and parents participate in a twenty-question survey focused on perceptions of engagement in learning, engagement in school, and feelings of belonging in the school community. These ratings are reviewed annually. For the last several years, administration has not identified any specific problems for focus of disciplined inquiry.

Behavior Data. In the literature review, I rarely noticed behavior data identified as a measurement. In Southeastern District, one district office administrator considered and discussed

behavior data routinely when considering successes and opportunities for improvement. The high school principal echoed an emphasis on attendance and behavior data. In publicly reported data on the state department of education website, there is a slight disparity in attendance at Southeastern High School by race, with Black and Hispanic students having lower attendance rates. Additionally, behavior data on the state department of education website showed Hispanic students were suspended at five times the rate of their white peers, and Black students were suspended at four times the rate of their white peers. A district office administrator and the high school principal both targeted this disparate data for improvement. For them, in order to achieve the goal of preparing all students for college and career, it was important to have all students in school.

Problem Specific Focus

Improvement Science identifies a problem-specific focus as a core principle to accelerate progress. The school district website states that the “Big Aim” for the schools is for “All students successfully transition to college and work prepared for their future.” As part of this aim, leaders in each district department, school, and school teams identify a problem focus to make improvements within their area of responsibility. An example problem-specific focus surfaced in every interview in the Southeastern District, with each administrator identifying a specific problem that is preventing ALL students from being prepared for their future. One district office administrator was focused on a reading problem at the elementary level. One district office administrator was focused on discrepant suspension rates for students of color and students with disabilities. Another district office administrator was focused on improving reporting of all data points by identity markers of race, socio-economic status, disability status, gender, and language proficiency. At the school level, the high school principal highlighted focus on a problem with

math achievement for the lowest twenty percent of students in the school and the disparate suspension rates for students of color.

Bryk et al. (2015) describe a focused learning journey when an organization prioritizes a specific problem. In the Southeastern District, individual leaders were focused on different problems within the system. In the interviews, each administrator described a similar process as they focus on the problem within their scope of responsibility. The superintendent described how instructional teams focus on problems within their realm of control and collectively support improvement towards the “Big Aim,” collaborating through routine reports and conversations of all school and district leaders about their improvement focus.

Understanding the System that Produces the Outcome. Southeastern administrators and teachers developed a process of classroom observations as a strategy to analyze the system that produced the current student achievement outcomes. In many of the interviews, an administrator identified a specific problem in the data, which may include state test results, screening assessments, or unit assessment measures. Then, the data was compared across classrooms and buildings within the district. If one classroom or one building had success with the data point in consideration, then administrators and instructional coaches would observe the instruction in the specific area of focus and analyze what was different in the classroom with the desired results compared with the classrooms in which results were less desirable. Focusing on positive variants led to identifying effective, innovative instructional theories of practice.

Positive Variants. The focus on the system in one of the Southeastern elementary schools was on successful classrooms. Disciplined inquiry for Improvement Science addressed the question: which teachers were getting the best results on end of unit assessments in reading and what were they doing differently. Specifically, a district office administrator discussed

focusing on variation in performance as a principle of Improvement Science. She highlighted an elementary case in which a teacher had significantly higher scores on reading unit assessment measures than the other classrooms. In order to pinpoint the differences between the instruction in the classroom with higher assessment scores, a team including an administrator and an instructional coach conducted a series of observations of classroom instruction in reading across the elementary school. A district office administrator described an observation process used by a group of instructional coaches. After hours of observation, coaches theorized that a teacher whose students were achieving higher reading scores conferred with students individually on a regular basis and identified a specific goal or task for the student to work on at the end of the conference. The coaches worked to provide other teachers with training on this conferencing instructional practice and then observed teachers implementing the conferencing protocol, offering feedback on fidelity to the protocol. Over the course of several units, the elementary school in which the strategy was implemented saw gains in student reading achievement data.

At the high school, the principal described a replication of this process when focusing on low math achievement for students in a remedial math track. The math department at the high school was challenged because they had difficulty identifying teachers with higher rates of success and therefore, the team was struggling to develop a theory of effective practice that could improve student performance in math. After multiple cycles of disciplined inquiry, the math team had yet to measure improvement. Based on improvement cycles presented to the Board of Education over the previous two years, some of the changes that had been tried included time before leaving class to begin homework to increase homework completion rates and reducing student-teacher ratios by co-teaching the lowest level math track classes.

The principal has been working with the district office administrators for curriculum and special education as well as the math department leadership to find bright spots in other high schools with similar demographics. Through a network, the team aspired to find a strategy to consider within their math classrooms. The principal noted that other schools who are raising the achievement of the bottom twenty percent do not have low track courses. Schools analyzed by the leadership group had raised expectations for all students, and the principal shared that efforts to make this change had been met with concerns from teachers that the students would not be able to do the work. The positive variants in other systems offered insight for future possibilities for change that could result in an improvement.

Disaggregating Data. While the achievement data, attendance data, and behavior data always had the potential to be disaggregated, one district office administrator acknowledged that discrepancies in data by demographic markers had not been “the focus of our curiosity.” Other administrators described a focus on the aggregate to raise academic achievement. During the first five years of implementation of Improvement Science in the district, academic achievement was below state average on many standardized measures. Therefore, improvement efforts focused on quality instructional practices that achieved high levels of student proficiency on unit assessments and standardized test measures. Over the last three years, an improved data dashboard increased the frequency with which district and building leaders considered data disaggregated by demographic markers. I understood that problems of disparate achievement outcomes among racial and lower economic status had not risen to the surface for improvement efforts until last year.

A district level administrator and the high school principal both focused their curiosity last year and this year on discrepancies in discipline data and AP course participation by race and

ethnicity. The change process engaged only a few actors, which included the high school administrators and the high school guidance counselors. The administrators identified the problem in a data review at the end of the 2019-2020 school year.

Due to restricted in person attendance due to COVID, the team had not initiated a PDSA cycle focused on this specific problem. This focus appears to be the first improvement efforts focused on disaggregated groups of students.

Disciplined Inquiry to Improve

After a decade, the practice of Plan, Do, Study, Act (PDSA) has become routine in Southeastern School District. Every building leader and every classroom teacher is required to use the PDSA process for their annual goal. The routine of the PDSA cycle at the teacher and classroom level appears to include the recycling of the same instructional goal each year, and it is unclear if the process continues to measure improvement. The classroom PDSA tools are used to measure curriculum progress, and students do not engage in the principles around defining the problem focus or analyzing the system leading to the problem. The classroom PDSA focus is on a curriculum learning goal defined by the teacher.

Plan, Do, Study, Act Cycles (PDSA). Annually, the District leadership team sets goals for reading achievement, math achievement, and other areas documented in the strategic plan. At the building level, the principal identifies goal teams for each of the focus areas. The goals cascade from the district level to the school level to the classroom level. Each goal area documents four PDSA cycles over the course of the year. Changes, actions, and measures are documented on a form called an A3 (Appendix E), and building leaders publicly report on improvement cycles.

The high school principal shared a goal focusing on improving math performance for the lowest 20% of students in the school. The goal team includes the high school principal, a school counselor, an instructional coach, and all of the math teachers. Using a fishbone analysis, the leadership had highlighted that the students achieving at the lowest levels were taking courses in a low-level math track with lower standards and lower expectations. The instructional coach had completed a series of observations to better understand the actual practices that are producing current outcomes. The coach shared observations of instruction below grade level, noting she observed teachers making modifications to curriculum that dropped the rigor of the standards to a middle school standard. Most math teachers rejected this perception of the observed instruction. I was not able to obtain the fishbone analysis document to fully review the other hypotheses generated.

The principal shared his concern that the team was stuck in a cycle of blaming the student behaviors, attitudes, and motivation rather than understanding the system that was producing the problem. He said,

That's an area where we are struggling. [Teachers say] 'It's the kids' fault. They don't do their homework.' It is everything that they (math teachers) don't control that is the reason instead of listening to the feedback that you're getting from kids and wondering: wait a minute! There has got to be something *we* can be doing differently.

As he described the process, I wondered if asking students in the low track math course why they were scoring below proficient on assessment measures, the team could build a greater understanding of how to change the course sequence to increase expectations for the students in the lowest level course.

The PDSA documentation presented to the Board of Education supports the principal's analysis of the lack of improvement. The math teachers attempted the following changes to impact math test performance, "Establish a procedure so that more students utilize the math resource center." In addition, the math teachers stated, "Continue to improve homework and review completion rates through more parent communication, online keys and investing more time in HW before students leave the classroom." The principal expressed frustration that the theories of practice being employed and tested were not aligned with what was learned from the observations of the instructional coach and the positive variants in other systems in which low track courses were eliminated. He said, "It's been a struggle" to convince the math teachers to make a change in the course sequence or the curriculum.

Equity Work

I will now turn attention to the experiences with diversity, equity, and inclusion in the Southeastern School District. I will describe the differences between the dominant normative culture of the adults in the schools and the student populations of color. In so doing, I will provide an overview of disparities in outcomes and the attention recently given to the voices of students of color.

After significant improvements in academic outcomes during the first five years of using the tools of Improvement Science, academic outcomes have plateaued at Southeastern High School. The most significant problem contributing to the current outcomes appears to be high levels of disparity in participation in high-track courses and standardized test performance between students of color and white students. At the time of the interviews and case analysis, the team had recently reflected on potential root causes related to lower rates of students of color participating in AP courses. Following the video testimonials, the ARI group shared with the

principal and subsequently staff, the high school principal set a meeting with the two student advocacy groups focused on racial injustices. The principal reflected that he had worked with teacher teams on problem solving, but the meetings were the first time he truly listened to the students of color, some of whom were taking AP courses and some of whom were not taking AP courses. Students shared frustration that none of their teachers looked like them. The principal stated, “one of their main points to me was you don’t hire black teachers.” The students shared that they felt alone when they were the only student of color in an AP course with a white teacher.

Based on publicly available data on the state department of education website, the teaching staff at Southeastern High School is over 95% white. The efforts to see the system that produces the current outcomes depends on the perspective of those analyzing the system. In the Southeastern School District, the process of analyzing the system is conducted by school leaders, all of whom are white. Rounding conversations have not intentionally targeted students of color who are closest to the experiences that may be impacting disparate access to courses and achievement. The fishbone analysis and instructional observations were primarily conducted by a predominantly white team of adults. As such, the perceived root causes and proposed changes are filtered through the lens of the dominant normative culture in the school.

In the Southeastern School District, racial tensions boiled over in 2019. Consistent with a national trend to eliminate stereotypical depictions of people of color as mascots for sports teams, a member of the Southeastern School District school community submitted a proposal to change the Native American mascot for the high school. While the District did not focus tools of Improvement Science on this problem, I found evidence that the experience influenced current perceptions and contributed to a problem focus following the mascot review.

Mascot Change. Southeastern High School had been represented by a Native American mascot until last school year. In 2019, the superintendent, influenced by the request from the community member and state and national conversations around stereotypes and mascots, initiated a process to change the mascot for the high school. A small number of community members from the Native American community supported the change and advocated at the board level. A sizable majority of community members, including two white members of the school board, did not see the representation as problematic and loudly protested the proposed change.

During the 2019-2020 school year, the community conversations were heated and contentious. Participants in the study shared specific instances of interactions between members of the dominant white culture, representative members of the local native community, and the School Board. Many spoke publicly at Board meetings on several occasions. Some advocated for maintaining the current mascot as tradition. Members of the Indian and Native American community shared personal responses to the stereotypical images and representation of their identity. The passionate pleas to the Board about maintaining the mascot despite the perspectives of the marginalized ethnic communities sent a clear message to community members and students of color that their voice was not understood by the majority white community. Leadership recommended the mascot change, and the Board approved the elimination of the Native American mascot on a vote of five to two.

The superintendent described this issue as a flashpoint that ignited racial tension. Subsequently, students of color were moved to action, seeking to be understood and access to opportunities they had perceived were not for them. Students saw the white leaders proceeding with changes that valued the perspective of students of color despite resistance from white normative culture. Students spoke up to trusted adults prompting the formation of several student

groups to advocate for students of color. The adult advisors navigated the school systems to recognize the affinity group and advocacy group.

Student Advocacy. I interviewed a teacher who leads a student group called Awareness of Racial Inequality (ARI) at the high school. The student group is comprised of a majority of students of color. The group formed this school year, and students advocated for an opportunity to share experiences as a student of color in the United States and in the Southeastern School District. Students who joined the group perceived that their voices had not been heard. The teacher leader of the group identifies as white, and she appreciates that her race impacts her interpretations of student experiences. For this reason, she wanted all staff to hear students of color directly rather than through a white filter.

Because leaders and teachers were not asking them questions about their experiences, students had not shared some of their negative experiences as a person of color in the Southeastern Schools. Moreover, students hesitated to speak up and speak out in contradiction to white normative narratives in spaces that were predominantly occupied by white staff and white peers. As the teacher leader of ARI gave students space to share their experiences in a safe environment, she convinced members of the group to document their experiences in video. The group took video testimonials of their experiences in the Southeastern School District that were interfering with their ability to be prepared for their future. These video testimonials were presented to staff in late fall of the current school year, with student permission.

Sharing the videos with staff required encouragement from the advisors. The teacher I interviewed shared how uncomfortable students felt making their stories public. They already felt like “everyone was looking at them differently” and sharing counternarratives made them feel uncomfortable. The advisors were able to point to the advocacy around the mascot change

and the media coverage around racial inequities to assure students there was readiness among staff to see and hear them.

The teacher requested and received permission from those students over 18 to share the videos with me. I watched several of the testimonials as part of my document review. Students described experiences of being the only person of color in certain spaces. One student explained, “No matter how many friends I made, I couldn’t get over the fact that I was the only brown girl.” She felt like “all eyes were on me,” and she had greater scrutiny on her behavior because of her skin color. A senior girl who identifies as black highlighted that she had white friends and took advanced classes, but when she sat with friends of color, her white friends texted “me from four tables away because they were scared of the black table.” This was the same table that during freshman year, a boy passed and commented that “the table smelled like poverty.” In this environment, some of the students of color felt alone and isolated. In one video, a student explained how these experiences influenced class choices and the difficulty for a student of color to take college level courses with all of the white students.

The principal acknowledged the impact of these videos on him and members of the staff. He recognized that the lack of diversity on the staff and the interpretations of adults were silencing the voices of students of color. Therefore, the Improvement Science processes were not manifesting changes that result in improvements for students of color. As a leader, he was considering data on the aggregate more often than disaggregated and therefore, problems of disparity were not recognized as the root cause of problems on the aggregate. The student voices in the videos have inspired a new direction and resulted in the identification of a problem. The school principal plans to refocus change efforts next school year and apply the Improvement Science principles of measurement with disaggregated results, seeing the system that produces

the current outcomes with input from students of color, and disciplined inquiry focused on changes that consider the voices of those most impacted. The student videos are shaping a new problem focus, which began with videos and continues with the principal meeting with the ARI student group and another student group focused on equity. Recently, he listened to their perspective on the demographic composition of the faculty and heard their interest in increasing the number of faculty who identify as a person of color.

Case Summary

For over a decade, the Southeastern School District used tools of Improvement Science. School leaders and staff throughout the system were trained to analyze data, identify problems, and used disciplined inquiry to make changes and determine if the changes resulted in improvement. For the first five years, many improvements were made across the system including increased academic achievement. More recently, changes in some areas of the high school have not resulted in significant improvements and student achievement has plateaued. Some of the staff, specifically the math team, have not been engaging with the level of curiosity needed to truly see the system and conduct authentic disciplined inquiry.

In the Southeastern School District, the problem-specific focus shared throughout the participant interviews and documentation has not focused on questions or problems of equity until recently. Academic and behavior data can be disaggregated by race/ ethnicity. Perception data (student surveys) are only available as a set of aggregate data. Administrators acknowledge that the disaggregated data points were rarely considered in problem identification. Leaders identified problems on the aggregate to engage in the process of Improvement Science.

With the mascot change in 2019 and the media coverage of Black Lives Matter protests, students of color at Southeastern High School began to share their stories and perspectives on

how they felt being a person of color in a predominantly white suburban school with nearly all white teachers. At first, a few students found a teacher who they felt would listen to them. I heard how the process to identify a new mascot and the national protests led students to feel more empowered.

At the high school, the problems identified and implemented changes had only recently considered the voices of those most impacted by the problems. The contentious process of eliminating a mascot considered to be an inappropriate racial representation sparked a movement led by students. The staff who were still committed to listening to those most impacted by the problem when attempting to see the system provided students with hope that they could be heard. As the high school moves into next school year, the results of this renewed effort to apply principles of Improvement Science to these newly identified problems of racial disparities in access to high-track classes, discipline situations, and other outcome measures are yet to be determined.

Chapter 6

In each case, I articulated the application of principles of Improvement Science and highlighted how equity was or was not considered and impacted by improvement efforts. Improvements in disparities among students qualifying for free or reduced lunch or students of color appeared unplanned or had not yet occurred. My main research question for this study is: In schools using Improvement Science, how and to what extent do school leaders attend to diverse populations and equity when implementing the improvement process? I will begin with an overview of the commitment to Improvement Science principles and strategies to solidify the ethos of Improvement Science within the schools. I will outline recent considerations in each district, including problems surfacing in recent years. Next, I will turn my attention to the research question: Under what conditions do problems of inequity rise to a level of problem identification in which schools focus the principles of Improvement Science? I will analyze the conditions which led to problems of inequity rising to a level of problem identification including findings of the influence of a *champion for equity*, how a *catalyzing event* can accelerate focus on equity, and the use of a *disaggregated lens for data* analysis to identify potential problems for focus. Finally, I will summarize implications for school leaders and future considerations for continued study.

Findings

I focused significant attention in this study on the principles of Improvement Science. The cases did not surface new findings about the implementation of Improvement Science. In outlining and documenting how and to what extent Improvement Science was applied, I was able to clarify the intersection of Improvement Science with the challenges of addressing equity in schools.

Southeastern School District has been invested in Improvement Science for over a decade. Northwestern School District began their Improvement Science work five years ago. In both cases, efforts were implemented by a new superintendent charged with improving overall academic performance. In both cases, overall student achievement measures were below state averages when the district implemented the principles of Improvement Science, and the districts made a significant investment in professional development for teachers and school leaders to build capacity to employ the processes of Improvement Science.

Implementation of Improvement Science

Initially, the problem specific focus in each district centered around standardized test scores. Plan Do Study Act (PDSA) cycles focused on curriculum and instructional strategies for all students. Both districts had success in implementing principles of Improvement Science. Standardized test performance in both districts climbed in the first three years of implementation. As standardized test measures began to rise, both districts refined their metrics and considered data beyond test scores. Data focus expanded to survey perception data, attendance data, and behavior data. With an increased number of data points, leaders and staff became distracted at times making changes outside of the Improvement Science focus.

In both districts, disaggregated data was available, but it did not rise to the level of analysis to identify problems of discrepant outcomes by race, ethnicity, economic status, or language proficiency status until very recently. In both cases, dominant, normative culture was white and middle class. Considerations of equitable outcomes and recognizing disparities in the data surfaced within the last two years in each district. While it is not central to the current set of district goals in either case, administrators within each system are beginning to turn their

attention to understanding the system that contributed to the current disparate outcomes for either students of color (Southeastern) or students living in poverty (Northwestern).

As each district worked to examine the system that led to the current outcomes, administrators employed tools to consider ideas and impact from those closest to the problem. In some cases, that took the form of individual conversations with students or staff called rounding, and in other cases, it took the form of observing the work within the system, including instruction. I did not find evidence that conversations and questions focused on the experiences of specific racial/ ethnic groups or economic status. Again, seeking input from those impacted by the problem took place on the aggregate.

The findings in these two cases support existing understandings that equity is impacted only when it is the focus of the improvement process (O'Day & Smith, 2016). Furthermore, the disjointed implementation of Improvement Science in the Northwestern School District illustrates the findings in the literature review that schools identified as exemplars may not have solidified the full system of practice. Additionally, the findings affirm the questions Capper and Jamison (1993) considered, inasmuch as the processes initially focused on the voices and ideas of teachers and leaders, all of whom were white and middle class. Voices of students and parents were “muffled by the culture we see” as one administrator reflected. The subtle insight of the building administrator in Northwestern High School highlights the primary barrier to addressing problems of inequities using Improvement Science principles in schools, given the teachers and leaders reflect the white dominant culture.

In the two cases I studied, the schools were not equipped to address problems of equity using the tools of Improvement Science *yet*. Consistent with ideas of Datnow and Park (2018), Gil and Kim (2018), Jimerson (2016), and O'Day and Smith (2016), leaders must intentionally

point principles of Improvement Science at problems of inequity to impact change for improvement. Neither district case had yet fully applied Improvement Science to these problems. More recently, both high schools recognized problems of inequities within the school. O'Day and Smith (2016) argue that Improvement Science provides an avenue for leaders focused on problems with equity to address gaps. In each case, leaders were grappling with how to shift the problem-specific focus and see the system that was producing the academic, engagement, and belonging gaps among students of color and students with lower economic status.

Ideas to Improve Implementation of Improvement Science. To maximize the impact and effect of the principles of Improvement Science, each school system could make some adjustments to practice. First, focus on documentation of processes and publicly sharing the practices of Improvement Science in order to maintain focus. The District Scorecards in each district were intended to maintain focus, but this only set out goals and measures. Southeastern schools employed a form called an A3 (Appendix E) to document disciplined cycles of inquiry. Using this tool to document the focus of PDSA cycles would support improvement efforts. In the Southeastern High School, PDSA documents at the building level are publicly shared at regular intervals. This practice is valuable and Northwestern School District would benefit from this practice. Leaders could better build a shared focus and collective action to advance hypothesis and test solutions developed in collective inquiry.

Each school employed tools to seek input from voices within the system. Leaders conducted rounding with staff and built a fishbone analysis at the onset of problem identification. Leaders have an opportunity to increase the impact of this tool by actively seeking out specific voices and letting the voices of those impacted by the perceived problem define the problem. Standardized test scores surfaced as a problem, and the board and administration assumed this

problem was paramount, setting this focus for the district. Student voices were heard in the form of test performance, and parent voices were not identified in the participant interviews. Students may or may not have perceived academic or instructional problems as the most important problem or the root of a problem. By engaging those voices in the fishbone process, new perspectives and insights would likely emerge. In both cases, seeing the system through a fishbone analysis occurred with a team of staff. I did not find evidence that students were involved. Seeking out voices in the fishbone process could elevate the clarity of understanding of the root cause. Schools should continue to engage those voices in developing hypothesis to make a change.

At both schools, students of color had stepped forward to speak with adults in the school about problems they were experiencing. These problems had not surfaced from the routine data review previously. Leaders and staff are beginning to see problems that had been previously unrecognized. Southeastern High School, leaders anticipated shifting into disciplined inquiry around the problem as it comes into clear focus. I grappled with what prompted students to volunteer their voice when the leaders had not sought them out. The following factors appeared significant in surfacing problems of inequities, leading to a future problem focus using Improvement Science.

Focusing on Problems of Equity. The second research question asks: Under what conditions do problems of inequity rise to a level of problem identification in which schools focus the principles of Improvement Science? I will now highlight conditions which led to recognizing problems of disparate experiences for students of color and students qualifying for free or reduced lunch. These conditions include: *a champion for equity* which includes staff who have lived experiences of marginalization and are able to perceive perspectives incongruent with

the dominant culture (in these two cases), *a catalyzing event* that brings the problem to the collective consciousness in the school community, and *the measurement of discrepancies* based on data disaggregated by demographic markers.

Champion for Equity

In each district, I identified a leader who had become a champion for equity. In the Northwestern School District, the high school principal truly saw the assets of students and families living in poverty and understood the challenges in the school setting. She recognized there were voices missing in the efforts for improvement. During her interview, she passionately shared the exploration of identity and her identity journey, recognizing how her point of view and dominant culture influenced her leadership and the school setting. Because economic status is not visible to others on a daily basis, staff do not see or notice differences among the population of students living in poverty.

While Northwestern School District has a small percentage of students of color, the principal spoke to both economic and racial diversity in her reflections. Speaking about the staff, she said, “We’re even afraid to talk about [race] because ... we don’t know how to say it. We are afraid of making a mistake.” As a champion, she made herself vulnerable by investing time in trying to understand different perspectives on race and leading staff to reflect on identity and begin to see race.

In the Southeastern School District, the director of pupil services heard and valued the voices of students with disabilities and students of color. She recognized the discrepancies in the discipline data at the high school and the disproportionate participation in AP courses. She developed a sense of wonder and engaged the high school principal who was learning to listen to the voices of the students of color.

More recently, a teacher leader has emerged as an advocate for students of color, willing to elevate their voices in the schoolwide conversation. The teacher leader articulated “students need a safe space in which they can speak freely” about their experiences as a student of color. As a champion, the teacher leader acted to disrupt the power structure that may limit students of color from expressing their understandings of the problems (Capper & Jamison, 1993).

To date, the champions for equity in each district have not advanced the problem-specific focus of equity to a full analysis using the principles of Improvement Science. In each case, the champion led others into uncomfortable conversations around racial identity and began to help others see the differences and potential impact of silence on students and families of color and families living in poverty in the school community.

Lived Experiences. In each situation where equity became a consideration recently, the champion had lived experiences that were incongruent with the dominant culture in a structured setting. These periods of time shaped their perspective and increased their awareness of differing perspectives in other settings and for other marginalized groups. In one case, the champion was raised in a low-income home and had pride in the values and work ethic that circumstance instilled in her. In one case, the champion parented a student with a disability. These lived experiences shaped their lens. It became a pair of glasses through which they saw the humanity, the dignity, and the outcomes for underrepresented voices in their schools. In each case, the champion recognized students and voices that were being muffled in the existing processes and practices.

At Northwestern High School, several white male administrators from middle class homes looked at engagement data, perception data, and grades. They saw deficits in the students. They saw a need to lend a helping hand to make sure they had food to eat or the tools to do their

schoolwork. The female principal who was raised in poverty saw students who are “living in cycles of abuse and addiction.” She understood it wasn’t just about giving them food for the weekend, but it was also about “asking the right questions and getting the right resources so kids are able to see other options for themselves.” With this perspective, she recognized the humanity of students living in poverty.

The lived experiences impacted how each saw the problem and interpreted the root causes and perceived the voices of students experiencing the problem. This affirms hypotheses from Capper and Jamison (1993), Datnow and Park (2018), Gil and Kim (2018), and Jimerson (2016) that Improvement Science principles alone do not address equity. While O’Day and Smith (2016) suggest Improvement Science principles can be focused on problems of equity for meaningful change and ultimately improvement, Southeastern and Northwestern school leaders had not actively sought to focus the principles, primarily because the majority of lived experiences were ignorant of the lived experiences of those most impacted by problems of inequity. Those with lived experiences relied on their lived experiences in interpreting the sources of data, becoming a champion for seeing the problems differently.

In the absence of lived experience, Improvement Science provides a set of tools to engage the voices of those with the lived experience that differs from staff. External champions such as parents and community members could provide insights for problem identification, solution generation, and monitor measurement for accountability to these stakeholders. Most importantly, students could be champions when a staff member steps forward to empower their voice and include them at the table. In each case, the champion for equity was a staff member. These champions did not act alone, but they served as the conduit to give power to those whose voices had been previously silenced or underrepresented.

Perspective Taking. Sometimes the champion did not have a lived experience in the identity group which had discrepant outcomes. In the Southeastern School District, the champion for equity was not a person of color, but she shared how she immersed herself in developing an understanding by reading and listening to pod casts that helped her see perspective. Her readiness to engage in this work stemmed from her lived experiences being marginalized as a member of a different marginalized group.

In the Southeastern District, teams had begun to focus on racial and ethnic identity and understanding racial bias in theory. A teacher leader at the high school invested in listening to a group of students who had been inspired by national events and the local process of changing the mascot. In the absence of a lived experience as a person of color or a person living in poverty, champions for equity actively seek to understand perspective by listening and elevating the voices of those muffled by the dominant culture in the school.

Catalyzing Event

In history class, students study events that initiated a change or a movement. In recent years, several events of racial injustices have spurred the formation of national advocacy groups with local chapters. Events in 2020 impacted both case studies in reframing their perspective. These events included the murder of George Floyd by police in Minneapolis, a series of Black Lives Matter protests and local efforts to eliminate a racially stereotypical mascot. Champions for equity in each school were able to draw upon the catalyzing event to accelerate recognition and escalate the problem into focus for future Improvement Science application. Students witnessed the changes and began to believe that their voices matter and would be heard.

National Ignition. In interviews in both districts, leaders discussed how students engaged in conversations with adults around the Black Lives Matter protests following the

murder of George Floyd. In the Northwestern School District, the student leader was a recent graduate, and the conversations led to identity reflections for district and school leaders. While all those interviewed were apprehensive and uncomfortable with the conversations, there was a realization that these perspectives are important in seeing the system and considering changes for improvement.

In the Southeastern School District, the summer protests coupled with a local mascot change inspired students of color to seek a faculty advisor for a student advocacy group. Three current students spearheaded a project to share their experiences and the voice of other students of color to shine a light on the school system. In those video narratives, students surfaced a theme of historically suppressing their feelings and experiences around race with their white friends and minimizing their struggles at school with adults, who are almost exclusively white. Even when asked for feedback through surveys or individual conversations, students did not reveal aspects of their experiences tied to race.

The principal and teacher advisors perceived that the Improvement Science processes sought out the voices of all students. After the summer of protests, the advocacy and affinity groups that formed empowered a collective voice from students of color. The teacher advisors heard some stories, ideas, and perspectives about how students of color in Southeastern High School experienced their school for the first time in the fall of 2020. Even though students had been asked on perception surveys, students indicated they held back on sharing ideas that ran counter to the white normative culture. The catalyzing event impacted students' readiness to share their truth with the adults in the school who were willing to listen. The catalyzing event also increased white adults' awareness of the voice of students of color in the school. The student

video perspectives are shifting how adults who identify as white understand the problems in the data and seek insight to propose changes for improvement.

Local Spark. A local spark can accelerate focus on disparities within the school system. The contentious process of changing the high school mascot in the Southeastern School District prompted new perspectives. Students recognized that their opinions mattered when leadership advocated to eliminate the racist mascot in the face of resistance from the white community. This change influenced students' and families' willingness to share more accurate reflections on problems they face in the current school and community culture because it brought hope that their voice would impact future outcomes. While this experience was quickly followed by a national event that also inspired advocacy, evidence in interviews with school administrators and teacher leaders point to student movement in early 2020. The formation of the student group and subsequent interest in understanding from school staff was influenced by both the local events and the national protests.

Measurement of Discrepancies

Bryk et al (2015) emphasizes measurement as an important principle to improvement. If a system only looks at aggregate data, it is difficult to see and identify problems of inequity. In both cases, early implementation of principles of Improvement Science focused exclusively on aggregate data.

Disaggregating the Data. As student voices of color were elevated, leaders in both school districts had an increased curiosity in understanding the impact on student academic outcomes and perceptions. While standardized test score data had been seen in a disaggregated format by race, economic status, disability status, and language proficiency status, school leaders did not focus on this problem as a point of improvement. The problem of the aggregate was more

pressing initially. Both districts were in their infancy when it came to equity focus. Each district was in the first year or two of truly seeing and recognizing disparities in student experiences that lead to gaps in student outcomes.

Through these case studies, some speculation in the existing literature was affirmed. Equity must be considered deliberately and intentionally in the application of Improvement Science principles within a school system. In both cases, school leaders had not attended to diverse populations nor issues of equity in the defined improvement efforts. O'Day and Smith (2016) and Bryk (2017) identified opportunities and possibilities in which Improvement Science could be employed to impact change that results in improved equitable outcomes for students. In this study, I have outlined conditions in which schools arrive at a point in which attention shifts to equity for this application to be possible. School and district leaders have begun to shift their attention to problems of equity based on recent awareness of discrepancies in outcomes, following catalyzing events that raised attention and consideration in the school. I will now turn my attention to recommendations to employ these findings to accelerate focus on addressing problems of inequity.

Considerations for Improvement Science Work

In some systems, Improvement Science is seen as an initiative of its own. In both Northwestern School District and Southeastern School District, district level leadership focused on the implementation of Improvement Science as the goal. Bryk (2017) references Improvement Science as a set of tools. As such, who uses the tools, the perspective of stakeholders and the goals of the organization all matter. Improvement Science will only solve problems of equity when focused on problems of equity.

Leaders with power within the system have the ability to elevate or suppress voices and ideas. One of the tools of Improvement Science provides strategies to elicit input from stakeholders within the school community, which Bryk et al. (2015) refer to as user centered. It is possible to be user-centered and only truly hear a limited number of voices by elevating those ideas that resonate with the leader and ignoring ideas that may disrupt the system. Within this study, I identified actions which can support leaders in avoiding the pitfalls of replicating bias within the system. Champions for equity who view the system through a lens of Critical Theory (Capper, 2019) are instrumental to successful use of Improvement Science to address problems of inequality. Who is listening and acting on problems and ideas within the system is critical to which problems get addressed for the students and families.

Throughout the literature and again in the cases, leaders begin with measurement as the most important tool of Improvement Science. While measurement is important to determine if changes are an improvement and can be a source to identify problems, this tool cannot be the beginning. When leaders start with measurement, the system relies on information that has already been measured and more importantly, information valued by those who have had power historically. Apple (2012) and other critical theorists argue this is to be expected and not surprising. During participant interviews, I began with questions about what data is collected and analyzed. No participants started with sharing a problem that required them to figure out how to measure and collect data. All participants started with data an external group had created a tool to measure such as test scores, attendance, or discipline. These data sources empower the establishment and the existing bureaucratic structures that are perfectly designed to get the current outcomes. A user-centered approach with an equity focus empowers different voices.

In the text *Learning to Improve*, Bryk et al. (2015) begin with problem-focused, user-centered, and seeing the system before discussing measurement. Implementation of Improvement Science should happen in this order, with an equity focus helping to “see the system” rather than traditional measures defining the problem. In this way, measurements are pointed at the user-defined equity problems in the system rather than shifting problems to match the measurements the schools already collect.

The Southeastern High School is poised to apply Improvement Science tools in this order. In essence, a catalyzing event *is* a problem in the system. This school year, students identified a problem, which they articulated in video. The students told school leaders that students of color don’t feel like they fit in with their peers and have to suppress aspects of their identity to be successful on traditional measures. Student leader groups have begun to consider how the system is contributing to this experience, such as staff “don’t look like me” and white kids avoid a table when a group of black students sit together. Harnessing these ideas through the tools of Improvement Science, leaders should work with students to consider how to target a change at one aspect of the problem and how to measure if the change results in an improvement.

In schools implementing Improvement Science, all the tools are necessary to effectively address problems within the system. Applying the tools in the order intended by Bryk et al. (2015) empowers the voices impacted within the local context rather than empowering the historical structures. In order to address problems of inequity, leaders require a lens of critical theory to approach Improvement Science. This combination provides greater opportunity for equitable schools.

Implications for School Leaders

Since the early 1990's, achievement and opportunity gaps are well documented in the United States for students of color, students qualifying for free and reduced lunch, students learning English for the first time, and students with disabilities. Achievement gaps refer to discrepant outcomes on academic measures such as standardized tests and grades. Opportunity gaps refer to reduced access to resources such as rigorous, college-level coursework, extra-curricular activities, or learning materials and facilities. After thirty years, these gaps persist with limited examples of gap closures by race or economic status. Improvement Science emerged over the last two decades as a systematic approach to targeting improvement in a local context.

Improvement Science holds the promise “to detect and learn from variations in practice, to build and share knowledge among practitioners, to motivate frontline innovators” (Lewis, 2015, p. 55). At the same time, depending upon shared knowledge, shared identity, and common lived experiences among the dominant normative culture within the system may replicate, amplify, and perpetuate the bias of the system if the focus is not disrupted. If the shared knowledge is viewed through the lens of Critical Theory (Capper, 2019), the potential to see problems of inequity is more likely. Champions hold the knowledge of practitioners that can amplify voices of marginalized populations who could otherwise be viewed as outliers.

In schools implementing Improvement Science, I recommend two focused efforts to create conditions that support the likely emergence of problems stemming from inequities. First, equitable outcomes should be the problem focus in order to impact change that results in improvements. School leaders can support focus on problems of equity by ensuring all data collected and reviewed in the system is disaggregated by demographic markers. Furthermore, school leaders should recognize outcome measures beyond traditional means defined by the

dominant class and race interests. This means focusing on the outcomes valued by stakeholders in the school community and will likely require seeking out additional perception data and differing forms of measuring academic success. To accomplish this, leaders will need target voices of underrepresented races and classes to recognize and consider alternative outcome measures. In addition, attending to the differences in outcomes by these demographic markers are critical to recognizing and elevating disparities to a level of problem focus.

Then, leaders who identify white and have experiences congruent with the dominant culture in the school community should reflect on this identity and lived experiences in the context of understanding the system. Students who had recently spoken out in both the Northwestern and the Southeastern High Schools shared that they had suppressed their truth in experiencing problems because they perceived that it would not be heard and would present a risk to their success in the environment. Students needed to feel emotionally safe revealing their true perspective. In order to elevate voices of identities and lived experiences most impacted, leaders could identify staff, students, community members, and other leaders who share a lived experience as a member of a traditionally marginalized group. These intentional efforts can support the identification of champions for equity. Once a school leader identifies champions for equity, school leaders should put them in positions to lead Improvement Science activities that generate root causes, target changes, and set measures to engage in discipline inquiry. Students and families most impacted by the problem should be engaged in defining the problem and developing hypothesis for improvement. In so doing, the improvement efforts can focus on sharing knowledge and building knowledge from varying perspectives. These reflections on lived experiences and identities of staff coupled with champions of equity among staff and students could serve to amplify the voices of muffled demographic groups.

Most importantly, school leaders should recognize catalyzing events as symptoms of larger problems. Events occur in our communities, in our regions, and throughout the country every day that make for moments of learning and elevate stories of marginalization and oppression. School leaders and champions of equity should recognize these as potential catalyzing events. Leaders can raise events of oppression and marginalization to the collective consciousness as a strategy to turn a moment into a catalyst for learning and growth.

Opportunities for Future Research

These case studies offer opportunity for future study. The limitation of time prohibited a longitudinal analysis of impact on problems of inequities. Because the cases had not fully identified problems of inequities or disparities at the time of the study in previous problem identification and cycles of disciplined inquiry, further study would focus on full application of Improvement Science principles focused on disparate experiences based on race/ ethnicity, economic status, English language proficiency or other demographic markers which surfaces substantial differences in outcomes.

As a practitioner in a school system in the formidable years of implementation of the principles of Improvement Science, I am beginning to internalize some of the findings to address problems of inequity. All student outcome data is considered both on the aggregate and disaggregated by race, economic status, English language proficiency status and disability status. Similar to the two cases, the community I serve has experienced the national ignition associated with Black Lives Matter protests during the summer of 2020, and a suspension in 2018 centered around racial slurs leading to student protests served as a local spark that raised collective awareness. Every school has identified a number of staff who serve as champions for equity. The champions are working to create safe spaces for students to express their truth about experiences

in the school system and seek their input in defining problems. I will continue to use my learning to focus principles of Improvement Science on problems of disparate experiences in the schools I serve.

Conclusion

Education in the United States has long been lauded as the great equalizer. Education is the difference for everyone to live their American Dream and change their social class if they work hard enough. I wholly believed this promise when I became a teacher. I claimed that I did not see race or class. I only saw students full of promise and opportunity.

My research and my experiences as an administrator in public schools has shone a light on inequities in the systems. Gates and barriers and bias interfere with these promises made. Improvement Science is only as good as the knowledge of the leaders and people in the system to see the structures from different perspectives. In this study, I observed two cases in which school and district leaders were committed to making changes and measuring if the changes were improvements. In these systems, the leaders and the teachers did not share racial or economic identities with a significant number of the users of the system. As leaders in the system raised their awareness of disparate experiences through a lived experience that differed from the dominant culture in an institution or through a catalyzing event that brought to light the bias and inequity being replicated in the system, they sought to do better for *all* students.

In these cases, participants saw the system differently when white teachers and school leaders started seeing color. I recognized that being color blind and not seeing social classes interferes with opportunities to improve the structures in schools. School and district leaders in these cases recognized new problems that needed to be considered when they attended to

disaggregated data. Furthermore, they are beginning to listen to the voices of students of color and those from lower economic classes to understand the differences in their experiences.

As I continue to lead in the school system that educated me, I hope to apply these learnings to improve the school experience for all the students in my care. I want the students of color to know that I see them and hear them. I want students from lower economic classes to know that their voice matters as much as those with more money. I hope that every student who graduates from the public schools I serve will believe that our schools were designed for them and each of them achieves their dreams through an educational system that listened and made changes that are improvements to their experiences.

References

- Apple, M.W. (2012). *Education and Power*, revised classic edition. New York, NY: Routledge.
- Arif, M., & Smiley, F.M. (2003). Business-Like Accountability in Education: Commentary on Morrill to Baldrige. *Education*, 123(4), 740-762.
- Arif, M., Smiley, F.M. & Kulonda, D.J. (2005). Business and Education as Push-Pull Processes: An Alliance of Philosophy and Practice. *Education*, 125(4), 602-614.
- Aumiller, B.E., & Hackmann, D.G. (2009). Top-Down and Bottom-Up Leadership: Can They Meet in the Middle? *ERS Spectrum*, 27(4), 9-21.
- Blankstein, A. (1996). Eight Reasons TQM Can't Work. *Contemporary Education*, 67(2), 65-69.
- Bryk, A. S. (2017, March 27). *Redressing Inequities: An Aspiration in Search of a Method*. Address presented at Summit on Improvement in Education, San Francisco, CA. Retrieved March 21, 2019, from <https://www.carnegiefoundation.org/resources/publications/redressing-inequities-an-aspiration-in-search-of-a-method/>
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How Americas schools can get better at getting better*. Cambridge, MA: Harvard Education Press.
- Burris, C.C. & Welner, K.G. (2005). Closing the Achievement Gap by Detracking. *Phi Delta Kappan*, 86(8), 594-598.
- Capper, C.A. (2019). *Organizational Theory for Equity and Diversity: Leading Integrated, Socially Just Education*. New York, NY: Routledge.
- Capper, C.A., & Frattura, E.M. (2007). *Leading for Social Justice: Transforming Schools for All Learners*. Thousand Oaks, CA: Corwin Press.

- Capper, C. A., & Jamison, M.T. (1993). Let the Buyer Beware: Total Quality Management and Educational Research and Practice. *Educational Researcher*, 22(8), 25-30.
- Cohen-Vogel, L., Cannata, M., Rutledge, S.A., & Socol, A.R. (2016). A Model of Continuous Improvement in High Schools: A Process for Research, Innovation Design, Implementation, and Scale. *Teachers College Record*, 118(13), 1-26.
- Cresswell, J.W. (2013). *Qualitative Inquiry & Research Design: Choosing Among Five Approaches*. Thousand Oaks, CA: SAGE Publications, Inc.
- Datnow, A., & Park, V. (2018). Opening or closing doors for students? Equity and data use in schools. *Journal Educational Change* 19(2), 131-152.
- Deming, W. E. (1993). *The new economics: For industry, government, education*. Cambridge, MA: MIT Press.
- Gelberg, D. (2007). The Business Agenda for School Reform: A Parallel Universe. *Teacher Education Quarterly*, 34(2), 45-58.
- Gil, E., & Kim, T. (2018). Response to “Redesigning Systems of School Accountability”: Addressing Underlying Inequities. *Education Policy Analysis Archives*, 26(9), 1-5.
- Glaser, B. (1965). The Constant Comparative Method of Qualitative Analysis. *Social Problems*, 12(4), 436-435.
- Hattie, J. (2012). *Visible Learning for Teachers: Maximizing Impact on Learning*. New York, NY: Routledge.
- Jimerson, J.B. (2016). How are we approaching data-informed practice? Development of the Survey of Data Use and Professional Learning. *Educational Assessment, Evaluation & Accountability*, 28(1), 61-87.

- Jimerson, J.B., & Reames, E. (2015). Student-Involved Data Use: Establishing the Evidence Base. *Journal of Educational Change*, 16(3), 281-304.
- Kelso, R. S. (2012). Origins of the Service Quality Evolution: From Industry to Academia. *Global Education Journal*, 2012(2), 110-122.
- LeMahieu, P.G., Nordstrum, L.E., & Greco, P. (2017). Lean for Education. *Quality Assurance in Education: An International Perspective*, 25(1), 74-90.
- Lemire, S., Christie, C. A., & Inkelas, M. (2017). The methods and tools of improvement science. In C. A. Christie, M. Inkelas & S. Lemire (Eds.), *Improvement Science in Evaluation: Methods and Uses. New Directions for Evaluation*, 153, 23–33.
- Lewis, A.E., & Diamond, J.B. (2015). *Despite the Best Intentions: How Racial Inequality Thrives in Good Schools*. New York, NY: Oxford University Press.
- Lewis, C. (2015). What is Improvement Science? Do We Need It in Education? *Educational Researcher*, 44(1), 54-61.
- Marshall, C., & Rossman, G.B. (2016). *Designing Qualitative Research*. Thousand Oaks, CA: SAGE Publications, Inc.
- Merriam, S.B. (2009). *Qualitative Research: A Guide to Design and Implementation*. San Francisco, CA: Jossey-Bass.
- O'Day, J.A., & Smith, M.S. (2016). Quality and Equality in American Education: Systemic Problems, Systemic Solutions. In I. Kirsch & H. Braun (Eds), *The Dynamics of Opportunity in America* (pp. 297-357). Educational Testing Service.
- Park, V., Daly, A.J., & Guerra, A.W. (2012). Strategic Framing: How Leaders Craft the Meaning of Data Use for Equity and Learning. *Educational Policy*, 27(4), 645-675.

- Rappaport, L.A. (1996). Total Quality Management: One High School's Experience. *Contemporary Education*, 67(2), 72-75.
- Schargel, F.P. (1996). Why We Need Total Quality Management in Education. *Total Quality Management*, 7(2), 213-215.
- Schmoker, M., & Wilson, R.B. (1995). Results: The Key to Renewal. *Educational Leadership*, 52(7), 62-64.
- Steyn, G. (2006). The Realisation of Quality Inviting Schools in the United States of America. *International Journal of Educational Leadership Preparation*, 1(2), 1-13.
- Weller, L.D., & McElwee, G. (1997). Strategic Management of Quality: An American and British Perspective. *Journal of Research and Development in Education*, 30(4), 201-213.
- Weller, L.D., & Weller, S.J. (1998). Raising Test Scores Through the Continuous Improvement Model. *Clearing House*, 71(3), 159-165.
- Westfall, J.E., Peltier, J.W., & Sheehan, J. (2005). Improve Schools with Empowerment Based Models. *Quality Progress*, 38(8), 55-61.

Appendix A

Administrator Interview Protocol

Introduction

1. Purpose of Study: I am interested in understanding the role of equity in improvement science. I am specifically curious how schools consider and address problems of inequity using the tools of continuous improvement.
2. Internal Review Board Protections
3. Informed Consent
 - a. Do you have any questions before we begin?

Background Information

4. What is your current position and how long have you been in this school district?
5. What made you decide to become an administrator here?

Implementation of Improvement Science

6. What does Improvement Science look like in your school/ district?
7. As you think about your schools use of improvement science, who are the people you serve?
8. What data is collected from and/ or for the people you serve?
9. Can you tell me about how data identified a problem and became the focus for improvement? How did you know it was a problem?
10. How are data collected and shared in your school(s)?
11. How did you examine the system that produced that outcome? Who is involved?

12. What process did you use to generate solutions? Who is involved? Who are the relevant actors? How did you get the trust of the marginalized groups so that they participated and had their voices heard?
13. How did you plan to measure the improvement effort?
14. Tell me about the improvement cycles. What did the data tell you about the improvement you made?
15. What documentation is used to monitor improvement efforts?
16. What members of your team are responsible for disciplined inquiry or PDSA cycles?

Addressing Diverse Populations and Equity Focus

17. Thinking about yourself as a leader and a person, have you engaged in any diversity training or equity work? If so, how has this learning impacted your perspective?
18. Have other leaders or teachers in the district engaged in diversity training or equity work? How much time has been committed to this work?
19. Has equity work in your school or district impacted the use of Improvement Science?
20. The state reports demographic composition and achievement in your school by race, socio-economic status, English Language proficiency status, and disability status. Are there any demographic markers you consider beyond those reported by the State?
21. As you think about the data you collect, how often do you disaggregate the data by race, economic status, disability status, or English Language proficiency levels?
 - a. If so, has the data surfaced any problems for focus for improvement between demographic groups?
 - b. If so, how did you generate ideas for improvement? What solutions were considered?

22. How do you ensure all people you serve have the power and ability to communicate their needs with the school?
23. Do you perceive improvement efforts have impacted groups of students differently?
How do you know?
- a. How have improvement efforts impacted students of color in your district?
 - b. How have improvement efforts impacted students living in poverty in your district?
 - c. How have improvement efforts impacted students labeled English Learners?
 - d. How have improvement efforts impacted students with disabilities?

Conclusion

24. Is there anything else that you think it is important for me to know or understand about the implementation of Improvement in your district?
25. Is there anything else that you think is important for me to know or understand about the diversity in your district?
26. Are there staff who have been critical of the use of Improvement Science?
27. Do you have any questions for me?

Appendix B

Teacher Focus Group Protocol

Introduction

1. Purpose of Study: I am interested in understanding the role of equity in improvement science. I am specifically curious how schools consider and address problems of inequity using the tools of continuous improvement.
2. Internal Review Board Protections
 - a. Informed Consent
 - b. Do you have any questions before we begin?

Background Information

3. What is your current position and how long have you been in this school district?
4. What made you decide to become a teacher here?

Implementation of Improvement Science

5. What does Improvement Science look like in your school?
6. What training did you receive in this method?
7. As you think about your schools use of improvement science, who are the people you serve?
 - a. What data is collected from the people you serve?
 - b. Can you tell me about how data identified a problem and became the focus for improvement? How did you know it was a problem?
8. What process did you use to generate solutions? Who is involved?
9. How did you plan to measure the improvement effort?

10. Tell me about the improvement cycles. What did the data tell you about the improvement you made?

Addressing Diverse Populations and Equity Focus

11. How has the school engaged in any diversity training or equity work?
12. How much time has been committed to this work?
13. Has equity work in your school or district impacted the use of improvement science?
14. As you think about the data you collect, how often do you disaggregate the data by race, economic status, disability status, or English Language proficiency levels?
- a. If so, has the data surfaced any problems for focus for improvement between demographic groups?
 - b. If so, how did you generate ideas for improvement? What solutions were considered?
15. Do you perceive improvement efforts have impacted groups of students differently? How do you know?
- a. How have improvement efforts impacted students of color in your district?
 - b. How have improvement efforts impacted students living in poverty in your district?
 - c. How have improvement efforts impacted students labeled English Learners?
 - d. How have improvement efforts impacted students with disabilities?

Conclusion

16. Is there anything else that you think it is important for me to know or understand about the implementation of Improvement in your district?

17. Is there anything else that you think is important for me to know or understand about the diversity in your district?

18. Do you have any questions for me?

If you have ideas that you need to share with me after this session, please reach out to me.

I can arrange for a conversation privately.

Appendix C

Document Analysis Protocol

Name of Document:

How is the document used in the school?

1. What components of Improvement Science are evident in the document?
2. To what extent are demographic disparities or demographic considerations identified in the document?
3. How are students of color, students receiving free or reduced lunch, students labeled English Learners, or students with disabilities treated within the document? Is there evidence of isolation or inclusion?
4. Explicit references to equity...

Appendix D

	% Asian	% Black	% Hispanic	% White	% of students receiving Free/ Reduced Lunch	% of students labeled Limited English Proficiency	% of students with disabilities
Total School Population							
Proficient on state assessment (2019)							
Proficient on state assessment (2018)							
Proficient on state assessment (2017)							
Proficient on state assessment (2016)							
Attendance rate or graduation rate (2019)							
Attendance rate or graduation rate (2018)							
Attendance rate or graduation rate (2017)							
Attendance rate or graduation rate (2016)							
Participation in AP (2018)							
Participation in AP (2017)							
Participation in AP (2016)							
Participation in extra-curricular activities (2017)							
Participation in extra-curricular activities (2016)							

Appendix E**A3 Template to Document Disciplined Inquiry**

School/Division:	Owner:	Strategic Priority:									
Annual Goal(s) Being Addressed:											
Background Information:											
First Short Cycle (August-November):											
Current Measurement Overview (August-Sept): What do data tell us about current state? What is the problem focus?											
Measurement Target (August-Sept): How will we know if our changes are improvements or not?											
Proposed Improvement Actions for Short Cycle (Sept-Nov): What will we do this short cycle to reach the intended results? What theory of practice will you test?											
<table border="1"> <thead> <tr> <th data-bbox="82 1262 930 1329">Action</th> <th data-bbox="930 1262 1227 1329">Action Owner</th> <th data-bbox="1227 1262 1528 1329">Timeline</th> </tr> </thead> <tbody> <tr> <td data-bbox="82 1329 930 1430"></td> <td data-bbox="930 1329 1227 1430"></td> <td data-bbox="1227 1329 1528 1430"></td> </tr> <tr> <td data-bbox="82 1430 930 1535"></td> <td data-bbox="930 1430 1227 1535"></td> <td data-bbox="1227 1430 1528 1535"></td> </tr> </tbody> </table>	Action	Action Owner	Timeline								
Action	Action Owner	Timeline									
End of Cycle Results (November): How did team do in completing outlined improvement actions? What was the short cycle result? Did you reach or exceed your predicted result? What was learned to guide priority for next short cycle?											

 Second Short Cycle (December-March)

Current Measurement Overview (Nov-Dec): What do data tell us about current state? What is the problem focus?

Measurement Target (Nov-Dec): How will we know if our changes are improvements or not?

Proposed Improvement Actions for Short Cycle (Dec-Mar): What will we do this short cycle to reach the intended results? What theory of practice will you test?

Action	Action Owner	Timeline

End of Cycle Results (March): How did team do in completing outlined improvement actions? What was the short cycle result? Did you reach or exceed your predicted result? What was learned to guide priority for next short cycle?

Third Short Cycle (April-June)

Current Measurement Overview (April): What do data tell us about current state? What is the problem focus?

Measurement Target (April): How will we know if our changes are improvements or not?

Proposed Improvement Actions for Short Cycle (April-June): What will we do this short cycle to reach the intended results? What theory of practice will you test?

Action	Action Owner	Timeline

End of Cycle Results (June): How did team do in completing outlined improvement actions? What was the short cycle result? Did you reach or exceed your predicted result? What was learned to guide priority for next short cycle?