Title

Preparing Schools to Successfully Participate in Networked Improvement Communities: A Case Study of Year 1 of a Math Instructional Network

Author

Rohanna, Kristen

Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Preparing Schools to Successfully Participate in Networked Improvement Communities: A Case Study of Year 1 of a Math Instructional Network

A dissertation submitted in partial satisfaction

of the requirements for the degree Doctor of Philosophy

in Education

by

Kristen Rohanna

2018

ABSTRACT OF THE DISSERTATION

Preparing Schools to Successfully Participate in Networked Improvement Communities: A Case Study of Year 1 of a Math Instructional Network

by

Kristen Rohanna

Doctor of Philosophy in Education

University of California, Los Angeles, 2018

Professor Christina A. Christie, Chair

Networked improvement communities provide a promising approach for improving education's most pressing problems. By uniting diverse practitioner experiences with expert subject knowledge and an improvement science evaluative framework, networks of multiple schools can collectively solve persistent educational challenges. However, networks have their own challenges. Teachers and administrators are required to learn and apply improvement science methods in order to successfully participate in the network. Given their extensive daily duties, learning these technical tools may be difficult and daunting. The challenge of building improvement science capacity can be further exacerbated by the diversity of multiple school contexts.

Through an explanatory single case study, this study examines how to prepare five schools to successfully participate in one networked improvement community, which sought to improve math instruction. This study utilizes interview, observation, surveys, and document and artifact analyses to provide a detailed narrative of the network's first year. Specifically, it investigates how the network hub built the improvement science capacity of its members, and what facilitated improvement science implementation by the schools, including how meaningful learnings were generated.

The study imparts three broad categories of findings. First, the findings confirm the importance of building a network hub team with expertise in the both the content area and improvement science. These experts should work closely together to establish two synergistic visions for building mutual capacity, whereby teachers can learn the technical tools and methods within the context of their classrooms. Second, by working closely with schools and establishing consistent in-school collaborative work structures, the network hub team can foster consensus and coherence in the fundamental network components and support educators' ability to complete Plan, Do, Study, Act (PDSA) cycles. Third, network learnings are primarily generated through the PDSA structure, particularly its reflective practice component. In education, where complex problems may be related to underlying instructional assumptions about student learning, networks should consider how to cultivate double-loop learning in addition to the more instrumental improvements of single-loop learning. Finally, this study offers a model for new networked improvement communities to follow when building the improvement science capacity of educators.

The dissertation of Kristen Rohanna is approved.

Kathryn M. Anderson

Louis M. Gomez

Jody Z. Priselac

James W. Stigler

Christina A. Christie, Committee Chair

University of California, Los Angeles

2018

To Dara, my biggest supporter, without you none of this would be possible;

my mother, who instilled my love of learning and encouraged me to dream;

and my father, who modeled integrity and hard work, and continually inspires me to be the best person I can be.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I have been fortunate to have wonderful people guide me throughout my graduate and research career. Foremost, Tina Christie, my advisor and friend, who shared and championed my practical research interests. I could not have asked for a better advisor. Her support led me to this network and this dissertation. Her professional and scholarly guidance over the last four years has been invaluable, and has made me a better researcher. I am grateful to the rest of my committee members whose thoughtful proposal critiques re-directed my questions and helped me shape this research. I believe this research is better for it. I am also indebted to Mike Rose, my writing guru, mentor, and friend, who provided guidance on my writing.

Incredible thanks is owed to the teachers and administrators who participated in this networked improvement community. Their dedication, commitment, and desire to improve for their own students inspired me every day, and encouraged me to do my best, both in my research and the network. Great appreciation is due to Jody Priselac and the hub team, Theo, Sarah, and June. It was a privilege and joy to learn and work with them throughout this project. Special gratitude is also owed to the Carnegie Corporation of New York, whose funding made this important network possible.

To the SRM faculty, when I walked into Moore Hall four years ago, I had no idea how lucky I was. You truly are the most supportive, thoughtful, and kind faculty. I have enjoyed learning from you all.

I would be remiss if I did not acknowledge Jeff Tayman. I am truly grateful for his mentorship and friendship over the last 15 years. His no-nonsense advice, encouragement, and exemplar of research integrity helped me be the researcher I am today.

Last, but not least, my wife, Dara. She is my “secret weapon.” Always cheering me on, inspiring me to aim high. None of this would be possible without her unwavering emotional support. You make me a better person, and my life a lot more fun.

VITA

EDUCATION

2006	M.A., Demographic and Social Analysis, University of California Irvine, Irvine, CA.
1998	Social Studies Teaching Certificate, Millersville University, Millersville, PA.
1996	B.A., History, University of Pittsburgh, Pittsburgh, PA.

RELEVANT PROFESSIONAL EXPERIENCE

2017 – 2018	Project Coordinator, Graduate Student Researcher UCLA Partner School Network University of California, Los Angeles, CA
2015 – 2018	Graduate Statistics Teaching Assistant Graduate School of Education & Information Studies University of California, Los Angeles, CA
2015 – 2018	Senior Research Advisor Nexus Planning & Research Los Angeles, CA
2014 – 2018	Project Manager, Graduate Student Researcher CDE Early Childhood Education Program Evaluations University of California, Los Angeles, CA
2014 – 2017	Independent Contractor, Evaluation and Assessment San Jose Unified School District Los Angeles, CA
2012 – 2014	Program Manager, Research and Evaluation San Jose Unified School District San Jose, CA
2003 – 2012	Senior/Associate Research Analyst San Diego Association of Governments San Diego, CA
1997 – 2000	Senior Demographic Analyst/Legislative Research Analyst Pennsylvania House of Representatives Harrisburg, PA

PUBLICATIONS

2017	Rohanna, K. Breaking the adopt, attack, abandon cycle: A case for improvement science in K-12 education. New Directions for Evaluation, 153. 65-77.
2016	Vo, A.T., Christie, C.A., & Rohanna, K. Understanding evaluation practice within the context of social investment. Evaluation, 1-19.

2007	Rohanna, K., & Tayman, J. Census data for transportation planning, analysis, and implementation. Journal of Social and Economic Measurement, 31, 167-183.

PRESENTATIONS

2018	Rohanna, K. Engaging educators in productive conversations that generate deeper learning and foster change. Presented at the Harvard Strategic Data Project Annual Convening, Boston, MA.
2016	Rohanna, K. & McGinnis. More than outcomes: Using process, implementation data for better information. Presented at the Harvard Strategic Data Project Annual Convening, Boston, MA.
2015	Rohanna, K. Meaningful evaluation in the context of schools: Rapid cycles of evaluation for school improvement. Paper presented at American Evaluation Association Annual Conference, Chicago, IL.
2015	Rohanna, K. Improving education through improvement science. Paper presented at American Evaluation Association Annual Conference, Chicago, IL.
2010	Rohanna, K. Invited panelist. Should the American Community Survey (ACS) continue to collect group quarters data? National Academy of Sciences, Washington D.C.
2007	Rohanna, K. Invited speaker. Demographic indicators and trends: Income, race, poverty. ACS Capacity Building Peer Exchange. Daytona Beach, FL.
2006	Rohanna, K. & Tayman, J. The Census at the dawn of the 21st Century: Census data for transportation planners. Paper presented at Southern Demographic Association Meeting, Durham, NC.

AWARDS, HONORS, AND APPOINTMENTS

2017	Distinguished Teaching Assistant Award, University of California, Los Angeles
2012 – 2014	Harvard Strategic Data Project Fellowship, Harvard University Center for Education Policy Research
2012	Elected Member, U.S. Census Bureau, Census State Data Center Steering Committee
2011	Appointed Member, National Academy of Sciences Transportation Research Board (TRB) Planning Committee for Using Census Data for Transportation Applications
2011 – 2012	Appointed Member, Census Data for Transportation Planning Products Oversight Committee
2009 – 2012	Co-chair, National Academy of Sciences TRB Census Data for Transportation Planning Sub-committee
2008 – 2012	Appointed Member, National Academy of Sciences TRB Urban Data and Information Systems Committee
1996	Phi Beta Kappa, University of Pittsburgh Chapter

CHAPTER 1

INTRODUCTION

"The only real voyage of discovery consists not in seeking new landscapes, but in having new eyes; in seeing the universe through the eyes of another, one hundred others - in seeing the hundred universes that each of them sees."
-Marcel Proust

My foray into education reform began in 2012 when I landed at a northern California school district through a Gates Foundation fellowship. I was tasked with being an "agent of change" around data use and research. It was during the education reform zeitgeist: Many private former businessmen-turned-philanthropists felt they had the knowledge to fix education's most pressing problems. Documentaries like *Waiting for Superman* (Guggenheim & Chilcott, 2011) and a *Time* magazine cover photo featuring Michelle Rhee and a broom (Ripley, 2008) had sent the message that education needed housekeeping or a hero to save it from the people in it.

But education reform is nothing new. In the mid-1960s, the federal government first attempted to address educational equality with a constellation of "War on Poverty" programs (e.g., the Elementary and Secondary Education Act of 1965). Yet, unequal access to education resources and opportunities in the United States' education system remain, and educators and researchers are still tackling the challenges faced by those 50 years ago.

Importantly, no single group possesses the unique knowledge to fix the system. Perhaps the most striking example of this was the unsuccessful effort to transform the failing Newark, New Jersey public school system. Fueled by a one hundred million dollar pledge from Facebook founder, Mark Zuckerberg, misguided top-down efforts valued the ideas and opinions of outsiders, while dismissing knowledge from the people they were there to serve (Russakoff, 2014). Four years after the reform efforts began, most of the money was spent and the outsiders had moved on, leaving behind many angry parents and teachers, and a fleeing superintendent. *Why did it fail?* Among other things, one could theorize that it did not honor the expertise and experiences of those on the frontline – Proust's "one hundred eyes."

The example of Newark demonstrates the importance of embracing a variety of knowledge; knowledge and experiences from a diverse set of individuals at all levels within an organization, along with expertise from outside the organization. In recent years, networked improvement communities have aspired to provide that diversity of knowledge, combining practical frontline teacher and principal experience with external content expertise and improvement know-how.

Networked Improvement Communities

In the context of education, networked improvement communities are *inter*-organizational networks, whereby multiple schools join forces to tackle complex challenges. Schools unite around improving one common problem of practice, such as low early literacy rates. They jointly dig deeper into the problem by sharing a diversity of experiences, and then develop a plan for how to improve it. Teachers and administrators provide the practitioner experience and content knowledge, while specialists, typically from universities or other research institutions, share research-related content (e.g. literacy expertise) and provide a framework for systematically examining potential solutions and evaluating short-term progress.

Networked improvement communities were a concept first envisioned by Douglas Engelbart, an engineer and technological innovator who coined the term "Collective IQ," and envisioned uniting a diversity of knowledge to solve the world's most complex problems (Engelbart, 2004). Engelbart believed that by leveraging a plethora of experiences, information, and abilities, people could collectively generate new insight into demanding problems. His ABC model for organizational improvement advanced how to develop collective IQ by generating and disseminating knowledge at multiple intra- and inter-organizational levels, whereby higher levels aim to enhance the improvement capability of lower levels (Engelbart, 1992).

In education, networked improvement communities seek to incorporate the notions underpinning Engelbart's ABC model of improvement (Bryk, Gomez, Grunow & LeMahieu, 2015). A network improvement community can be conceptualized as having multiple levels: The network level consisting of multiple schools (Level C), the individual school level (Level B), and the teacher/classroom level within each school (Level A). At the highest level, the network is committed to enhancing both the school's and the teacher's capability for improving complex problems of practice.

In practice, the network is typically coordinated by a hub, such as a university or other research institution, and is formed when the hub invites teachers and principals from multiple schools to unite around a persistent shared problem, such as math failure rates. Networked improvement communities generally utilize improvement science as its continuous improvement or problem-solving framework. As such, the hub team usually includes a subject matter expert(s) (e.g. math instruction), and an improvement science specialists(s).

The network hub's initial role is to lead the network in defining the shared problem of practice, and in more deeply analyzing the problem through root cause and causal systems analyses. Once these are established, the network hub directs the development of a theory of change framework for improving practices, which guides the work, also known as a *theory of practice improvement or driver diagram*. Individual schools test this theory in practice, which typically involves multiple teachers experimenting with potential solutions in their classrooms. Through this process, educators within one school ideally form their own *intra*-organizational improvement community, where they share, learn, and generate new knowledge on a smaller scale than the network. These within-school learnings are then shared with all the other schools during periodic network gatherings. Multiple, varied new insights or knowledge from each school are integrated to continually inform the theory of practice improvement. A key characteristic of networked improvement communities is that they embrace the variation in school experiences; thus it is crucial to implement in a diversity of school contexts (Bryk, Gomez, Grunow & LeMahieu, 2015). The network's role at this stage is to generate new knowledge and formalize both the collective content and improvement process learnings (Bryk, Gomez, Grunow & LeMahieu, 2015). By participating in the network, the within-school improvement communities not only potentially solve the problem of practice but they also "get better at getting better" (Bryk, Gomez, Grunow & LeMahieu, 2015).

Statement of the Problem

Networked improvement communities provide a promising approach for solving education's most pressing problems today. However, the promising feature that embraces practitioner experience and knowledge can also create a challenge when considering the capacity needs of those engaging in the improvement activities. Networks who apply an improvement science framework require teachers and administrators to learn and apply technical methods related to causal systems analysis, theories of change, disciplined inquiry, and data collection and analysis. Learning how to engage in these technical activities can place an extra burden on teachers and administrators, which may be daunting and/or unrealistic given their existing duties. However, the network's ability to achieve success and solve complex problems of practice likely hinges on its ability to build improvement science capacity (or another continuous improvement process) because knowledge is primarily generated through these improvement science tools. Without this crucial capacity, teachers and administrators may become frustrated when participating in these time-intensive network activities, with little improvement shown for their effort.

Additionally, networked improvement communities by design, seek differing school contexts to encompass a diversity of knowledge and experiences. This might inadvertently lead to inconsistent implementation of the improvement activities. Previous research has shown that school reform undertakings and capacity building can be influenced by organizational factors, such as school leadership and culture (Bryk, Camburn, & Louis, 1999; Louis, 2006; Louis & Lee, 2016; Preskill & Boyle, 2008). These organizational differences could either help or hinder a school's ability to implement improvement science and fully participate in the networked improvement community. While this idea seems obvious, it is also an understudied area in related research. Existing literature focuses on the role of the network hub with minimal discussion of the complexities surrounding the implementation of improvement science by distinct schools within a network.

Even under the best circumstances, implementing reform activities is messy and can have negative consequences (Rose, 2016). This messiness can be due to a myriad of reasons: poor school leadership, lack of structures, or even misunderstandings about the level and type of capacity needed to implement an initiative. Teachers and administrators may become frustrated because they do not have a clear understanding of why the gap between idealism and realism arises, often leading to the "adopt, attack, abandon" phenomenon where educators frequently give up on an initiative, instead of addressing the complex operational challenges (Rohanna, 2017). Thus, if networked improvement communities are to fulfill their promise, there needs to be more empirically-based research explaining how to successfully prepare schools to participate in networks, including how to build improvement science capacity and foster its implementation by schools with diverse organizational contexts.

This present study addresses these issues through an explanatory single case study of one networked improvement community with five diverse schools. The network was coordinated by a university hub, and utilized improvement science as its continuous improvement framework. The use of a case study for this research was most appropriate because it was an empirical study of a phenomenon within its real-world context where the boundary between the object of study (e.g. improvement implementation) and context could be blurred (Yin, 2014).

Research Purpose and Questions

The primary purposes of this research were to: 1) examine how schools were prepared to successfully participate in the network, and 2) explain what facilitated successful participation. Successful implementation of improvement science was defined as a school's ability to execute Plan-Do-Study-Act (PDSA) cycles within Year 1 of the network. (PDSA cycles are defined in Chapter 2). The study was guided by the following research questions.

1. How were schools prepared to successfully participate in the University networked improvement community?
    a. How did the hub (University) teach the schools improvement science?
    b. How did schools learn improvement science?
    c. How were the fundamental network components developed, i.e. theory of practice improvement, change ideas, measures?
2. What facilitated a school being prepared to successfully participate in a networked improvement community?
    a. What capacities, structures, and/or conditions were needed for schools to successfully implement improvement science?
    b. What capacities, structures, and/or conditions were needed for schools to generate learnings that contributed to the learnings of the network?

Positionality

At this point, it is important to be forthcoming about my positionality in this study. In addition to researching this network, I also served as graduate student researcher (GSR) on the hub team who participated in the planning and coordination of the network. My role was further elevated mid-year (December 2018) and I became responsible for leading the network activities, both at the network convenings and at individual school meetings held between the convenings. My positionality posed benefits and challenges while conducting this study. While it provided me access to all network-related meetings and decisions, enabling me to write a detailed narrative, my positionality also impacted my research. That is, my research lens was that of an insider, rather than outsider. This insider perspective, combined with the methodological limitations of my primarily qualitative data, subjected my research to potential biases. As such, I took steps to address this issue. I attended to my own reflexivity throughout the study by writing in a journal and engaging in frequent conversations with my advisor regarding how my research was influencing my network role, and vice versa. As described in more detail in the Chapter 3 (Research Methods), I collected and triangulated data from multiple sources, systematically considered alternative hypotheses and disconfirming evidence, and engaged in peer debriefing and member checks with other hub staff. I further validated my findings by comparing to other similar research (Proger, Bhatt, Cirks, and Gurke, 2017).

Study Significance and Implications

This study has the potential to benefit both educators and researchers, because it is an empirically-based case study about preparing schools to implement improvement science and successfully participate in a networked improvement community. Existing literature provides little direction regarding the specific processes for establishing a new networked improvement community, including how to build the improvement science capacity of diverse member schools. This study aimed to fill that void by providing a detailed account of this topic, including the theorization of a new improvement science capacity building model in the hopes that it could serve as a roadmap for others.

Furthermore, the ability of networked improvement communities to solve complex problems is dependent upon the learnings being generated both at the school level, and the network level (Engelbart, 1992; Bryk, Gomez, Grunow & LeMahieu, 2015). In my experience, the push to have teachers and administrators complete the procedural aspects of improvement science, particularly PDSA cycles, can overshadow the process's true purpose: generate actionable learnings. Consequently, if we expect schools to apply the improvement science framework, we would benefit from a deeper examination of how and what types of knowledge are generated through this process, including what crucial elements should be in place to foster meaningful learnings. This study aims to provide that examination.

Finally, this subject is timely, as more attention is being directed to the promise of networks to improve complex, persistent education problems of practice. For example, in January 2018, the Bill & Melinda Gates Foundation announced their initiative to fund Networks for School Improvement (NSI). Analogous to a networked improvement community (also known as a NIC), an NSI is defined as a group of schools that collectively and individually work through a continuous improvement process to improve outcomes for Black, Latino, and low-income students (Bill & Melinda Gates Foundation, 2018). It is centered around a coordinating organization that "brings together multiple school leadership teams to tackle common problems and work towards common aims." (Bill & Melinda Gates Foundation, 2018).

With attention and funding being directed toward school improvement networks, it is crucial to examine these processes more closely, and acknowledge the gap between the ideal and reality in practice. School reform efforts sometimes operate under the assumption that technical methods can overcome all; a sort of blind faith that once the structural pieces are in place, educators will somehow successfully sort through the implementation challenges. We have seen this presumptive idea in the recent data-based decision-making movement: "If you measure it, change will come." In my experience, the people and institutions leading this effort paid little attention to internal capacity needs and the school context. This study aims to provide a thorough investigation of this timely topic.

Manuscript Organization

This manuscript consists of eight chapters. The next chapter, Chapter 2, is a review of the relevant literature. It includes the conceptual framework, as well as a more thorough discussion about this study's contribution to existing literature. Chapter 3 provides information about this study's design, data collection procedures, and analytical methods. Chapter 4 provides a detailed narrative of the case of the network, told from the first-person perspective in an effort to be explicitly transparent about my dual roles as researcher and hub staff. Research question one is primarily analyzed through this narrative. Chapters 5, 6, and 7 are organized by themes, and provide the findings to the second research question by building upon the narrative in Chapter 4. These chapters further identify capacities, structures, and/or conditions that facilitate schools successfully implementing improvement science and generating meaningful learnings. In the last chapter, Chapter 8, I conclude with a review of the findings and a discussion of this study's significance and implications, including a theorized model of how to build a network's improvement science capacity in Year 1. This chapter also includes a discussion of this study's limitations and possible directions for future research.

CHAPTER 2

REVIEW OF RELEVANT LITERATURE

Introduction

As past reform efforts have shown, it is difficult to successfully implement improvement efforts in schools due to a myriad of reasons. This study aims to better understand how to prepare schools to successfully participate in networked improvement communities. In particular, it focuses on one improvement model, improvement science, and how it is implemented by five schools within one networked improvement community. These five schools provide the diversity necessary for the networked improvement community design. However, as already discussed, this design that seeks variation in school experiences may contribute to challenges of building improvement science capacity due to differing needs and organizational capacity.

Another, yet paradoxical challenge, and strength, of implementing improvement science in networked improvement communities is that it is a *participatory*-oriented evaluative framework (Cousins & Earl, 1992; Christie, Inkelas, & Lemire, 2017). Improvement science can be defined as:

> A data-driven change process that aims to systematically design, test, implement, and scale change towards systemic improvement, as informed and defined by the experience and knowledge of subject matter experts (Lemire, Christie, & Inkelas, 2017, p. 25).

The participatory nature of the framework means that the frontline workers who have the experience and practitioner subject knowledge, in this case, teachers and administrators, are the ones who actually implement the improvement activities. The previous Newark example illustrates why it is important to closely involve school staff in reform and improvement efforts. Yet this strength also becomes a challenge when the frontline staff are charged with leading this evaluative work at their school sites. In order for the networks to be successful, school teams need to build improvement science capacity. Learning improvement science can be difficult because it is very technical with distinctive terms and tools.

Furthermore, improvement science is considered an organizational learning model because of its systemic approach that embraces a continuous learning aspect that is ideally integrated into the organization's work activities, culture, and values (Torres & Preskill, 2002). As such, any discussion of networked improvement communities should also consider what type of learnings are generated and what facilitates that learning.

The connections among three areas – evaluation, organizational learning, and education – frames this literature review. In evaluation and organizational learning, I looked to the improvement science literature, grounded in Deming's "system of profound knowledge," and considered other relevant organizational learning theories. In education, I reviewed literature related to networked improvement communities, school improvement, organizational factors, and organizational learning. My analysis is not limited to only peer-reviewed or scholarly journals. I also included books and non-scholarly articles because practitioner-based writings can be found in non-scholarly sources. Organizational learning theorists also tend to write books that either provide more detail, or are the primary source of their writings.

Improvement Science and Networked Improvement Communities

Improvement science is founded upon much of W. Edward Deming's work (Langley et al., 2009). Deming was an engineer and statistician who advanced production and management effectiveness and quality improvement. His ideas shaped Japanese manufacturing and industrial practices after World War II (Walton, 1986). Deming (1994) presented the fundamental idea of a system of profound knowledge in his book *The New Economics for*

Industry, Government, and Education (Deming, cited in Langley, et al., 2009, p.75). Deming is also credited with the PDSA cycle, which was an evolution of Walter A. Shewhart's initial cycle of scientific testing (Moen & Norman, 2010). The PDSA cycle is format for rapidly experimenting with new practices. It has four stages: plan, do, study, and act. The experiment logistics are planned during the first stage (plan), and implemented during the second stage (do). During the next stage (study), the experimenter analyzes relevant data, reflects upon the process, and determine the next steps. These next steps are put into action in the final stage (act). Ideally, this cycle should occur within a short timeframe and on a small scale so that ideas can be quickly tested, and either adapted and tested again, gradually scaled up, or potentially abandoned if necessary.

Langley and his colleagues at Associates in Process Improvement expanded upon Deming's work and developed the "Model for Improvement" (Langley et al., 2009). The Model for Improvement encompasses three questions and the PDSA Cycle (Langley et al., 2009). The three questions are:

1. What are we trying to accomplish?
2. How will we know the change is an improvement?
3. What change can we make that will result in an improvement?

While improvement science is grounded in industry and manufacturing, a review of the literature today reveals that improvement science is more prevalent in healthcare than any other field. Don Berwick, the founder of the Institute for Healthcare Improvement (IHI), was one of the early champions of improvement science in healthcare. The IHI promotes the use of improvement science methods to improve healthcare, worldwide. Improvement science has successfully cracked taxing healthcare challenges, such as how to reduce the number of child asthma-related visits to the emergency room (C. E. Williams, 2015) and improve physician compliance with hand hygiene (White, et al., 2012).

Improvement science has expanded to education in recent years. While improvement networks can apply any continuous learning framework, the Carnegie Foundation has advanced the use of improvement science in networked improvement communities. They addressed the widespread problem of developmental math students not successfully passing a college-level mathematics course within three years by using improvement science and networked improvement communities (Sowers & Yamada, 2015). By reengineering the approach to developmental and college-level mathematics courses, participating community colleges were able to increase the percentage of students passing college-level math courses (Sowers & Yamada, 2015). K-12 schools also are embracing improvement science, which is demonstrated by the continually growing attendance at the Carnegie Foundation's annual summit and blogs from their website.

However, there is still more to be learned about how to successfully apply networked improvement communities and improvement science in K-12 education. The fairly recent publication of *Learning to Improve: How America's Schools Can Get Better at Getting Better* by Bryk, Gomez, Grunow, and LeMahieu (2015) has provided guidance, yet implementing a theoretical model in a school requires empirical studies to understand real-world complexities. As of this writing, I only found one published empirical study regarding the complexity of implementing improvement science in K-12 schools, and it limited its focus to PDSAs (Rohanna, 2017). Yet, it lacks detail regarding the processes for successfully implementing improvement science in the school context.

Two more recent publications have enhanced the conversation about initiating networked improvement communities. *A Framework for the Initiation of Networked Improvement Communities* offers more practical guidance but is still too broad to serve as a how-to guide for others hoping to launch networks (Russell, Bryk, Dolle, Gomez, LeMahieu & Grunow, 2017). Furthermore, it does not detail processes for building improvement science capacity. In their report *Establishing and Sustaining Networked Improvement Communities: Lessons from Michigan and Minnesota*, Proger, Bhatt, Cirks, and Gurke (2017) provided more information about processes; however, they provided guidance regarding *what* processes should occur rather than *how* to enact. They do not provide information regarding how to build the improvement science capacity of their educators either. These literature also tended to focus more on the hub's role than the experience of the teachers and school participating in the network.

Evaluation Capacity Building

Because improvement science is a participatory evaluation framework, we can look to evaluation capacity building literature for guidance. Though there are several similar definitions, evaluation capacity building can be succinctly defined as "an intentional process to increase individual motivation, knowledge, and skills, and to enhance a group or organization's ability to conduct or use evaluation" (Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012, p. 308). The topic of capacity building emerged among evaluators in the early 2000s as they became more intent on involving stakeholders, and building stakeholder understanding and internal evaluation practices within organizations (Preskill & Boyle, 2008; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012).

Evaluation capacity building was characterized as distinct from other evaluation practices. While evaluation practice could be crudely classified as conducting a program evaluation using an acceptable model (e.g., participatory evaluation), evaluation capacity building was (and is) an intentional and ongoing practice to develop and sustain regular evaluation use within and/or between organizations (Compton, Baizerman, & Stockdill, 2002).

Preskill and Boyle (2008) added to the evaluation capacity building literature when they conceptualized a multidisciplinary model to guide practitioners. Their model drew upon the fields of evaluation, organizational learning and change, and adult and workplace learning theories. First, the model designates the *goal* of evaluation capacity building as the development of evaluation knowledge, skills, and attitudes. Then, it further distinguishes that those who initiate evaluation capacity building activities have various motivations, assumptions, and expectations regarding what they hope to achieve. Depending on these and the intended objective, there are 10 different evaluation teaching and learning strategies that could be employed: internship, written materials, technology, meetings, appreciative inquiry, communities of practice, training, involvement in evaluation, technical assistance, and coaching.

According to their model, the learning needs to be transferred to the work context in order for this individual capacity to be sustained. Their model further deconstructs the processes, practices, policies, and resources required for sustainable evaluation practice. Examples of these include the use of evaluation findings, integrated knowledge management evaluation system, and continuous learning about evaluation, and so forth.

In the model, organizational learning capacity envelops teaching and learning strategies and sustainable evaluation practices. Preskill and Boyle (2008) posit that four areas of organizational learning capacity will influence the extent to which individuals will learn and build evaluation capacity, and the extent to which it will be sustained. These four areas are: leadership, culture, systems & structures, and communication.

The evaluation capacity building model developed by Preskill and Boyle (2008) provided a framework for conceptualizing evaluation capacity building efforts that influenced others. Labin and her colleagues (2012) continued to build upon the evaluation capacity building theory by reviewing and synthesizing evaluation capacity building empirical literature. They structured their research synthesis around their own model, the Integrative ECB Model, which was primarily based on Preskill and Boyle's model (2008). As such, it also identified the individual capacity building goals as attitudes, knowledge, and skills/behaviors; and it categorized organizational capacity outcomes as processes, policies, and practices, leadership, organizational culture, mainstreaming, and resources.

Organizational Factors in Education

To better understand what organizational capacity factors could hinder or foster evaluation capacity building efforts in network schools, I turned to education literature related to organizational factors, organizational learning, and school improvement. These literature are framed around professional learning communities (PLCs). Since there is little literature about improvement science in education and networked improvement communities, I turned to literature about PLCs and their network counterpart, networked *learning* communities (NLCs). NLCs are networks of school PLCs. According to Stoll and colleagues (2006), there is no universal definition of PLC but there is a "broad international consensus that it suggests a group of people sharing and critically interrogating their practice in an ongoing, reflective, collaborative, inclusive, learning-oriented, and growth-promoting way" (p.224). Thus, the PLC is similar to the within improvement community identified at Level B of Engelbart's ABC model for improvement. Because improvement communities are a fairly new occurrence in education, there is no literature that differentiates the two types of communities at the school level; a somewhat confusing point in practice. However, the distinction between the two is their emphasis – one is on learning, the other is on improving.

This distinction has been made at the network level by experts in both areas, and can logically be applied to the school community level. Katz and Earl (2010), NLC scholars, acknowledge that networks exist for many purposes, and "in networked learning communities, the emphasis is on learning" (p.28). Gomez, and his colleagues, who introduced networked improvement communities to education, further breakdown the distinction by defining two types of collective action networks in education: sharing networks and execution networks (Gomez, Russell, Bryk, LeMahieu, & Mejia, 2016). In sharing networks, like NLCs, Gomez et al. state that the purpose is to share information that supports individual school action and agency. Each school may be working towards solving their own *unique* problem of practice, and learning from others addressing a similar issue. By contrast, they assert that the execution network members explicitly agree to improve one *common* aim or problem of practice, typically with a collective target. Networked improvement communities are execution networks. They develop a common theory of how to improve the problem and use improvement methods (improvement science) to test and refine strategies and interventions. Networked improvement communities, and subsequently, in-school improvement communities, should have specific structures in place in order to collectively solve a complex problem, whereas NLCs and PLCs may also be working towards solving similar problems with their group but with less structure regarding specific theories of how to improve and improvement methods.

Organizational factors are considered those environmental characteristics or conditions that potentially influence workers' capabilities, and they can be conceptualized in numerous ways in education. Bryk, Camburn, and Louis (1999) researched structural, human, and social organizational conditions that facilitated school PLCs. They found that trust among faculty, facilitative principal leadership, principal supervision (as an indicator of regular principal involvement), and small school size were related to the presence of professional learning communities in schools. Other studies have considered similar organizational factors in schools, such as principal leadership style, school relationships, and supportive structures (Stoll & Louis, 2007; Katz & Earl, 2010; Louis & Lee, 2016).

Louis (2006) argued that three school cultural components – professional communities, organizational learning, and trust – improve student learning. Louis continued to study the interplay of culture, organizational learning, and school improvement, and broke out organizational learning from school culture, which she and Lee (2016) considered to be a possible mediator between a school's culture and its ability to improve outcomes. Their study found that the cultural factors – academic press, academic support for students, and trust and respect among colleagues – predicted a school's capacity for organizational learning. Furthermore, they included dimensions of professional learning in their analysis and showed that shared responsibility, reflective dialogue, and deprivatized practice (teachers observing and giving feedback to one another) among teachers also were related to organizational learning.

Other contextual factors, such as the percent of low income students, tend to be included in discussions of school organizational factors as well. Some factors, such as school size and school level (i.e. elementary, middle, and high) have been considered both contextual factors and structural, (Louis 2006; Stoll, 2006).

While the aforementioned school organizational factors are associated with school improvement efforts, it is important to more specifically investigate which factors are related to *implementation* of improvement initiatives. Implementation is the precursor to outcome changes. However, I found no studies in my review of the literature. Although, Katz and Earl (2010) attempted to tease out implementation as part of their study of influential key enablers (organizational factors) in learning networks. They asked teachers to rate their colleagues' changes in thinking and practice as an intermediate outcome to the ultimate student outcome. Because they relied on survey data, their change in practice measure was broad and was subject to biased perceptions from colleagues. Thus, it was not an objective measure of implementation.

Instead, these prior empirical studies tended to connect organizational factors to the capacity for organizational learning, instead of actual implementation (Bryk, Camburn, and Louis, 1999; Louis & Lee, 2016). One reason for incorporating organizational learning as an outcome is that the capacity for organizational learning suggests the presence of organizational processes that are expected to be a mediator between organizational factors and improvement in-school outcomes (Louis & Lee, 2016). Like implementation, organizational learning capacity is a precondition for school improvement and continuous inquiry (Louis & Lee, 2016).

Another reason is more practical than theoretical. These studies investigated relationships between organizational factors and capacity for organizational learning by utilizing multilevel models to analyze survey and other data from multiple schools (Bryk, Camburn, & Louis, 1999; Louis & Lee, 2016). Predictive models with implementation as the outcome would be problematic if each school were employing individual improvement initiatives. Differences in the outcome could be due to differences in the improvement activities. While studies could analyze student outcomes as the dependent variable, it does not necessarily represent a change in practice. If the goal is to better understand how organizational factors influence implementation, (as the necessary intermediate step between initiating an endeavor to solve a problem of practice and determining whether the desired impact occurred), then the actual changes in teacher and administrator practices needs to be assessed. This is an understandable gap in the literature. Unlike the research proposed here, very few studies, if any at all, have the opportunity to research capacities, structures, and/or conditions that facilitate the successful implementation of improvement efforts while holding the improvement activities constant.

Knowledge Generation and Organizational Learning

Knowledge acquisition and generation are essential ingredients for improvement activities (Engelbart, 1992; Langley et. al., 2009). Networked improvement communities are founded upon the idea of generating collective learnings to solve complex problems of practice (Engelbart, 2004). Furthermore, Engelbart's ABC model for organizational improvement necessitates that learnings are generated at multiple levels (Engelbart, 1992). Thus, in networked improvement communities, one would expect learnings to be generated at the teacher/classroom level, school level, and the network level. Thereby further indicating that networked improvement communities are organizational learning entities.

Existing networked improvement community literature does not specify the type of organizational learning to be generated from network improvement activities. Rather, in my experience, networked improvement communities rely upon improvement science classifications when considering learnings. However, even though improvement science does provide descriptions of knowledge generation, it is limited when trying to conceptualize the type of learning necessary for solving complex problems of practice.

Improvement science delineates two types of knowledge: subject knowledge and profound knowledge. Subject knowledge is considered the content knowledge within a particular area, often held by practitioners and/or researchers, while profound knowledge is the more systematic awareness of “how to make changes that will result in improvement in a variety of settings” (Langley et al., 2009, p.75). Deming defined profound knowledge “as the interplay of theories of systems, variation, knowledge, and psychology” (Deming, cited in Langley et al., 2009, p. 75).

Deming structured his system of profound knowledge around four types of knowledge (Christie, Inkelas, & Lemire, 2017):

1. Knowledge of systems: This type of knowledge refers to the interdependence of departments, people, and processes within an organization (Langley et al., 2009). Integration of these individual parts toward a common aim contributes to a successful organization (Deming, 1994; Langley et al., 2009).
2. Knowledge of variation: This component not only promotes the shift from analyzing averages to a deeper study of variation in data, but it also encourages an understanding of different types of variation and their implications for system performance (Christie, Inkelas, & Lemire, 2017; Langley et al., 2009).
3. Knowledge of how knowledge grows: This type of knowledge refers to learning by making predictions about potential changes, then actually making the changes and

measuring the results. The PDSA cycle is an example of how knowledge grows (Langley et al., 2009).

4. Knowledge of psychology: This component reflects the human side of change, and encompasses how attention to people's values, attitudes, and motivations can influence change (Langley et al., 2009).

Within Deming's third type of knowledge – knowledge of how knowledge grows – is where the knowledge generation is expected to occur. The PDSA cycle is the primary source of this knowledge generation. Langley and colleagues (2009) describe how learning is generated through the cycle:

> Deductive and inductive learning are built into Plan-Do-Study-Act (PDSA) Cycles. From Plan to Do is the *deductive* approach. A theory is tested with the aid of a prediction. In the Do phase, observations are made and departures from the prediction are noted. From Do to Study the *inductive* learning process takes place. Gaps to the prediction are studied and the theory is updated accordingly. Action is then taken on the new learning. (p. 82)

Other than the PDSA cycle, improvement science does not further conceptualize the type of learning needed for improvement, thereby leaving a void when more precisely considering what type of learning should be transpiring in improvement communities. This may be due to the fact that improvement science was born out of manufacturing and extended into health care. The original intent behind PDSA cycles was to experiment with very small changes, which are often very process focused (Langley et al., 2009). Therefore, the expected learning is often precisely specified in the cycle, and may not require additional consideration. Network improvement communities, however, comprise different levels of learning (teacher/classroom, schools, networks) and learning generation not only from the PDSA cycle, but also from the collective intelligence of the diverse schools (Engelbart, 1992; Engelbart, 2004; Bryk, Gomez, Grunow, & LeMahieu, 2015).

Because networked improvement communities are organizational learning entities, we can look to organizational learning theory to conceptualize the type of learning required to solve complex problems of practice. One of the most well known organizational learning theories is Peter Senge's learning organization. It considers five dimensions of organizational learning: systems thinking, personal mastery, mental modes, building shared vision, and team learning (Senge, 2006). However, it does not focus on processes for knowledge generation.

Argyris and Schön (1996), on the other hand, advance the concepts of single-loop and double-loop learning that are relevant to knowledge generation for both individuals and organizations. Single-loop learnings are instrumental learnings that lead to improved performance without changing underlying values, norms, or strategies regarding current practices (Argyris & Schön, 1996). Double-loop learnings question underlying values, norms, or strategies, ultimately leading to changes in how and/or why certain practices are being done (Argyris & Schön, 1996). Senge (1990) incorporates Argyris and Schön's ideas in his team learning and mental modes dimensions, likely because they more precisely postulate how learning can occur.

Double loop-learning transpires when the inquiry results in changes to individual or organizational values of "theories-in-use" (Argyris & Schön, 1996). Theories-in-use are the patterns that are implicit in individual or organizational behaviors, and can be compared to "espoused theory" (Argyris & Schön, 1996). Espoused theory represents the strategies and values that individuals or organizations communicate to explain their actions (Argyris & Schön, 1996). Fundamentally, espoused theory represents the notion of "what we say we do" which can be compared to theories-in-use that signifies "what we actually do" (Senge, 2006).

The notions of single-loop and double-loop learning complement networked improvement communities because PDSA cycles and other network activities could lead to either single-loop or double-loop learnings. Both types of learning would be beneficial. However, single-loop learning may not be enough due to the specific purpose of networked improvement communities. Networks are intended to solve complex problems of practice, which may require educators to change their value, norms, and/or strategies regarding their current practices, thus, necessitating a need for double-loop learnings. Therefore, Argyris and Schön's (1996) organizational learning theory provides a useful model for conceptualizing learning in networked improvement communities.

The Current Study

From the literature, I created a conceptual framework to guide this research, shown in Figure 1. Most notably, this framework supported my development of data collection instruments and protocols by helping me operationalize what "preparedness" means in this context.

Preparing schools to successfully participate in a networked improvement community requires network hubs to develop the improvement science capacity of its members. At the individual level, the network hub should consider capacity related to knowledge, skills/behaviors, and attitudes. The framework posits how these concepts can be conceptualized in the networked improvement community context (column 3). At the organizational level, schools need to embed processes and structures, and the network hub should attend to the culture and leadership support. The framework further suggests that building teacher and school capacity in these areas will contribute to the completion of PDSA cycles, which will then lead to network learnings. Again, based on the literature, these learnings are conceptualized as single-loop and double-loop learnings.

Preparing schools to successfully participate	Evaluation capacity-building literature	Improvement science conceptualization in a networked improvement community
Individual capacity	Knowledge, skills/behaviors, attitudes	Knowledge: Participants understand terms and tools, causal analysis, driver diagram, PDSA process. Skills/behaviors: Participants apply and demonstrate use of tools and methods with integrity. Attitudes: Motivated to improve practices, values continuous learning and the process, believes worthwhile.
Organizational capacity	Systems, processes, structures to embed evaluation activities. Leadership and culture values inquiry, has necessary resources, structures, and policies for engaging inquiry.	School teams (teachers) build structures to effectively collaborate, conduct PDSA cycles, and engage in inquiry. Culture: trust, shared responsibility. Leadership sets up structures and provides resources for teachers to participate in improvement and network activities. Leadership is facilitative.

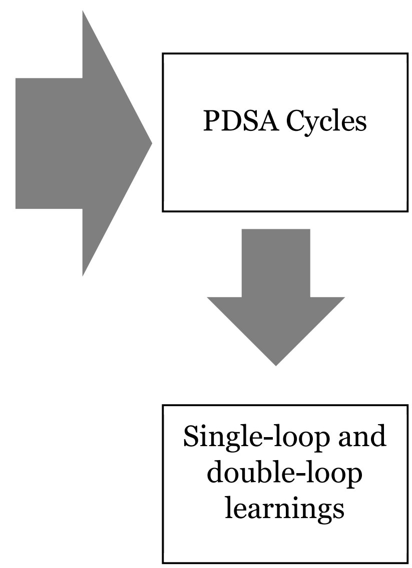

Figure 1. Conceptual framework

In sum, the present study has the potential to add to the scholarship regarding networked improvement communities. While existing network literature tends to be practical, it still does not provide a detailed empirical account of how to establish a network. Current literature provides guidance on what broad steps should be taken, but it is limited in information regarding how to undertake those steps. The present study provides a more in-depth account of starting a network through narrative process tracing, and examines what facilitates successful participation.

Furthermore, this study contributes to existing network literature by integrating evaluation capacity building theory, which fills the void left by improvement science literature regarding how to build the necessary improvement science capacity. By incorporating Preskill and Boyle's multidisciplinary model into the conceptual framework, this research contributes to both network and improvement science literature by positing a model for building improvement science capacity.

Finally, this study contributes to the network literature by recognizing that networked improvement communities are organizational learning entities and should more explicitly consider the type of learning that should transpire. Because Engelbart's ABC model of organizational improvement is more encompassing than improvement science when it comes to knowledge generation, network literature should also more explicitly look to organizational learning theories. By incorporating Argyris & Schön's (1996) concepts of single-loop and double-loop learning, this research theorizes that different types of learning may be required to solve complex problems of practice in education.

CHAPTER 3

RESEARCH METHODS

Introduction

This chapter presents the research methods and analyses used to examine how schools were prepared to successfully participate in a networked improvement community. First, this chapter provides an overview of the study procedures, which is followed by information about the case study design and setting, participant selection and recruitment, and data collection. Lastly, the analytic procedures are described. The specific research questions for this study were:

1. How were schools prepared to successfully participate in the University networked improvement community?
    a. How did the hub (University) teach the schools improvement science?
    b. How did schools learn improvement science?
    c. How were the fundamental network components developed, i.e. theory of practice improvement, change ideas, measures?
2. What facilitated a school being prepared to successfully participate in a networked improvement community?
    a. What capacities, structures, and/or conditions were needed for schools to successfully implement improvement science?
    b. What capacities, structures, and/or conditions were needed for schools to generate learnings that contributed to the learnings of the network?

Overview of the Study Design

This study employed an explanatory, single case study approach with embedded school units (Yin, 2014). A case study was most appropriate for this study because its purpose was to examine the phenomenon of how to prepare schools to successfully participate in a networked improvement community within the real-life context of which it occurred (Yin, 2003). The primary outcome of a success was specified as the completion of PDSA cycles by schools, with individual teachers completing their own PDSA cycles. The networked improvement community, as the single case, was bounded by the activities occurring through the end of the network's first year. The embedded school units were also bounded by only including the people and activities related to the network. This multiple-method case study utilized detailed narrative process-tracing methods (George & Bennett, 2005), along with within-case, cross-unit (schools) analysis. These multiple modes of data and analytical techniques are described later in this chapter.

Study Setting

This case study examined one newly established networked improvement community founded by a large public university in Southern California, hereinafter called the University. The University was founded in 1919, and is a world-renowned research institution with a top-rated graduate school of education (U.S. News & World Report, 2018). The University's Graduate School of Education resides in a large, urban, and diverse city, and has a strong commitment to the study of urban education and social justice, and solving today's most pressing problems in education.

The University networked improvement community was born out of this commitment and housed in the Graduate School of Education. The network fell under the purview of the Associate Dean. The University staff working on the network, known as the hub team, initially consisted of four individuals, including the Associate Dean. One more person joined the team mid-year. The team included the Associate Dean as the project director, a math coach, an improvement science specialist (me), and two school liaisons. My role was elevated to network coordinator mid-year when I took a larger role in the planning and presentation of network content. The initial focus of the network was to improve student outcomes in Algebra I, and was later expanded to include all of middle school math. The University began planning the network in the fall of 2016, and it was launched at the beginning of the 2017 – 2018 school year. Seven network convenings were held throughout Year 1.

Five schools were recruited to participate in the network in the spring of 2017. These schools resided in one large urban school district in Southern California. There was one K-12 school, one high school, two middle schools (6th grade to 8th grade), and one junior high (6th grade to 9th grade). Within these schools, math teachers from each of the relevant grades participated in the network. Administrators were also invited, and participated to varying degrees. More details about the schools and their participating teams can be found in Chapter 4.

Case Study

Case Study Design

A single case study, and more specifically this particular network case, was deemed appropriate for two reasons. First, because of my positioning as both a researcher and a graduate student researcher (GSR) supporting the network, I had intimate access to the decision-making, planning, and processes of this particular network. This access provided the opportunity for a detailed examination of how to prepare schools to successfully participate in the network, including how to build improvement science capacity, which had not been undertaken before in an empirical study of networked improvement communities. Second, the network represented a *common* case. The network hub was a research institution, whose network was grant-funded, that recruited public schools to participate. Thus, it was foreseeable that this network was a usual case, and provided an opportunity to capture the conditions and circumstances of a typical network situation. Yin (2014) considers both of these rationale as acceptable reasons to employ a single case study design instead of a multiple-case design. It should also be noted there that although this particular case was not *selected* due to its successful outcome (i.e., I was unaware of the network's Year 1 outcome when I began collecting data), it did represent a successful case by the time I conducted my analyses, which influenced my analytical decisions.

This case study is further specified as a single case study with embedded school units (Yin, 2014), because this, and any networked improvement community, consists of distinct schools. Because the research questions seek to explain how schools can be prepared to successfully participate in a networked improvement community, this study's analysis occurs at multiple levels, i.e, the completion of PDSA cycles by individual schools and by the network as a whole. This design required data collected from hub staff, network convenings, school meetings, and teachers and administrators, along with network and school artifacts. The UCLA Institutional Review Board (UCLA IRB# 17-000847) approved this study design, the survey instruments, and the observation and interview protocols.

Participant Selection and Recruitment[1]

Participant selection for hub team. By the end of Year 1, the network hub team consisted of five people. Three of these individuals – the project director, the math coach, and the school liaison were recruited into the study because they represented the core team members who were present throughout the year. (I represented the improvement science specialist.) These individuals were recruited for interviews via email and in-person conversations.

Participant selection for schools. All five schools were included in the case study in order to obtain more potential outcome variation (# of PDSA cycles completed by school), as well as obtaining a wider range of capacities, structures, and conditions that could explain outcome variations. The general characteristics of the five network schools are shown in Table 1.

[1] The names of schools have been changed to protect the confidentiality of network participants.

(More detailed descriptions of the schools are provided in Chapter 4.). The grade level ranges for each school are not shown due to the increased likelihood of identifying schools based on this characteristic.

Table 1

Network School Characteristics, 2017-2018

School	Number of Network Teachers	Total Students	Demographics (%)					
			Low SES	Afr. Am.	Asian	Hisp./Lat.	White	Other
Roosevelt	6	1,004	91.8	1.7	9.7	80.8	1.6	6.2
Marshall	6	613	56.3	20.6	5.2	40.6	28.1	5.5
Middleview	4	394	77.9	51.5	0	46.7	0.3	1.5
Sawyer	4	660	75.3	13.0	3.5	69.2	10.9	3.4
Central	3	1,564	75.6	24.6	7.0	52.9	11.6	3.9

Note. Source is the California Department of Education. The number of teachers represents those participating in the network as of May 2018.

Participants from each school were recruited into the overall study at the first network convening, with the exception of teachers who joined the network mid-year and were approached individually later. The purpose of the study was explained to them, and all network participants were provided an informed consent form with more information. Participants could choose to participate in all or some of the research activities (e.g., observations, interviews). Teachers who did not consent to network or in-school meeting observations were not included in the analysis of data related to those meetings.

The interviews further employed purposive sampling. Eight teachers and administrators from three schools – Middleview, Sawyer, and Central – were interviewed as part of this study. These schools were selected because they included teachers *within* a school that had the most variation in the number of PDSA cycles completed by mid-April. The purpose of this sampling technique was to gather data related to school and individual capacity differences. Two teachers from each school were recruited via email. Of those, one teacher from each school represented either an individual who completed a "high" number of PDSA cycles or a "low" number, shown in Table 2. For Sawyer and Central, the network-participating administrator also was recruited via email and interviewed. The Middleview administrator was not interviewed because he did not participate in the network, and thus, would not have been able to answer the protocol questions. All of those who were recruited agreed to participate in an interview.

Table 2

Number of PDSA Cycles Completed by Interviewed Teachers

School	Number of PDSA Cycles Completed	
	Teacher: Low Number	Teacher: High Number
Middleview	0	4
Sawyer	2	5
Central	1	7

Data Collection

This multiple-method case study examined and triangulated data from surveys, interviews, observations, and document/artifacts. Due to my own positionality and the potential bias that could ensue, it was important to include a variety of data and methods to increase the credibility of my findings (Guba & Lincoln, 1986). Because this case study developed a process-tracing narrative (George & Bennett, 2005), it was also vital to collect an abundance of data from multiple sources to generate this detailed narrative. Furthermore, as Yin (2014) points out, incorporating multiple methods into case studies allows the researcher to collect a stronger array of evidence than could be obtained by any single method alone, thus, further strengthening the findings against my own potential biases. Table 3 first shows the data sources, with each category further described in this section.

Table 3

Data Sources for Case Study

Network Sources	
	Interviews: Network hub team (3 interviews)
	Observations of network convenings (6 observations)
	Surveys (2)
	Documents/Artifacts
	Convening agenda, presentation materials, feedback forms
	Causal analyses artifacts: Fishbone diagram, process maps
	Worksheets from meeting activities
	Network driver diagram
	GSR notes and calendar
School Sources	
	Interviews: Teachers and administrators (9 interviews)
	Participant observations of school network-related meetings (23 observations)
	Documents/Artifacts
	Fishbone diagrams, process maps, beginning driver diagrams
	PDSA forms
	Secondary data: Demographic, enrollment, and achievement data from the California Dept. of Education

Note. Three of the network convening observations were conducted by someone other than myself due to my role leading those meetings. GSR notes and calendar are my own.

Interviews. As shown in Table 3, network hub staff, teachers, and administrators were interviewed for this study. Semi-structured interview protocols were designed for each distinct group (Appendix A). The protocols were designed mid-year, and expanded upon information collected through the first survey and observations. It included items specific to the research questions while also providing enough flexibility to explore emergent ideas (Merriam, 2009). The first interview, in each group, also served as a pretest of the process and protocol. After the first interview, the protocol and procedures were revised as necessary.

Network hub. For the network hub team, the protocol primarily collected information regarding their perspective on the research questions, e.g., "To what extent, do you think the schools learned the different components of improvement science?" In addition to gaining their perspective, these interview results served as a form of members checks on my own analyses (Guba & Lincoln, 1986). Additionally, the protocol collected background information, albeit abbreviated, to better understand how their own history could shape their network experiences and perspectives (Seidman, 2013.) Each person was interviewed once. Three interviews were conducted: project director, math coach, and school liaison. The interviews occurred in May 2018 through video conferencing. The three interviews lasted approximately one hour each.

School teachers. The teacher protocol also asked questions related to teachers' backgrounds and teaching style, the school's background, and their experiences in the network. Specific questions were designed to gather data regarding their individual improvement science capacity, any pre-existing capacity, and processes that facilitated their learning. Each person was interviewed once. A total of six teachers were selected and interviewed following the purposive sampling previously described. The interviews occurred in April and May 2018, and were conducted via in-person (3), video conferencing (2), and telephone (1). Mixed modes were available because I had previous relationships with them, and thus felt each interviewee would be comfortable and open regardless of the mode. The choice of mode depended upon what was most convenient to the interviewer and interviewee due to the schedule and/or location. Each interview lasted approximately one hour.

Administrators. Per the interview sampling previously described, two administrators were interviewed. The protocol first asked questions regarding their background, leadership style, and the school. Then, questions were designed to learn more about their network experiences, their own and the school's improvement-related capacity, and whether anything has changed since participating in the network. The interviews occurred in April and May 2018. Each person was interviewed once, and it lasted approximately one hour. One administrator was interviewed by phone while the other occurred through video conference.

Observations, network convenings meetings. There were seven network convenvings, with all schools, held throughout the first year. For the most part, the primary objectives of these meetings were to build a collective network identity, develop and decide upon the fundamental network components (e.g., theory of practice improvement, build improvement science capacity, and engage in inquiry and dialogue. Six of these meetings were observed using an unstructured protocol (Appendix B). The first three network convenings (September, October, and November) were observed by me, as a participant. The last three meeting observations (January, March, and April) were conducted by another person because I was unable to observe, even as a participant, because of my role leading the meetings. The last meeting was not formally observed; however, my own notes and meeting artifacts provided data about this meeting. These observation periods lasted from approximately 8 a.m. to 3 p.m., with no observations occurring during the break periods (e.g., lunch).

Participant observations, school meetings. In addition to network convenings, the hub team also met with school teams at their schools. These meetings typically occurred during existing math department meetings and professional development time, and/or during a specifically designated time for network activities. The primary purpose of these meetings was for the hub team to facilitate improvement activities and build improvement science capacity. Participant observations were conducted at these meetings, using a *semi-structured* observation protocol that was designed to collect data specific to the research questions (Appendix C). This protocol was completed after the meeting and entered into an Excel spreadsheet. These meetings were also recorded for the purpose of further review if needed.

Observations were conducted on all network-related in-school meetings that occurred between October 2017 and April 2018, with the exception of the first round of meetings when expectations were discussed but no improvement-related activities commenced. The use of a more structured protocol allowed me to feasibly collect data from all of these meetings. The number of observed meetings by school varied depending upon the network hub's ability to schedule regular meetings, as discussed more in subsequent chapters. In the case of Middleview, one formal network-related meeting occurred at their school, but its purpose was to present the content shared at the January 2018 network convening because almost of none of them attended it. While hub staff attended other Middleview department meetings, these were not specific to the network, and thus, were not observed nor considered network in-school meetings. In those cases, I took informal notes when the network-related content was discussed, typically in the last 15 minutes of the meeting. In total 23 observations were conducted between October 2017 to April 2018, as shown in Table 4.

Table 4

Number of School Observations by Month (October 2017 – April 2018)

School	Total Number	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	April
Roosevelt	5	1	0	1	0	1	1	1
Marshall	4	1	1	0	0	1	0	1
Middleview	0	0	0	0	0	0	0	0
Sawyer	7	1	2	1	0	1	2	0
Central	7	2	2	0	0	1	1	1

Note. Even though no formal observations occurred at Middleview, informal notes were collected as part of my network role.

Network surveys. Network participants were surveyed twice throughout the first year (Appendix D). The purpose of these surveys was to collect data related to individual and organizational improvement science capacity, including organizational factors such as trust, collaboration, shared responsibility, and leadership. The first survey was conducted at the end of the first network convening (September 2017), and administered by paper. Twenty-three (23) surveys were completed, representing a response rate of 96 percent. The data were entered into Excel. The second survey was administered at the end of the last network convening (May 2018), and administered online via Qualtrics. The total number of surveys completed by teachers (17) and administrators (2) was 19, representing a response rate of 85 percent of the meeting attendees.

The survey items were designed by first reviewing relevant literature. Existing items were used or slightly modified when available in order to increase the validity and reliability of the survey instrument. When relevant items did not exist, I designed my own by using well-documented item development guidelines (Fowler, 2009). The first survey instrument was not pretested due to the short timing between its development and the first network convening; however, other team members reviewed and provided feedback. The second survey was pretested with a teacher and administrator who were familiar with improvement science but were not participating in the network. Slight revisions were made based on their feedback.

Documents and artifacts. Network and school meeting documents and artifacts were also collected throughout this study. With the exception of the PDSA forms, the primary purpose of these artifacts was to assist with the development of the network's narrative. The PDSA forms were further integrated into the analyses of the second research question because they operationalized improvement science implementation (outcome). Each teacher completed their own PDSA form for each cycle they conducted. I collected agendas, sign-in sheets, presentation materials, meeting artifacts (fishbone diagram, process maps, driver diagram, and other worksheets), and feedback forms from network covenings. From the school meetings, I took pictures of their fishbone diagrams, process maps, and beginning driver diagrams. I also reviewed their individual PDSA forms, which were stored on a shared Google Drive.

School characteristic data (secondary data). This study utilized school characteristic data from the CDE. Data were gathered regarding student demographics, student enrollment trends, and achievement data. The data were downloaded from the CDE website.

GSR notebook/informal interviews and observations. Additionally, due to my positioning in the study – both as a researcher and GSR assisting the network – I had access to my notes and calendar. I kept my own network-related notes in Microsoft OneNote, along with a hard-copy notebook that I carried to my school meetings. In OneNote, I documented the network's planning activities and processes, and they often included detail regarding why certain decisions were made. In the notebook that I carried to schools, I often wrote an informal agenda for the meeting, along with informal observations and/or conversations that were notable. Also, in this notebook, and my journal, I wrote reflections that were applicable to both my GSR and researcher roles. These data assisted with the detailed network narrative.

Analytic Procedures

Data analyses occurred in several stages throughout the study. These steps were not necessarily chronological and were fairly iterative. These analyses utilized both a deductive approach, which was driven by the research questions and conceptual framework, and an inductive approach that allowed emergent themes to be incorporated into the analyses. This section describes the analytical stages and procedures that were employed.

Stage 1, Preliminary Analyses

In November 2018, preliminary analyses were conducted on the first network survey, in-school meeting observations, and my GSR notes. The purpose of these analyses were to identify any emergent themes related to school network participation that were not already captured in the conceptual framework. The first survey was analyzed by developing frequency tables for each question. The in-school meeting observation data were already in a structured format and entered into Excel. The results of the survey, observations, and my notes were first compared, somewhat informally by reviewing each one and writing down themes, and then more formally by creating an Excel sheet that identified evidence of the themes by school. Three broad themes emerged: collaboration and facilitation issues within school teams, skepticism of the process, and perception of more improvement science knowledge and skills than they actually have. The identification of these themes allowed me to purposefully incorporate these potential issues into my subsequent data collection, if not already included.

Stage 2, Data Preparation and Analyses

Interview coding. Once all of the data were collected, I analyzed interview data for the sampled network hub staff and Middleview, Sawyer, and Central. Interviews were coded primarily using a deductive coding scheme; however, the coding also allowed room for new codes to emerge.

During the first cycle of coding, a-priori codes were established from the conceptual framework and research questions using a structural coding method (Saldaña, 2013). For example, first-cycle structural codes included: “individual capacity,” “organizational capacity,” “leadership,” “improvement science learned (what),” “improvement science learned (how),” “learnings generated (what),” “learnings generated (how),” and so forth. Another round of coding occurred within these broader structural codes that used a hybrid of descriptive, versus, and process coding methods (Saldaña, 2013). These codes inductively permitted themes to emerge within the existing conceptual framework. Analytic memos were also written throughout the first cycle coding. Before second cycle coding was applied, these codes were recoded and/or collapsed due to new codes being developed in later interviews that were not present in initial interview coding.

Pattern coding was used during second cycle coding to organize codes and attribute meaning to them (Saldaña, 2013). In some cases, the initial structural codes remained as the highest level code with new code categories added such as “facilitates improvement science learning” and “embedded in-school work structures”; in other cases, new pattern codes were developed by combining lower level codes into higher-level categories, such as “network math vision” and “network challenges.” The coding process utilized MaxQDA software. The interview coding framework is shown in Appendix E.

In-school meeting observation coding. The semi-structured observation protocol, which was entered into an Excel sheet and hierarchically structured data by research question and nested sub-questions and sub-topics, allowed for analysis without additional coding. However, in an effort to more systematically triangulate the interview and in-school observation data, observation data were further coded for the same sample of schools whose teachers were interviewed as part of this study, with the exception of Middleview who did not have any in-school meeting observations as previously discussed. There were 14 observations included in this analysis.

The data were imported from Excel directly in MaxQDA software, which automatically created deductive high-level structural codes from the column headings in the Excel sheet. Once imported, I conducted one additional round of descriptive coding to identify more detailed and/or any other themes emerging from the in-school meeting observations that were not in the initial protocol design. These codes were informed by the interview coding themes. This coding framework is shown in Appendix F.

Network surveys. The two network surveys were analyzed using descriptive methods because of their small sample sizes (n = 23, n = 19). For both surveys, the data were imported into SPSS to compute frequencies. Open-ended questions were coded into categories. Summary tables were developed to compare items across the schools, as well as to the results for the network as a whole.

PDSA cycles. There were 63 PDSA cycles completed during Year 1 of the network. These PDSA cycles were analyzed by first developing an Excel sheet that listed each PDSA form with the teacher name, school name, date, and change idea and form content. Then, each of these cycles was reviewed to determine if the change idea was aligned to the network's driver diagram. Finally, each of them were assessed to determine if they indicated single-loop and/or double-loop learnings. Summary tables that showed the number of PDSA cycles completed by school and teacher were also computed.

Stage 3, Research Question 1 and Narrative Process Tracing

The first research question was primarily answered through process-tracing techniques that generated a historical narrative of the network's first year, including how the network was established and how the hub team taught the schools improvement science and developed the fundamental network components. Process tracing is a within-case analytical technique for identifying causal mechanisms of a case study outcome, and is widely used in political science (George & Bennett, 2005). It is a method for the researcher to trace the processes that may have led to a specific outcome in a single case, while also acknowledging equifinality; that is, there could be alternative causal paths to achieving the same outcome (George & Bennett, 2005). Yin (2014) refers to this analytical method as "explanation building" that specifies a set of a causal links about why or how something transpired. Yin, however, provides little direction for employing this technique, thus, I relied upon political science practice for guidance.

Bennett & Checkel (2015) recommend that all or some of the following criteria, depending upon the specific research needs, serve as a practices that should be followed for good process tracing. I followed these best practices.

1. Cast the net widely for alternative explanations.
2. Be equally tough on alternative explanations.
3. Consider the potential biases of evidentiary sources.
4. Take into account whether the case is most or least likely for alternative explanations.
5. Make a justifiable decision on when to start.
6. Be relentless in gathering diverse and relevant evidence, but make a justifiable decision on when to stop.
7. Combine process tracing with case comparisons when useful for the research goal and feasible.
8. Be open to inductive insights.
9. Use deduction to ask "if my explanation is true, what will be the specific process leading to the outcome?"
10. Remember that conclusive process tracing is good, but not all good process tracing is conclusive.

While process tracing is typically tied closely to theory, as either a theory-testing or theory-development case study strategy, in can also be atheoretical in the form a detailed narrative (George & Bennett, 2005). In this form, process tracing aims to "throw light on how an event came about" through a historical chronicle (George & Bennett, 2005, p. 210). Furthermore, as is the circumstance of this research, an initial narrative account can provide enough detail about possible causal processes that the analyses can then be *converted* into a more theoretical explanation of the phenomenon (George & Bennett, 2005).

For this case study, I constructed a chronological narrative of the network by first creating a sequential timeline of significant events, along with the related activities and corresponding data, such as my GSR notes and network artifact and documents. The detailed narrative was further constructed within this timeline by reviewing all of the data sources. For example, the network convening observations and artifacts provided evidence regarding what occurred at those meetings and participant reactions. Through the analyses conducted in Stage 2, I further identified key moments and themes to highlight in this narrative. Because of my positionality, I wrote this narrative from the first person perspective in an effort to be thoroughly transparent about my role in the network.

Stage 4, Cross-school Analyses

Because of the case study design with schools as embedded units, I could further analyze interview, observation, survey, and PDSA data across the schools to answer the second research question. While the network as a whole was deemed a successful case, there was variation in the number of PDSA cycles completed among the schools. This variation allowed for cross-school, and even cross-teacher, analyses to learn what capacities, structures, and/or conditions contributed to improvement science implementation and subsequent meaningful learnings within a networked improvement community context. Cross-school analyses were conducted by first analyzing each data source across the schools and teachers within those schools, and then by triangulating those data to identify similarities and differences that corresponded with the PDSA outcome.

Stage 5, Furthers Test of Empirical Data

To strengthen the rigor of my analyses and the validity of my findings, I conducted further analyses following the process tracing best practices by systematically considering a wide range of alternative explanations for the findings and what evidence would be expected if the findings were true (or not true) (Bennett & Checkel, 2015). The initial findings that fell under the second research question were tested by developing an Excel spreadsheet that listed those findings, along with each potential alternative hypothesis that could explain away the finding(s). For each of these, I populated a cell with data related to the two bullets below:

- The evidence one would expect to see if the finding or alternative hypothesis were true. This test was similar to Van Evera's Hoop Test (albeit modified to fit my purposes where I did not have specific testable hypotheses) where the lack of evidence suggests that the finding is likely incorrect but the presence of evidence on its own does not prove the finding (Beach & Pedersen, 2013).
- The disconfirming evidence one would *not* expect to see if the finding or alternative hypothesis were true. The presence of disconfirming evidence suggests that more careful consideration should be given to whether the finding is likely correct, while the absence of disconfirming evidence strengthens the finding or alternative hypothesis.

Stage 6, Connection to Theory

The last stage of this analytic process was the consideration of relevant theories, in this case adult learning theories, that could partially explain the outcome of this case study. Again, this was part of the overall iterative process. The potential theories initially materialized through the detailed narrative and were further distinguished as part of the posited Year 1 improvement science capacity building model described in the Discussion Chapter.

CHAPTER 4

THE CASE OF THE NETWORK

Introduction

In September 2016, I was sitting in my advisor's office discussing potential dissertation topics, *again*. It had been months, and I still had not landed on the right one. As an applied researcher and program evaluator for the past 15 years of my career, I had little interest in the theoretical. Still struggling to find something practical, yet significant enough to warrant a dissertation, I was sure this meeting would end no closer to an answer.

But then, my advisor told me the University had recently received a grant from a philanthropic foundation to establish a networked improvement community. The grant's purpose was to construct "a network of schools that will work together in a sustained partnership for improvement in teaching and learning in diverse K-12 schooling contexts" (University grant proposal). A world-renowned public institution, the University had a strong commitment to social justice and improving the quality of the education workforce in today's urban schools. They had pledged to develop a math networked improvement community grounded in improvement science for its disciplined inquiry framework. I was intrigued. A newly forming network using improvement science theories in a practical application? We had found my topic.

My advisor immediately connected me with Jackie, who was overseeing the grant, and charged with developing and running the network.[2] We agreed that in addition to researching the network for my dissertation, I would also help Jackie coordinate it as a graduate student

[2] The names of people, places, and schools have been changed to protect the confidentiality of network participants.

researcher. Neither of us had experience building a networked improvement community from the ground up. We started planning the network in January.

Network Narrative

The Network Hub Team

Introducing Jackie. Jackie was an Associate Dean in the University's Graduate School of Education. In that role, she oversaw programs that connected the University with local neighborhood and community schools and had substantial experience directing school initiatives and providing professional development.

> ***Jackie's background and experiences.*** *I started out my career as a high school math teacher where I taught for 12 years. And at that time, I became engaged, I participated in the University Mathematics Project, which is professional development for teachers. And the following year, they asked if I'd be willing to direct the project. So, I actually came. I did that part-time, taught part-time and did the directing of the Math Project part-time. But that was not working, so I came to the University full-time as the director of the Math Project. At that time, the University was looking at its role with K-12 and they proposed this Center that would actually link research and practice. And so, I was one of the team that worked to create what is now The Center, and became its first director. So, I was the Executive Director of The Center, which was about the intersection of research and practice. I was faculty in the teacher education program. I supported and prepared teachers going into the teaching of secondary mathematics. I did that until about four years ago, when I became Associate Dean [of the Graduate School of Education]. So, all of my work has been about engaging in schools, and looking at practice, and trying to work on how we inform practice, how we transform practice and schools. So, that's where my experience comes from in this area.*

Motivation for starting the network. *One of the challenges we faced was that we don't want to see our work at each of these schools…they kind of become silo-ed. Like, "Here's what we do here, here's what we do there." Once I was brought into this position as Associate Dean, I could see that there was this great need for bringing our work together. There were ways that we could learn from each of the schools. The experiences were all so different, but it didn't mean that there weren't things we could learn from each other. I wanted to figure out ways to bridge, to link, to learn from. It was at that time that I started learning about improvement networks from Louis Gomez. It was his work. He started doing professional development with one of the schools around improvement networks and improvement science. That then led us to this discussion about how do we engage our other schools that the graduate school is connected to. [It] made us start thinking about this notion of network and what it means. He and I collaborated on trying to envision what it might look like. So, that's how it got started. And then thinking, "OK, we need funding." And then that's how it led to the funder, who actually contacted us first, to say, "How can we partner with you around work?" And that's how we started thinking about this network piece.*

Introducing me. As a program evaluator by experience and training, I was, and still am, particularly interested in how to drive change through evaluative methods. Before undertaking my Ph.D., I served as the Manager for Research and Evaluation at a mid-to-large-size northern California school district, as part of a two-year fellowship. Through that experience, I learned first-hand just how difficult it is to change an entrenched system, and realized that rather than prioritizing outside knowledge, we needed to find new ways to combine it with the collective knowledge and experience of the educators within. From that experience, I came to believe that sustainable, systemic change would require practical means for evaluators, researchers, and practitioners to work closely together. I was excited that this new networked improvement community was going to embrace these ideas.

After my fellowship concluded, I remained at the school district and led a project that implemented improvement science with the goal of improving the academic perseverance of middle school students. The project included principals (instead of teachers) from six middle schools, and select district leaders. While we were able to strengthen relationships, and created a new culture of openness among the participating district staff and principals, the project never gained traction, and the group disbanded after three years. I deemed the effort a failure because nothing actionable resulted from it. Because of this experience, I had mixed feelings about my role in the new University network. I was excited for the opportunity to work closely with schools and share my evaluation experience, but I was reluctant to step up and lead our network. I did not want to fail again.

I was also concerned about my positionality for my dissertation. *Was it appropriate to research a process, if I am part of that process? Was I influencing what I was studying?* My advisor and others assured me that it was acceptable, as long as I was open and transparent about my role.

So, here it is. While this is an empirical account of starting a networked improvement community, my own story is woven into the fabric of its narrative. The story of how wearing two hats – practitioner and researcher – continually informed each other throughout my quest to answer the overarching research question:

How do you prepare schools to successfully participate in a network improvement community?

Building the rest of the team. Unsure of how to start a networked improvement community, I dove into the research. There was very little available. I turned to *Learning to Improve* (Bryk, Gomez, Grunow & LeMahieu, 2015) and an article entitled *A Framework for the Initiation of Networked Improvement Communities* (Russell, et al., 2017). Step one: Build a network initiation team.

The network initiation team needed to serve several purposes. Its role was to recruit members, secure needed resources, and provide expertise in subject and improvement knowledge (Bryk, Gomez, Grunow & LeMahieu, 2015). It also functioned to identify the problem of practice, analyze the system that contributes to the problem, develop an aim statement, and draft an initial theory of practice improvement (Bryk, Gomez, Grunow & LeMahieu, 2015).

Jackie and I struggled with selecting our initiation team members. Who had relevant, complementary expertise, and the time needed to participate? Concurrently, we had to consider how to develop the aim statement, and the initial theory of practice improvement (also known as a driver diagram). We debated whether to first develop the aim and initial driver diagram with the initiation team, or to include all the network members in that development. Bryk and his colleagues (2015) deem this a crucial initiation question:

> Do we convene a small team to orchestrate the up-front work of refining the problem and framing a prototypical driver diagram and measures? Or do we first assemble the interested partners and have them identify a problem to pursue together? This strikes us as important tactical decision. (p. 160)

The importance of this tactical decision could not be overstated. We grappled with this question for weeks, carefully deliberating the pros and cons. I knew from my previous work that teachers would not fully own a change idea, nor the improvement science process, if they were not engaged at the onset (Rohanna, 2017). From her substantial school experience, Jackie also understood the importance of doing things *with* teachers, not *to* them. However, we recognized that convening the schools and teachers without an initial problem focus and asking them to build a driver diagram from scratch could be endless. We worried that teachers would find it challenging and frustrating as they were being asked to learn improvement science and find consensus with others. There was no perfect answer.

Just as we were about to make our difficult decision, we punted. (We would revisit this question again later.) We learned we were not the only University network starting in the fall. Another group had received an endowment to launch a networked improvement community. Like us, they were interested in math, and brought expertise in networked improvement communities and improvement science through their stronger affiliation with the Carnegie Foundation for the Advancement of Teaching. It made sense to join their network. We would be a sub-network within their larger network.

Jackie and I still needed to build our own hub team to support our sub-network. A core principle in improvement science is the inclusion of members with subject knowledge, and members with profound knowledge (knowledge of how to improve) (Langley et al., 2009). As the improvement science specialist, I filled the latter role. But we still needed a math expert to provide the subject knowledge. We found that expertise in Tom.

Introducing Tom. Tom was a math coach when he joined the network. He worked for the University's Math Project and had an extensive background in math professional development and instruction. His vision for engaging and quality math instruction would prove to be fundamental to our network's work.

> ***Tom's background and experiences.*** *I have Doctorate in Education, with an emphasis on teacher education, teacher education in multicultural societies. I have a Master's Degree in Education from Stanford, with a specialization in teaching mathematics. And then, a mathematics degree. That was my undergrad degree.*
>
> *I have been a math facilitator, math consultant, math coach. All things math-related. I've also been a professional development provider. I've been at many schools providing training and Cognitively Guided Instruction. A lot of our work in the Math Project has revolved around that. But, because of my background in middle school and high school math, I've also done a lot of professional development specific to that grade*

span of middle school and high school. I've also done a lot of coaching, both in elementary and middle school and high school. And I think a lot of that has been connected to, like, my prior life, which, all of it, was math teaching. I taught middle school and high school. I taught General Math, Algebra 1, Geometry, and Calculus.

Motivation for joining the network. *I'd received an email from Jackie basically inquiring about whether we knew, my colleague and I who worked for the Math Project, whether we knew someone who might be interested in a middle school and high school coaching. But I think it was very specific to algebra at that time, as well as another condition, which was whether the person would happen to have enough background in College Preparatory Mathematics, CPM, that program. I was like, "Ooh, that's me!" What excited me about it was just the emphasis on working with specifically algebra teachers. So when I saw Jackie at a school where we were providing professional development, Jackie and I happened to be in the same space, and we got to talking and that was what got me even more interested in participating.*

Introducing Samantha. We also needed someone on our team who knew many of the teachers and principals. Samantha was that person, and as she tells it, joining our team was a "happy accident." At the University, Samantha was the director of an education initiative with University neighboring schools. She was already tending to relationships at the schools she served, and saw an opportunity to liaise between the schools and us. She proactively offered to help facilitate access, advocate on the network's behalf, and represent the schools' perspective for meetings and developed materials. Rather than being an "accident," Samantha had the foresight to understand her crucial role in the network.

Samantha's background and experiences. *[In my University position,] I'm responsible for bridging enrichment, enrollment, and working closely with principals. We have eight schools and I'm responsible for four of them. Before that, I was at one of* *the network schools for about seven years. So, officially my job title there was Director of Bridging and Enrichment, but I was basically responsible for whatever the school needed. And that could be like working with the principal, to help with professional development, working with families, working with students, or helping place University resources at the school. Whether that's like student teachers or grants, volunteers, departments, donors, getting donors to come, helping distribute scholarships, to supporting kids, K-12, and beyond. I did my student teaching there, and I was a resident teacher there the first year it opened.*

I'm currently a doctoral candidate in the University's education leadership program. I got my masters through the teacher education program at the University. And I got my undergrad in sociology at the University.

Motivation for joining the network. *So, I first became aware when we started talking, you know. Jackie was talking about that there was a grant that she needed my help, like, talking to some of my schools. And at the time, I just thought it was another resource we were bringing. So I didn't even have any concrete plans. And then, as part of my job is to be on board, be on deck, to make sure that resources are happening at the schools. You know, schools are really chaotic and busy places. Sometimes everybody has the best intentions, but sometimes things don't happen just because either volunteers or groups don't really understand how schools work. Or schools really want the resources but they just don't have somebody to be there, like, navigating and making sure that the scheduling works. And, so I just kind of jumped on board that way. Then my role grew into something different, which is cool. But that's just sort of the nature of my job, is to always sort of be there in the beginning. To help facilitate and really to make sure things don't fall through the cracks.*

Recruiting the Network Schools

Five schools were recruited to participate in our nascent network. They all resided in a large, metropolitan, urban school district. Like many urban public school districts, it was experiencing years of declining enrollment, possibly attributed, in part, to the growth of independent charter schools (Rich, 2012). Between 2001 and 2017, district enrollment had declined by 15 percent (California Department of Education [CDE], 2017a). Declining enrollment was also due to the socio-economic factors of increasing area costs of living, including housing, and the demographic trend of lower birth rates.

Also similar to other urban districts, this district had overall lower-than-desired academic performance. Little more than half of the students in the secondary grades (55%) met or exceeded achievement standards for the state standards in English Language Arts. Less than a third (28%) met or exceeded the standards in math (CDE, 2017b).

The five network schools varied in their achievement scores, but all were concerned with improvement, especially in their math achievement. The schools had pre-existing partnerships with the University, under Jackie's purview. She had relationships with the principals, and personally called each one to ask if they would be interested in participating. They all said, "Yes." They were especially interested in the prospect of receiving additional math classroom coaching as part of their participation.

We commenced the network by meeting with the principals and teachers. From previous experience working with schools and conducting professional development, our team knew it would be difficult for teachers to use the improvement science tools on their own. We also knew it was hard for educators to meet and prioritize a continuous improvement process unless it was part of a set schedule (Rohanna, 2017). Thus, the purpose of these meetings was to hear their expectations, concerns and hopes, clarify our commitment expectations, establish a structure and schedule, assure facilitation support, and importantly, establish our expectation that the

University team would meet with the teachers at their schools, in-between the network convenings. These in-between meetings were a condition of their participation in the network. Maintaining this commitment would later prove to be trying at times.

Descriptions of the network schools. While the general characteristics of the schools are shown in Chapter 3, the five schools are described in more detail in this section. In total, 24 math teachers participated in the network.[3] Five principals were asked to actively participate, but not all did. One assistant principal also participated.

Roosevelt. Roosevelt was a district pilot school that was launched in the 2009 – 2010 school year, and enjoyed a close University partnership. Its enrollment had tripled since opening, its teachers and administrators considered their school "unique" and were proud of their accomplishments. A commitment to social justice and their active role in the community were part of Roosevelt's self-identity.

For the most part, the school's math performance was higher than the overall district's (CDE, 2017b). More than 40 percent of the 6th graders (44%) and 7th graders (45%) met or exceeded standards, which was higher than the district's (28% for both grades). However, 8th grade figures were lower than the district's (22% compared to 28%). While Roosevelt was outperforming the district, on average, there was still significant room for improvement: Fewer than half of the students were meeting or exceeding standards, and Algebra I grades data from the past few years showed that approximately half of the students were receiving Ds or Fs.

Through several years of working with the University on other improvement initiatives, Roosevelt teachers already had exposure to and/or experience with improvement science tools and processes. Their principal valued and promoted a strong culture of improvement, and expected the teachers to conduct PDSA cycles. However, some of the teachers expressed

---

[3] At the end of the first year.

concerns that they still did not completely understand the process, and suspected that they were possibly doing it incorrectly. This turned out to be an accurate self-assessment, based on my observation of one of their in-school meetings in October.

During this meeting, the teachers were sharing results from a PDSA cycle they had conducted independently from the network. They had one collective PDSA form for *all* of the teachers. Issues became obvious when they tried to find consensus and/or identify strategies that were conducive to each teacher's specific needs. The teachers taught a variety of grade levels and courses, and their one-size-fits-all PDSA form was not practical. Additionally, they were unclear what their prediction for the cycle was, or why there even was one. Thus, there was still substantial room to improve upon their improvement science skills. Despite that, their existing knowledge assisted them greatly throughout the year, and they had in-school work structures already in place. Their foundation was more developed than other schools. It was something to build upon.

Their network attendance was fairly consistent throughout the year. One teacher, who had been on leave missed most of the meetings in the fall, but returned by January. Two other teachers missed a couple meetings each, but the rest of the group attended all of the network meetings.

Marshall. Marshall was a school that aimed to provide its students with quality instruction and "experiences in art, technology, leadership, and athletics." The school's enrollment declined substantially between 2005 and 2015 (66%), but was starting to grow again in recent years.

The school's math performance was a little higher than the overall district's (CDE, 2017b). A little more than a third of the 6th graders (34%) and 7th graders (38%) met or exceeded standards. Thirty-one percent of 8th graders met or exceeded standards. While

Marshall was outperforming the district in those grade levels, there was clearly still cause for concern, since more than half of the students were not meeting standards.

Marshall's team consisted of six teachers who attended network convenings consistently, however, we struggled to schedule time to work with them between those meetings. Even though we had emphasized this expectation to the administrator and the teachers at our first in-school meeting, we had difficulty getting them to follow through on that commitment. They could not confirm a consistent day and time for the intermediate meetings, and often did not reply to Samantha's emails. We were only able to work with the teachers at their school four times throughout the year, and that was due to Samantha's diligence, tenacity, and relationships with them.

The administrator attendance at meetings was inconsistent. He attended the first network convening, and a portion of one other meeting in April. Importantly, at our first meeting with him and his teachers in October, he expressed concerns about following the improvement science process. He wanted flexibility to continue following his own approach for working with his teachers, rather than strictly following the network's process. Through this conversation and others, I suspected that he did not value the improvement science process, which likely contributed to the lack of urgency in scheduling in-school meetings.

Middleview. Middleview was a school going through a major transition. At one point in its history, it had been prominent neighborhood school. The school experienced a sharp decline in enrollment during the 2000s with its enrollment falling by more than half (68%) between 2001 and 2011 (CDE, 2017a). After that, enrollment continued to slowly drop. The decline in enrollment was mostly due to the large number of charters opening in the surrounding neighborhood. By 2016, it was a school in crisis and entered a new partnership with the University with hopes of restoring its once-held position as a leading neighborhood school.

Middleview was described as a high needs school. According to those connected with the school and the CDE data (CDE, 2017a), it had a relatively high population of students in foster care, students classified as special education and/or having disabilities, and English Language Learners. Less than one percent in any of the grades met or exceeded the math standards (CDE, 2017b). With the University as a partner, the school was in the process of rebuilding its instructional capacity. The year prior to the network year, the school only had two permanent math teachers – 6th and 8th grade. Seventh grade had a long-term substitute for the year. It was, by far, the most challenged school in the network.

When the network started, the school still did not have all of its math teacher positions filled with permanent staff. Two teachers attended the first network meeting. Of those two, one was a first-year teacher. Another beginning teacher, who did not initially attend the first meeting, also was later added to the team. A third teacher, also in their first year, joined the school in the winter. Because of this significant instructional transition, and their high needs population, Jackie suggested we take a different approach than the other schools. Our plan was to invite them into the network and support them, but we would allow them to take the lead, rather than insist they follow the same structures as the other schools. We did not visit their principal, who was in his third year leading the school, to clarify commitment expectations before the network started. We assumed that full participation in this process would be too overwhelming for the teachers, so we did not dictate that they focus on the to-be-determined shared problem of practice, and we did not initially ask them to schedule regular school meetings with us.

Notably, the assumptions we made about Middleview turned out to be wrong and had unforeseen consequences. Our hands-off approach allowed the district's continuous improvement initiative to swoop in. The teachers met regularly as part of their process, thus, making it more difficult for them to meet with us.

Middleview's principal was not engaged with the network. He did not attend any convenings, even though he was regularly invited. Participation in the network by the teachers was inconsistent. Administration was concerned that leaving their classrooms for the full-day network meetings was too disruptive for their students. More than once, they only allowed one or two teachers to attend, rather than the whole team. Most of the teachers missed many network meetings. We did not meet with them regularly at their schools, although we did attend other meetings at their school with them.

Sawyer. After a steady decline in enrollment, Sawyer was a school on the rise (CDE, 2017a). For years, parents from the neighborhood chose other schools for their children. The current principal, who had been in the position for a few years, had been working to rebuild goodwill and trust with the surrounding community and staff, and now, the school was experiencing an increase in attendance from local families.

On average, the school's math performance was a little lower than the overall district's. About a third of the 6th graders (31%) met or exceeded standards, which was similar to the district's percentage (CDE, 2017b). However, its 7th and 8th student percentages were less than the district, with 18% and 11% meeting or exceeding standards (CDE, 2017b).

When the first network meeting convened, the math department was short in permanent math teachers. The principal was still in the process of "right-sizing" the staff – some existing teachers were let go, and new teachers were being brought onboard. The initial math team that attended the first network meeting consisted of two veteran teachers and one new teacher. During the year, two other teachers joined the team. One teacher was already at the school but was teaching both math and science courses, and a new teacher joined in late fall.

Attendance was inconsistent at the fall network convenings. While the whole team attended the first meeting, and included their assistant principal who oversaw math, only one teacher and the assistant principal attended the second meeting, and they had to leave early. No one attended the third meeting. The drop off in attendance at the second meeting could be attributed to it falling on the same day as their parent-teacher conferences. However, their complete absence at the third meeting was due to a lack of communication with the teachers. They were either unaware of, or had forgotten, about the meeting. The assistant principal expressed his apologies.

However, attendance improved and was consistent during the second part of the year, after they had expanded their team. Most of the teachers came to the meetings. One did not attend in the winter and spring because she was participating in a teacher fellowship program that required her to miss more full days.

Central. Central could be described as stable and “welcoming.” Its principal had been there for years. Both the teachers and the principal depicted a diverse staff that reflected the student body. Their population of students included those from the surrounding neighborhood, as well as other areas.

Following the district trend, Central was experiencing declining enrollment (CDE, 2017a); however, during my time working with the school, I never heard anyone express concern, unlike at the other two schools. This is likely because the decline occurred to a smaller and slower degree than the others, and it still retained a healthy size (1,564 students).

According to the 11th grade math state assessment data, Central performed higher than the district in math (35% versus 24% meeting or exceeding standards) (CDE, 2017b). Yet, for obvious reasons, there were still concerns: Only about a third of the students were at or above standards. Additionally, almost half of the students were failing 9th grade algebra.

There was a “nice” mix of older and newer teachers in the school. Its three 9th grade algebra teachers who participated in the network represented that nice mix. One of the teachers was newer, having been there for three years, while the other two were veteran teachers. While all three teachers initially attended the network convening meetings, one of the veteran teachers attended sparsely in the spring. For the most part, his lack of attendance was not explained. While he was out of town for one meeting, he was expected at the other three meetings by his principal, yet did not show up.

At Central, the principal was very involved and supportive of the network. He rarely missed a meeting. He provided time during the school day for the three teachers to meet with us regularly during a conference period. He alternated that time so the teachers, who did not all have one common conference period, would take turns missing class. He sought coverage for the teachers during those meetings, which for the most part, occurred every two weeks. Before those school meetings, our University team would also meet with him to get his perspective on what we were going to cover in the meeting (and sometimes the upcoming network convening). He helped us to better understand his teachers' needs and their concerns about the network process.

Launching the Network

As previously described, we were a sub-network within a larger network when we first launched. Jackie and I made the decision to be part of the larger network because of their expertise in starting networked improvement communities, and strong connections to the Carnegie Foundation for the Advancement of Teaching. We had none of those advantages.

The large network was comprised of 16 schools, excluding ours. They were all high schools, while our schools and teachers represented 6th-9th grades. Their network focused solely on high school algebra. Like them, we had originally planned to focus on algebra, but during early conversations with our schools they pointed out how crucial it was to involve the other grades when considering problems around algebra. They were highlighting a *systems* problem. Jackie explained.

And they pushed that Algebra I is also in middle school. The issues surrounding it don't start in eighth grade or ninth grade or with Algebra I. It's everything that happens before, you know. So that was the thinking, that, while at first we thought it was going to be all about Algebra I, we were pushed by the schools. Algebra I is an issue, but it's also what happens in the previous level.

Because we subsumed into the larger network, our network adhered to their training activities. Their team took the lead on organizing and planning the monthly network meetings. We provided input, and communicated with our five schools, but did not play a role in deciding what content would be delivered at the network meetings.

We did, however, have a blueprint for building the improvement science capacity of our five school teams. Per my initial research, we knew it was crucial to determine the objective(s) at the outset when building evaluation capacity, (Preskill & Boyle, 2008). Improvement science is considered a form of participatory evaluation (Cousins & Earl, 1992; Christie, Inkelas, & Lemire, 2017). As such, there were three broad capacity building objectives to consider: knowledge and understanding, skills and behaviors, and affective or attitudes (Preskill & Boyle, 2008; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012). Our primary objective in the first year was for teachers to learn processes and embed them in their schools. Thus, our focus was on building skills and behaviors, and transferring those skills to their work context (Preskill & Boyle, 2008). While gaining knowledge and understanding seems like a notable first step, we took the position that our team's improvement science specialist (me) would provide that expertise by facilitating the improvement science work at the schools. Of course, we were also concerned with attitudes. We knew behaviors would be more likely to stick if teachers saw the value in this work. Fortunately, as we found out later, Tom's math expertise provided this aspect.

Large network convenings (August to November). The network began with a convening of all the principals. This occurred in August after school started. Three fall network convenings followed. The purpose of these meetings was to teach the schools improvement tools and methods in order to define a common problem of practice, establish an aim statement, and eventually, develop a driver diagram. Figures 2 through 4 outline the content taught and the activities facilitated at each of these meetings.

Month of Convening	**Sessions/Content**	**Activities**
September (8 a.m. to 12 p.m.) (All network schools)	Introduction of network	Presentation of concepts
	6 Core Principles of Improvement	Presentation of concepts
	Identifying the problem	Presentation of concepts; Brainstorming landmark problems by school; Fishbone Diagram (Cause and Effect Diagram); 5 Whys.
	Empathy interviews	Presentation of concepts; Active listening exercise; Practice empathy interviews.
September (12 p.m. to 3 p.m.) (5 schools)	Real-world example of school improvement process	Presentation by a principle.
	Reflection on plans for own improvement	Discussions within school teams.
	Understanding the problem	Using one tool (Fishbone Diagram, 5 Whys) continued brainstorming around their landmark problem in school teams.
	Reflection	Whole group reflection on the day.

Figure 2. September large network convening content and activities.

Before the October convening, the large network (including our schools) were given several pieces of homework to complete:

- Continue working on their fishbone diagram, and possibly include other colleagues.
- Gather any data related to their landmark problem and bring to the next meeting.
- Conduct empathy interviews and/or observations with at least two people in their school.
- Find and read literature associated with their landmark problem to gain an understanding of what the field says regarding this problem.

We met with all of our schools, except Middleview, between the October and November convenings. The primary purpose of these meetings was to outline our expectations, learn about their expectations, and facilitate the fishbone diagram as a part of their root cause analysis. To facilitate these meetings, I developed “kits” that included Post-It notes, a hand-made 14 x 18.5 inch fishbone diagram, and fishbone diagram examples. The secondary purpose was to establish network-related work structures within the school.

We knew time was a scarce commodity for teachers. Even with the best intentions, it is difficult for educators to find the time and energy to commit to an improvement endeavor (Rohanna, 2017). As enticement, we offered to help them complete their homework and prepare for the meetings, if they agreed to commit to one or two set meetings a month. This tactic created a set time for network-related work, and minimized any additional work for them outside of that set time.

Month of Convening	Sessions/Content	Activities
October (8 a.m. to 2 p.m.)	Fishbone and high-leverage problems	Presentation of concepts and examples; Re-working fishbone diagrams in school teams; Paired schools shared with each other; Whole group share-out, one school at a time.
	Finding a shared network problem of practice	Identify, categorize, and prioritize potential problems of practice through school discussions and whole group voting (stickies and stickers).
	Further prioritize top problems	Effort and benefit continuum prioritization, within school teams and then as a whole group (stickers).
	Fishbone on one of the two top priorities (lack of pre-requisite skills and mindsets)	Re-do fishbone in school teams. Shared in small inter-school groups.
	Pareto chart	Presentation of concept and directions.

Figure 3. October large network convening content and activities.

The schools were not given any homework between the October and November meetings. We continued to meet with our schools individually.

Month of Convening	Sessions/Content	Activities
November (8 a.m. to 3 p.m.)	Reminder of the network process and purpose	Presentation
	Continue causal analysis with fishbone diagrams	School teams revisited, and possibly revised, fishbone diagrams from the last network meeting. Whole group share-out by a couple schools.
	Pareto charts	Presentation of concepts
	Process maps	Presentation of concepts, components, and directions; Scaffolded exercise; School teams developed their own.
	Aim statements	Presentation of concepts and directions. School teams developed an aim statement. Whole group share-out.

Figure 4. November large network convening content and activities.

Decision to separate from the larger network. From the beginning, we were able to structure our sub-network differently from the larger network. Our sub-network was small and manageable. We had five schools and existing relationships with them. With our core team – an improvement science specialist, a math coach, and a liaison with knowledge of and relationships with the schools – we had the resources to support our schools in both the improvement science and math content.

The larger network also had a team, but was limited by resources. They supported 16 schools, but only had two or three people available for additional support within those schools. Unlike our team, no one in their network had math instruction content expertise, and only one, other than the project director, had improvement science experience. During our meetings together, they acknowledged this made it challenging to provide the additional within-school support that our network provided. However, they did make themselves available for individual school support. Initially, it was by request, and then eventually, they began making more regular visits.

Due to the support of our team, our network progressed more quickly than the larger network. As previously discussed, we met with the schools individually between large network meetings to facilitate the improvement science activities. We also engaged them in collaborative dialogue around math instructional issues. By the end of the second meeting (October), it was clear our schools were further along because we had already facilitated some of the activities that were occurring at the network meetings. So much so, some of the activities at the second meeting felt redundant to our teachers. At the third meeting in November, one of our principals questioned the purpose of doing a fishbone diagram because through our in-school support, they had already progressed past the causal analysis and were beginning to brainstorm potential drivers on how to improve Algebra I pass rates. It had become apparent we would have to re-evaluate our participation in the larger network. This reconsideration was based on a few factors:

- Our teachers were frustrated. By November, they had spent three months trying to understand the problem of practice. Although we did not want to rush to find solutions, we risked losing their interest and motivation if they did not get to try something new in their classroom soon.
- We had concerns about the direction of the larger network's common problems of practice. Both problems were centered on student deficits: students' gaps in knowledge and lack of growth mindsets (students and teachers). We felt the teachers needed more direction framing a problem of practice that was within their locus of control and integrated with their instructional practices.

- The schools were not uniting around a common problem of practice. At the end of the third network meeting, the schools developed and shared aim statements, i.e., measurable goals for improvement. (Three of our five schools shared their aims. One school's team did not write one because of internal disagreements. The fifth school did not attend the meeting.) The foci of the three aim statements were disparate and distant: improving student problem-solving skills, improving the percentage of students passing Algebra I, and improving the process of student grouping in the CPM curriculum. Our schools were not moving closer in their focus. They were actually moving further apart.

Given those developments, we realized we needed to bring more cohesion and a common purpose within our five schools. And so, we made the decision to separate from the larger network and relaunch our small network.

The Re-launch

At this moment, I valued Tom and Samantha's expertise even more. I was also more confident in my improvement science expertise. While participating in those larger network meetings, I realized my knowledge was on par with theirs. I even wondered if I had a distinct advantage in my program evaluation background. As discussed earlier, improvement science is a form of participatory evaluation (Cousins & Earl, 1992; Christie, Inkelas, & Lemire, 2017). My comfort and deep understanding around disciplined inquiry, theories of change, and data and measures, along with my experience working for a school district, would inform my efforts to unpack technical ideas, and hopefully, render them more accessible for teachers.

Even with an amazing team and a newfound confidence in my improvement science knowledge, I was still unsettled about next steps. It was December. To keep momentum, we needed our teachers experimenting in their classroom by early February, which left us with one network meeting in January. *How were we going to develop a new shared problem of practice, a common aim, a driver diagram, and change ideas to test in less than two months?*

I sought the counsel of another established math networked improvement community, and arranged a conference call with the two people who led that network. They shared their experiences and provided valuable guidance. Having previous experience introducing improvement science to principals, I was unsurprised by two experiences that they shared because they mirrored my own. First, they purposely minimized the importance of the driver diagram. Teachers were not engaged by it and became bored. They suggested that I move the driver diagram to the background, rather than the foreground of the work. Second, they were flexible with the requirement that PDSA cycles be short iterative cycles. They were comfortable with the longer cycles (i.e., more than a couple weeks), which seemed to embrace the idea of creating a longer plan, rather than testing small changes. Ironically, these were my two non-negotiables.

I understood their sentiment toward driver diagrams. It reminded me of my own previous experience working with middle school principals trying to improve the academic perseverance of students. The principals' eyes would glaze over any time I presented the driver diagram. It seemed too technical to them. As a result, I, too, minimized its importance. The principals all had access to it, but I rarely showed it after its initial introduction. The result: They never remembered our theory for improvement and did not develop change ideas *intentionally* aligned with it, if at all. Thus, it was not serving its theory of change purpose.

Also based on previous experience, I was resolved to short iterative PDSA cycles. The intention behind PDSA cycles is learning, generated from the cycles (Langley et al., 2009). Each time a small change is tested, a new learning is generated. Thus, the formal cycle of learning is the same length as the PDSA cycle. With my former academic perseverance project, the learning was slow. The principals committed to one PDSA cycle in the fall, and one in the spring. They met after each cycle, shared the results, and reflected together. Beneficial, yes, but not likely to be truly transformative. (This was the aforementioned process that never got traction and the group disbanded.) Rather than generating rapid incremental learnings, completing cycles that fit within their schedule became the goal. During that project, we lost sight of the PDSA purpose when we tried to adapt it the school setting. I did not want to make the same mistake twice.

I also subscribed to a view from the *Toyota Kata* (Rother, 2010). We were in this for the long haul. Changing organizational learning behaviors meant building skills and establishing structures, routines, and habits (Rother, 2010; Preskill & Boyle, 2008). And, it was important to teach the *right* habits (Rother, 2010). Being too accommodating could lead to the wrong habits. Through my experience with Roosevelt, I realized that the wrong habits could lead to extreme frustration if teachers were authentically engaging in the improvement science process, but then getting stuck, or not seeing promised results. Of course, the trick is understanding when to be flexible and when to be strict, when to adhere to the rigorousness of the process and when to make adaptations. For me, I drew that line at this whole network's purpose: generating meaningful learnings that could solve complex problems of practice. I wanted to accelerate learning.

I shared my perspective and my own commitment to keeping the rigor of these two improvement science processes with the two people who led the other network. They offered two invaluable pieces of advice.

1. Own the driver diagram. This meant letting the hub develop it, rather than trying to bring the teachers to consensus.
2. Establish a math instructional vision. This meant giving the math expert (Tom) a larger role in the network convenings so that he could more explicitly guide teachers' experimentation in their classrooms.

Tom, Samantha, and I brought these two pieces of advice together as we developed the essential network elements needed to propel our small network forward. Once again, we faced the initiation team question that Jackie and I had dodged (or so we thought). The tactical question that Bryk and his colleagues (2015) consider essential:

> Do we convene a small team to orchestrate the up-front work of refining the problem and framing a prototypical driver diagram and measures? Or do we first assemble the interested partners and have them identify a problem a problem to pursue together? (p. 160)

Even though only a few months had passed, we were in a better position to answer this question. Our answer: Do a hybrid. Because we facilitated activities and conversations at our schools – we had artifacts from their causal analysis, and had even begun drafting drivers with some of the schools – we knew what problems they had identified, and some of the instructional issues they were struggling with. We gathered their fishbone diagrams, process maps, beginning driver diagrams, and our notes. We three would use these artifacts to refine the problem and develop the preliminary driver diagram. We felt this hybrid strategy was advantageous because it represented our teachers' views and experiences, without placing the burden on them to learn and develop a driver diagram, or to find consensus, both within their own teams, and among the other schools. We would serve as the consensus-makers. Yet, we were still uncertain how they would receive it. *Would they embrace it as their own?*

Before establishing the driver diagram, we needed to determine the network problem of practice. We had two criteria: teacher-focused, rather than student deficit-focused, and sufficient flexibility for each school's needs, while begetting coherence among them.

I remembered something one of our former larger network colleagues said at a planning meeting. Several of us were sharing our concerns with the two initial problems of practice selected by the large network. He articulated that the real issue was the "lack of alignment between the learning needs of kids and teacher skills." He was right. In one form or another, our teachers had expressed that they were unsure how to deal with incoming students who did not have prerequisite math knowledge, lacked problem-solving skills, struggled to make mathematical connections with real-world applications, and/or simply had no interest in mathematics. We also heard many of them express that today's students processed information differently because of the Internet and social media.

This provided an opportunity to pivot from what teachers perceived students were lacking to the real issue: Our current instructional practices were not aligned with students' learning needs today. While possibly too broad of a problem statement for improvement science purists, it provided coherence with flexibility, and importantly, a path forward for our teachers from diverse schools who represented multiple grade levels.

Initial network problem of practice**:** *Our current practices are not aligned with students' learning needs today. Many students are failing math.*

It was almost poetic that shortly after New Year's Day, Tom, Samantha, and I "locked ourselves in a room" to develop the brand new initial driver diagram. We first agreed upon a potential network aim. Rather than set measurable targets for the schools, we decided upon a global aim that was broad. We would let the schools set their own targets, again recognizing that each school had different needs.

Initial network global aim statement: *Our practices actively engage students in math and meet their variety of learning needs. We know we're improving when more students are learning and passing math classes with a C grade or better.*

Tom's notions of math instruction guided our development of the driver diagram. During the fall, Tom sent me a copy of the National Council of Teachers of Mathematics (NCTM) (2014) report, *Principles to Actions: Ensuring Mathematical Success for All.* We made connections between specific passages and what we were hearing at the schools. Importantly, the report also depicted an image of effective teaching and learning in mathematics.

> An excellent mathematics program requires effective teaching that engages students in meaningful learning through individual and collaborative experiences that promote their ability to make sense of mathematical ideas and reason mathematically. (p. 7)

With this report and the school artifacts, Tom, Samantha, and I drafted a preliminary driver diagram. We sequestered ourselves in one of the University's classrooms, equipped with the ubiquitous wall-sized whiteboard at the front of the room. They sat at the student tables, while I stood in front. I facilitated the discussion – asking Tom and Samantha questions, pushing their thinking, and pressing them to explain why one action would lead to another – while I jotted down ideas on the whiteboard and sketched out how they were linked, working and reworking primary and secondary driver descriptions. My hopes in that room were that we would complete the diagram, but at the end of the third hour, in what was supposed to be a two-hour meeting, we had only drafted one primary, and several secondary drivers. And, as Samantha stated, "My head hurts." Her comment reminded me of the time I taught a friend how to create a logic model. After two hours, he said that he could not think anymore. I shared that story with Samantha, and offered that maybe I took the effort for granted. My brain had been trained to think in terms of causal pathways from years of evaluating programs. She likened creating the driver diagram to training for a marathon: You have to start small. You cannot just go out and run 26 miles. You run a couple miles and build up endurance. For you (me), it is like you can already run the marathon. I reflected upon her comment. Theories of change and driver diagrams are technical and complex. *Why would we ever expect educators to be capable of doing this with no experience and little training?*

It took a second rigorous three-hour meeting to complete the driver diagram. Armed with the phrase, "possibly wrong and definitely incomplete" (Bryk, et al., 2015), we felt that we had a solid preliminary driver diagram that could be shared with the network. Our roll-out plan was to detail how we created the diagram from their perspectives and artifacts, and then solicit their feedback for revisions. Figure 5 shows the initial driver diagram.

Even though we were satisfied with the preliminary driver diagram, I was still left with one nagging feeling. It was not very systems-oriented in the traditional sense of improvement science.

One of the four tenets of profound knowledge is the knowledge of systems (Langley et al., 2009; Christie, Inkelas, & Lemire, 2017). As mentioned previously, profound knowledge is defined "as the interplay of theories of systems, variation, knowledge, and psychology" (Deming, cited in Langley et al., 2009, p. 75). Systems knowledge refers to the interdependence of departments, people, and processes within an organization. Langley and his colleagues (2009) assert that "understanding the organization as a system" is one of the components of a system of improvement (p. 312).

The idea of systems thinking in education was a concept that I grappled with frequently. The term itself, "systems," was very abstract to me, even though I understood its significance. For example, a school district's assessment policy could inadvertently reduce the number of instructional hours in a classroom, or that the incoming mathematical knowledge of new 7th grade students greatly depends upon the 6th grade math instruction. Yet, in education, these interdependent policies, people, and processes are often outside the teacher's locus of control. For me, the system needed to be bounded by the people in our room, whom were primarily teachers.

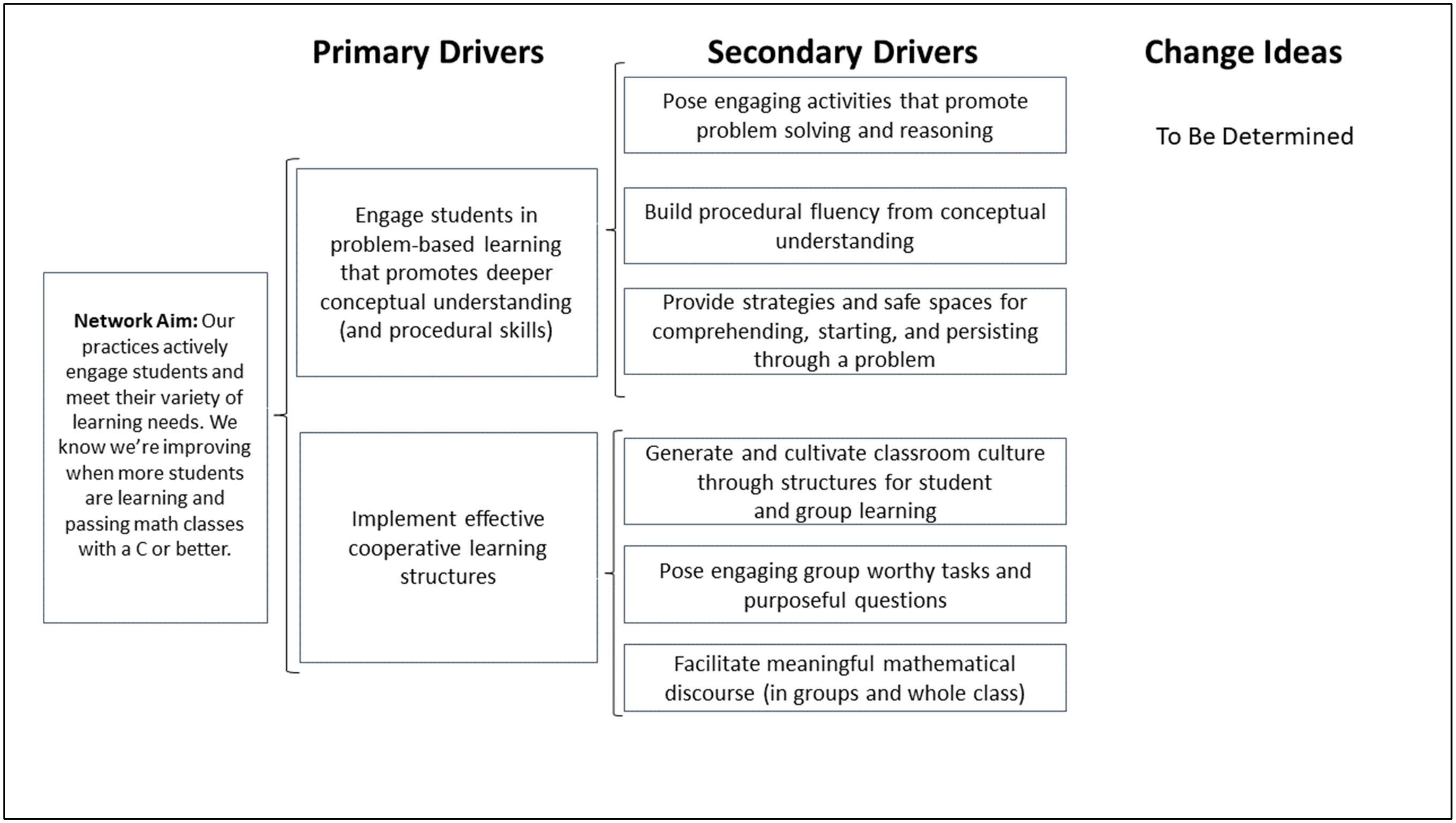

Figure 5. Preliminary network driver diagram developed by the University hub. *Note*. This version includes teacher feedback regarding minor wording changes.

Langley and his colleagues (2009) assert this idea that systems need to have boundaries. "The larger the system, the more difficult it is to optimize" (Langley et al., 2009, p.78). This phrase reminded me of an observation shared by one the principals from my previous project. During one of those meetings, I projected a slide with the Central Law of Improvement: "Every system is perfectly designed to deliver the results it produces," and explained the importance of interdependent parts of a whole (Langley et al., 2009, p.79). One principal immediately spoke up and lamented: Every time someone brings up systems, they never mention that there are a lot of parts outside of their control. For example, family life is part of this system, but they cannot control that. His obvious frustration and disdain towards systems thinking had stuck with me, years later.

Remembering that previous experience, I felt the system needed to be small and manageable. Because our network was very instructionally focused, we considered their classrooms to be their systems. I also resolved to drop the abstract term "systems" when talking with educators, and instead replace it with a more concrete depiction, i.e., the classroom. While this all made sense from a practical standpoint, it was still unsettling. I worried that the boundary was too small from a more traditional systems-thinking sense. Yet, returning to that idea of where to draw the line between rigor and flexibility, I chose flexibility. Narrowing the system was more likely to propel us forward and foster learnings, rather than inhibit us.

Network convenings (January to May). We held four network convenings between January and May. For each of those meetings, we established objectives and planned relevant activities and content, shown in Figures 6 through 9. We posted meeting materials on our newly developed network website. We continued to meet with schools between the convenings as we did during the first half of the year.

January convening. Our goal was to have the schools identify a change idea aligned with the driver diagram that they could experiment with for their first PDSA cycle. We wanted them to start and complete the cycle within two weeks of that first meeting. Toward that meeting goal, we provided the following content and professional development.

Month of Convening	Sessions/Content	Activities
January (8 a.m. to 3 p.m.)	Reminder and overview of the network's purpose and the improvement science process	Presentation of our progress thus far, and a timeline.
	Network problem of practice and aim statement	Presentation of how we developed from their artifacts and discussions.
	Math professional development regarding implementing tasks that promote reasoning and problem solving and mathematical discourse	Presentation and interactive activities where teachers experienced being the students for number sense routines.
	Driver diagram: What it is and how to use one?	Presentation of concepts; School teams provided input on poster-sized diagram and selected drivers where they could improve.
	PDSA: What, why, and how.	Presentation of concepts.
	Change idea	Introduction and brainstorming ideas in school teams.
	Planning our PDSA cycles	Individual work time for teachers to plan PDSA cycle. Fishbowl activity where two teachers shared and other teachers asked questions.
	Reflection	Group reflection on the day.

Figure 6. January network convening content and activities.

Because we were still concerned with whether our teachers would accept the problem of practice, aim statement, and the driver diagram as their own, we explicitly described how we developed these network essentials from their artifacts and dialogue. To our surprise and gratification, we received positive feedback regarding the driver diagram. Rather than dismissing it, the teachers expressed their appreciation that they were not forced to spend hours trying to create one and find consensus.

The teachers were asked to complete one PDSA cycle before the next network convening. Because we considered the classroom to be each teacher's system, each one was expected to complete and document their own PDSA cycle on our network website. That is, each teacher was asked to experiment with a new practice in their classroom, collect data that could inform whether they met their prediction regarding what they expected the new practice to achieve, reflect upon those data, and then decide next steps.

We scheduled time to visit each school, except Middleview, who only had one teacher attend the January convening. The purpose of those meetings was for teachers to share the results of their PDSA cycles with each other and engage in inquiry and dialogue.

March convening. When we visited the schools after the last network convening, I noticed that teachers were reluctant to ask each other questions. They seemed comfortable sharing their results, but often when I would ask, "What questions do you have for [teacher's name]," an awkward silence would fall over the room. In some cases, teachers would ask for some small clarification on what the other teacher had done, or offer a suggestion for doing something differently next time, but I judged the conversations to be polite and benign. On rare occasion, a teacher would really push another teacher's thinking, or forced them to justify why they had done something. This concerned me because the purpose of the PDSA cycle is to build knowledge through the reflection component, but the teachers seemed hesitant to challenge each other. (Langley et al., 2009). Langley and his colleagues (2009) state:

> Deductive and inductive learning are built into Plan-Do-Study-Act (PDSA) Cycles. From Plan to Do is the *deductive* approach. A theory is tested with the aid of a prediction. In the Do phase, observations are made and departures from the prediction are noted.

> From Do to Study the *inductive* learning process takes place. Gaps to the prediction are studied and the theory is updated accordingly. Action is then taken on the new learning. (p. 82)

I worried that meaningful learnings would not occur if teachers were reluctant to ask deeper probing questions. Furthermore, following Engelbart's ABC Model for Organizational Improvement, the goal was not just individual learning but team learning that would then ascend to network learning (Engelbart, 1992). To build team learnings, individuals need reflection, inquiry, dialogue, and discussion skills (Senge, 2006). In my experience with disciplined inquiry cycles in education, I have witnessed on numerous occasions (and participated in organizing) educators being brought together with expectations of engaging in inquiry and dialogue around some important issue and/or data, yet no one had first taught them how to do it. *Why did we expect productive inquiry and dialogue to occur just because we put them all in a room together?* I reflected upon this question and decided to explicitly teach them inquiry and dialogue skills.

Toward this end, the primary objective introduced at the March convening was instruction on how to engage in productive inquiry and dialogue, use those new skills to engage with teachers from other schools, and then plan their next PDSA cycles. Our secondary objective was to introduce data and practical measures, including classroom observation rubrics aligned to our driver diagram (developed by the hub team). For the primary productive inquiry and dialogue objective, we began by unpacking:

- What it means to engage in inquiry: asking questions, clarifying information, probing.
- What it means to engage in dialogue: exploration of issues through multiple people and/or perspectives, with a purpose of expanding our own individual understanding (Senge, 2006).

- What it means to engage in productive inquiry and dialogue: improve their own practices and reflect upon personal assumptions, potentially leading to a change in how they see and do things.

Month of Convening	**Sessions/Content**	**Activities**
March (8 a.m. to 3 p.m.)	Reminder of network process, problem of practice, and aim statement	Presentation
	Math professional development regarding implementing tasks that promote reasoning and problem solving and mathematical discourse (building upon January's content)	Presentation and interactive activities where teachers experienced being the students for number sense routines.
	Engaging in productive inquiry and dialogue	Presentation of concepts; Modeling and practicing inquiry and dialogue
	PDSAs: Productive inquiry and dialogue across schools	Each teacher shared their PDSAs results with teachers from other schools. Engaged in inquiry and dialogue using a protocol.
	Math professional development regarding making sense of students' thinking	Presentation and interactive activities
	Planning next PDSA cycle	Individual work time for teachers to plan PDSA cycle.
	Data for improvement (practical measures) and system of measures	Presentation of concepts; brainstorming measures aligned to drivers; discussion around potential rubrics and run chart.
	Reflection	Group reflection on the day.

Figure 7. March network convening content and activities.

Next, we reviewed norms for engaging in productive inquiry and dialogue. We relayed our expectation that productive inquiry and dialogue is a skill that can be honed through practice. Our team modeled asking questions and gave teachers a chance to practice. We also provided a protocol to guide their inquiry and dialogue around PDSA cycles.

While we wanted teachers to learn to engage in inquiry and dialogue with teachers from their own schools, it was also important for us to build this routine at the network level. To learn from other schools and build a network identity, teachers needed the opportunity to interact with one other. Therefore, teachers from different schools were strategically grouped together to engage in inquiry and dialogue. We grouped teachers who we felt would connect and build relationships, provide valuable guidance, and/or push each other's thinking.

Regarding the secondary data and practical measures objective, teachers were asked to choose and test one classroom observation rubric and incorporate it into their current PDSA cycle. We sought to collect data regarding:

- Engaging *all* students in mathematical activity.
- Providing students with opportunities to explain their thinking.
- Providing opportunities for student-to-student questioning.

The assignment was to collect the data before, during, and after they tested their change idea, using a run chart format (similar to a line chart), which we developed as an Excel template for each of the rubrics. We hoped they would see variations between the day(s) when they implemented their change idea – an increase in student participation or mathematical discourse would be typically expected – compared to the other days. They were also asked to provide feedback on the rubric itself, e.g., whether it was suited for the task, whether the wording could be clarified or improved, and other impressions. **Not one teacher from any of our five schools did it.** (Although two teachers did provide rubric feedback.)

I became aware of the situation before the April convening, thanks to our intervening in-school visits. During one of those visits, I asked if anyone had a chance to pilot the rubric or the run chart. Two teachers replied, “What is that?” Other teachers vaguely remembered the rubric exercise, but only after I described the procedure again to jog their memories. One teacher did remember the assignment, and kept proclaiming, "It was the yellow piece of paper!” to the others, as if that were a better clue than my description. To her credit, some teachers did remember a yellow piece of paper. Unfortunately, no one could explain what was written on it.

To their defense, the rubric and run chart assignment was fairly technical. Some teachers stated they did not fully understand the directions. This confirmed what I observed at another school, a few days after the March convening. Two teachers were collecting data for the run chart, but upon review, I realized they were not looking at the rubric descriptors while giving the ratings. Thus, they were not doing the activity as intended, and also, did not complete it.

Teachers also described feeling stressed. "It was a long day,” they explained, during a time of year that already demanded a hectic pace at their schools due to the state testing calendar. As one teacher described:

> *I’m involved at the meetings [but] the minute I leave… I got to come [to school] and deal with who didn’t learn what, and I got to re-teach….and [then] get ready for progress reports. All that.*

This was a common concern among teachers from all our schools. They described feeling “overwhelmed” and expressed concerns about missing a classroom day for a network convening, with state testing right around the corner. Even the "yellow paper" teacher revealed that she had forgotten the details of the assignment by the time she returned to her school. (One positive result of this meeting was the realization that teachers were more comfortable being candid and asking questions about technical content in this setting, rather than in the larger network meetings.)

April convening. We responded to their concerns. At the University team pre-meeting, we discussed a need to ease off the technical rigidity for the April convening. We designed activities where teachers would lead discussions, instead of being presented with technical ideas from us. Our re-calibrated objectives for the upcoming convening would be: build relationships among schools, have teachers reflect upon their own beliefs regarding student engagement in mathematics, and plan their last PDSA cycle of the school year.

Month of Convening	**Sessions/Content**	**Activities**
April (8 a.m. to 3 p.m.)	Math professional development regarding implementing tasks that promote reasoning and problem solving and mathematical discourse	Presentation and interactive activities where teachers experienced being the students for number sense routines
	Circles of engagement: Beliefs and assumptions around student engagement in the math classroom	Reading with protocol and facilitated small group discussion (mixed schools)
	PDSAs: Productive inquiry and dialogue across schools	Time to complete their reflection from the last cycle. Then, each teacher shared their PDSAs results with teachers from other schools. Engaged in inquiry and dialogue using a protocol.
	Spotlight from PDSA cycles	Whole group sharing out of inspiring practices heard from the PDSA discussion
	Mapping change ideas to driver diagrams	Team activities to develop latest change idea, map it to the driver diagram, and present to whole group
	Planning next PDSA cycle	Individual work time for teachers to plan PDSA cycle.
	Reflection	Group reflection on the day.

Figure 8. April network convening content and activities.

We also continued to build network routines around the PDSA process. Again, we grouped teachers from multiple schools and asked them to engage in inquiry and dialogue about their PDSA cycles using the established protocol. Additionally, at the end of this session, we asked the teachers to volunteer an inspiring practice that they heard in their group. The purpose of this activity was to highlight the learning that was occurring across schools, and acknowledge the positive practices of our teachers.

May convening. Due to state testing occurring between the April and May convening, we were only able to meet with one school before the May convening. Our team also internally relaxed our requirement to complete the last PDSA cycle. Not only did we want to remove further pressure on the teachers, but I did not want them to be inauthentic to the process. By reviewing the shared PDSA forms on our network website and through discussions with several teachers, I knew many of them did not complete their last PDSA cycle. My fear for the May convening was that they would *say* they conducted a PDSA cycle, without actually having completed one. Again, I had experienced this with the principals on my previous project. Therefore, even though we had asked them to plan a PDSA cycle at the April convening, we did not plan for them to share the results at the May convening. Our objectives for that final meeting were to continue developing a shared understanding of the network's purpose, and to plan for next school year by setting measurable aims, and revisiting our root cause analysis and driver diagram. Importantly, we also conducted a PDSA "showcase" – a recommendation from one of our teachers – that asked several teachers to demonstrate one of their change ideas from the year and/or a PDSA-related resource. The goal was for teachers to learn new ideas and/or strategies from each other, particularly those from other schools, rather than engage in disciplined inquiry. The showcase was a success. Teachers excitingly asked each other questions, and gained new ideas to try in their classrooms.

Month of Convening	Sessions/Content	Activities
May (8 a.m. to 3 p.m.)	Reminder of network process and purpose	Presentation with reflection activity.
	Reflection upon math professional development, and the idea of eliciting student thinking to respond to student needs	Presentation connecting all of the math professional development from the previous meetings.
	Setting aim statement targets for upcoming year	School teams reviewed grades data for the previous three years and discussed.
	PDSA showcase	Teachers presented their change ideas and resources at stations, while other teachers rotated.
	Revisiting 5 Whys: Deepening our understanding of the problem	Facilitated 5 Whys activity with each school team.
	Revisiting Driver Diagram	School teams discussed and suggested revisions on poster-size driver diagrams.
	Reflection	Group reflection on the day and year.

Figure 9. May network convening content and activities.

Network's Outcome, End of Year 1

Unlike my previous experience with principals, I deemed this network a success. I began the 2017-2018 school year with a specific evaluation capacity building objective: We would build improvement science capacity, and embed processes within the schools. My hope was that by the end of the year, teachers would be able to:

- Understand and use a driver diagram.
- Develop aligned change ideas.
- Gather meaningful evidence.
- Engage in inquiry, dialogue, and reflection.

Network teachers demonstrated these improvement science skills and behaviors (although to varying degrees). Three out of the five schools established regular within-school meetings to work on network-related activities. We did have difficulty establishing consistent within-school meetings with Middleview and Marshall due to lack of leadership participation and support as previously discussed (with more in Chapter 6).

For our network as a whole, the evidence of successfully building these skills, behaviors, and processes was evident in the number of completed PDSA cycles, a measure of success I had chosen before the network began. The network had completed 63 PDSA cycles between February and May. Of those PDSA cycles, 97 percent included a change idea that was aligned to the network's theory of practice improvement (i.e., driver diagram.) All of the participating teachers, save one, completed at least one PDSA cycle. The number completed by each school is shown in Table 5.

Table 5

Number of PDSA Cycles Completed

School	Total Number Completed	Number of Teachers	Ave. Completed per Teacher
Roosevelt	22	6	3.67
Marshall	12	6	2.00
Middleview	7	4	1.75
Sawyer	13	4	3.25
Central	9	3	3.00

While I was primarily interested in building capacity and processes during the inaugural year, it became clear that the intended PDSA learning outcome was occurring too. Through their PDSA reflections, and observations and interviews, teachers exhibited meaningful and actionable learnings. This new knowledge was generated through the PDSA cycles, and by teachers engaging in inquiry and dialogue and sharing with their colleagues, both within their own schools and from other schools. These learnings are discussed more in Chapter 7.

Furthermore, we made progress helping the schools come together as a network. The five schools had a common aim and driver diagram. We had established network routines that fostered teachers engaging in inquiry and dialogue with teachers from other schools. Through these routines, teachers had the opportunity to hear experiences and perspectives outside of their own school teams, and learn new practices and strategies. Teachers also had access to all of the PDSA forms through our network website. Through our own observations and teacher feedback from network meetings, it was clear that teachers were enjoying the collaboration that occurred within the network.

Although the network demonstrated a positive end-of-year outcome, we still faced challenges. While an average number of PDSAs completed per teachers is reported in Table 5 for comparison purposes, there was variation within the number completed by teachers within in a school. For example, and most disparate, one teacher at Central completed seven PDSA cycles, while the other teachers completed one cycle, each. We also struggled to determine common network measures to monitor our improvements. Even though teachers were collecting meaningful data as part of their PDSA cycles, we still needed to develop practical measures aligned to the driver diagram in order to evaluate whether we were making progress towards our network aim. We will continue to work on these areas and try to improve in the next year.

The next three chapters further examine what facilitated our network schools, and the teachers within those schools, being prepared to successfully participate in the network. That is, what factors, structures and/or conditions were needed for schools and individuals to build the improvement science capacity and implement PDSA cycles? And, what factors, structures and/or conditions were needed to generate meaningful individual and school learnings? Chapters 5 and 6 impart findings regarding what helped the University build teachers' and schools' improvement science capacity. Chapter 7 unpacks the networks learnings that were generated, and what contributed to them.

CHAPTER 5

A TALE OF TWO VISIONS

Introduction

As mentioned previously, a core principle in improvement science is the inclusion of team members with subject knowledge and profound knowledge (how to improve) (Langley et al., 2009). Therefore, as a networked improvement community that was grounded in improvement science principles, this became a fundamental tenet of our network. This point cannot be overstated, as illustrated by what transpired in the early days of the network.

When we started the network, I knew nothing about math instruction and Tom knew little about improvement science and its underlying evaluation concepts. As the network lead, I was responsible for meeting with all our schools, in-between the network convenings. Tom was already responsible for providing math professional development to numerous schools, but the plan was for him to attend the majority of our school meetings. I presumed the teachers would be able to fill the math subject knowledge void if Tom was unavailable for any reason. My assumption was wrong. The teachers did possess subject knowledge to certain degrees, *but* not the type of math instructional knowledge needed to solve complex problems of practice.

This became obvious to me during a meeting with Sawyer's teachers in the fall. Tom was out sick that day, so I was the sole facilitator. Building upon the previous month's network convening, I was facilitating a process map regarding students' lack of prerequisite skills. The math teachers and I were in the school library, huddled around a large rectangular table. We mapped out the steps (processes) they currently take when students lack the necessary prerequisite skills for a lesson. When they began brainstorming an ideal process for addressing the problem, the teachers offered ideas including teaching a "Unit Zero" at the start of the school year, and teaching the missing skills during lesson warm-ups. Since we were already four months into the school year, I directed the conversation around the latter idea. The idea of teaching a lacking prerequisite skill, such as long division, during the warm-up sounded viable and reasonable to me because I lacked essential subject knowledge. I did not know any better. But Tom did.

Soon after learning Tom's full instructional vision while developing the driver diagram, I realized my deficiency. I not only realized the importance of subject knowledge, (particularly when facilitating a conversation such as the one at Sawyer,) I saw the need in a new light: It was not merely about content knowledge, but more importantly, about a vision for changing practice that possibly required an outside perspective. Teaching a skill during the lesson warm-up was a contributory learning that could improve an immediate problem, but it was unlikely to transform instructional practice or solve a complex problem. The teachers were brainstorming but only within their current, known instructional paradigm and system. *How would they know a different way to approach instruction, and why would we expect them to?*

The Math Instructional Vision

Tom envisioned another way of teaching. His instructional vision was grounded in the National Council of Teachers of Mathematics (NCTM) (2014) report, *Principles to Actions: Ensuring Mathematical Success for All*, and Cognitively Guided Instruction (CGI) principles (Carpenter, Fennema, Franke, Levi, & Empson, 1999). CGI is a teacher professional development program that focuses on eliciting, understanding, and building upon students' mathematical thinking (Carpenter, Fennema, Franke, Levi, & Empson, 2000). One of the principles is to build upon what students know and rely on an understanding of students' learning trajectories to make instructional decisions.

As mentioned in the previous chapter, *Principles to Actions: Ensuring Mathematical Success for All* (NCTM, 2014) played a prominent role in the development of our driver diagram. It provided a framework of eight mathematics teaching practices:

1. Establish mathematics goals to focus learning.
2. Implement tasks that promote reasoning and problem solving.
3. Use and connect mathematical representations.
4. Facilitate meaningful mathematical discourse.
5. Pose purposeful questions.
6. Build procedural fluency from conceptual understanding.
7. Support productive struggle in learning mathematics.
8. Elicit and use evidence of student thinking.

Tom's concept of quality math instruction guided the network. At the network convenings, he always kicked off the meeting with a number sense routine, such as "Which One Doesn't Belong," or choral counting. He also thoughtfully prepared problem examples that would apply to grades 6 through 9. He would project the problem on the room screen, and then have the teachers play the role of student. Nimbly moving around the room, Tom would call on them, probing for answers, and pushing them to explain their thinking. His favorite phrase was, "Turn and talk," promoting mathematical discourse among the teachers. Using poster paper, he demonstrated how to easily collect evidence of students' thinking by writing down what was being said, and who said it. Tom also provided opportunities for the teachers to ask questions and reflect upon how they would apply these ideas to their own students. All his exercises connected to the mathematical standards, and modeled quality instruction in line with his stance, which was projected on one of his slides:

> Effective teaching of mathematics engages students in solving and discussing tasks that promote mathematical reasoning and problem solving and allow multiple entry points and varied solution strategies. (NCTM, p. 12)

The teachers enjoyed this time, often raising their hands, and laughing at Tom's jokes. As one teacher exclaimed, "I'd do math all day if we could do this. This is fun!" Another teacher further explained the value in Tom's activities.

> *I think as a math student. Like the math student piece, I've kind of seen how ideas could be presented or taught. That's another thing that I've appreciated ... is the process that we go through to kind of breakdown how a brain would attack this.*

Philosophically, Tom's instructional viewpoint also connected to equity. When we first started visiting schools together, he often confided his concern when a teacher would offhandedly refer to some of their students as the "low" students or the "high" students. Based on his experiences, he suspected that some of those "low" students were also former or current English Language Learners. His purpose for introducing number sense routines was not only to elicit student thinking, but also to provide opportunities to shift teacher thinking. He wanted teachers to understand that there are different ways to engage those "low" students and build upon existing student understanding, rather than focus on their lacking mathematical skills. To this end, he mindfully shifted conversations during our in-school meetings to discuss PDSA results, commonly asking "What did you learn about your student's thinking?" and "How can you build upon your student's thinking?"

The Vision for Building Improvement Science Capacity

I also had a vision. As briefly discussed in the previous chapter, my aim for building improvement science capacity during Year 1 centered on teachers' learning processes and embedding them in the schools. I approached this undertaking with an evaluation capacity building lens, which can be succinctly defined as "an intentional process to increase individual motivation, knowledge, and skills, and to enhance a group or organization's ability to conduct or use evaluation" (Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012, p. 308). Evaluation capacity building could focus on all three individual areas: knowledge and understanding, skills and behaviors, and affective or attitudes. Or, I could concentrate on one or two areas (Preskill & Boyle, 2008). My primary focus was on teachers building particular skills and behaviors. By the end of the first year, I wanted teachers to demonstrate the following:

- Use a driver diagram to develop change ideas that were aligned with the network's theory of practice improvement.
- Gather meaningful evidence that could be used to evaluate and modify the change idea.
- Engage in disciplined inquiry with their colleagues using the PDSA format.

Improvement science consists of more components – understanding and conducting causal and systems analysis, developing a driver diagram, creating measures to monitor improvements towards the aim – but I differentiated between what I believed teachers could realistically accomplish, and what expertise I needed to provide as the improvement science specialist. To me, those additional components were too technical for those not trained in evaluation. While it was certainly possible to impart these next-level components to teachers, it did not seem practical. Schools are hectic and chaotic environments. Teachers already have a lot on their plates, on a daily basis. In my view, their foremost role in the network was to bring their teaching experiences to the shared problem of practice, experiment with ideas in the classroom, engage in inquiry and dialogue with their colleagues, and reflect upon how they can continuously improve. My role was to provide the evaluation expertise, which included facilitating the technical aspects of improvement science.

At this point, it is also important to distinguish the ideas of individual and organizational learning capacity. My goal was to build both. Individual capacity, as described above, relates to building participants knowledge, skills, and attitudes (Preskill & Boyle, 2008; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012). Organizational learning capacity, in the context of evaluation capacity building in schools, refers to whether leadership values learning and the evaluation process, and whether the school has a culture of inquiry, has the necessary systems and structures for engaging in the evaluation process, and offers opportunities to access and disseminate evaluation information (Preskill & Boyle, 2008; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012).

I envisioned developing individual skills and processes, and embedding them at an organizational level. This served two purposes: 1) building the skills mentioned above through consistency and practice, and 2) the transfer of learning to the workplace or mainstreaming to develop a sustainable learning culture in schools (Preskill & Boyle, 2008; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012). I wanted to build organizational learning habits, so this was my concern in Year 1 rather than actually improving the problem of practice. I recognized that this was an incremental route, and my goal was process-oriented. Again, I was playing the long game. We would work with schools, both on-site and at network convenings, so that they would embed processes to continually engage in improvement and evaluative thinking, even after improving our identified problem of practice. Improvement would not be an “add-on project,” but would become a way of working (Rother, 2010).

Initial Improvement Science Capacity

To gain a sense of the network's initial capacity regarding improvement science, I conducted a survey at the end of our first convening in September. Tables 6 and 7 show the results for the capacity-related questions for the five case study schools. For the most part, participants (other than Roosevelt) were not familiar with improvement science concepts and tools. Notably, more than half of all participants (14) indicated that they were at least somewhat familiar with process maps, even though later work with the teachers demonstrated that they were unfamiliar with an *improvement science* process map. In this case, the disconnect likely reflected an instance of teachers not knowing what they do not know. "Process map" is a seemingly generic term, but has a specific meaning and application in improvement science. In this context, a process map is more technical, and includes start and end points, decision points, and open and closed loops. It was clear when working with teachers that they were not familiar with these more technical aspects, even though they understood the basic concept of how one process flowed to another.

Table 6

Teacher and Administrators Familiarity with Improvement Science Concepts

Before today, how familiar or unfamiliar were you with the following?	I don't know what this is	Not at all familiar	A little familiar	Somewhat familiar	Familiar	Extremely familiar	n
Improvement Science	0	10	2	2	5	4	23
Root cause analysis (e.g. 5 whys)	0	9	2	2	5	5	23
Driver diagrams	1	10	3	4	3	2	23
Process maps	0	4	5	5	5	4	23
PDSA cycles	1	7	2	4	3	6	23
Systems thinking	1	7	1	7	4	3	23

Note. Survey of network participants from the first network convening, September 2017.

Findings of actual unawareness of the more technical aspects of improvement activities existed elsewhere, too. Almost all teachers reported that they were somewhat confident to confident that they could engage in improvement-oriented activities from the outset (Table 7). Yet, it became clear while working with them that they were, in fact, unaware of specific processes for engaging in these undertakings. For example, as demonstrated by Table 6 above, most teachers were not familiar with root cause analysis, even though all of them were at least somewhat confident in their ability to identify barriers to student learning in math. This suggests that teachers may perceive they do not need to build the capacity, despite observations of skills and behaviors that demonstrated otherwise.

Table 7

Teacher Confidence in Improvement-Oriented Capacity

How confident or not confident are you that you can:	Not at all confident	A little confident	Somewhat confident	Confident	Extremely confident	n
Identify potential barriers to students' learning math/alg. concepts?	0	0	8	9	4	21
Make changes to your teaching or classroom practices that could improve student learning in math/alg.?	0	0	4	12	4	20
Test whether a change in your teaching or classroom practices resulted in an improvement of your practice?	0	3	7	8	3	21
Collect meaningful data to evaluate whether a change in your teaching or classroom practices resulted in an improvement of your practice?	0	1	6	12	2	21
Analyze data to evaluate whether a change in your teaching or classroom practices resulted in an improvement of your practice?	1	1	7	9	3	21

Note. Survey of network participants from the first network convening, September 2017. Teachers only.

The challenge of building improvement science capacity among teachers was not merely in helping them value improvement, because most of our network already did. On the first survey, 15 out of the 21 teachers indicated that it was extremely useful to experiment with new teaching practices, when it came to their own teaching. Rather, the challenge here was in demonstrating that the tools and evaluative processes for improving are valuable, too. This proved more difficult.

Throughout this first year, more than one teacher initially resisted conducting the PDSA cycle. They insisted that they already did the same process in their practices (without formally documenting it.) While we did not challenge whether this was true, we did respond by reminding them that the network was a collaborative endeavor. Its purpose was not only about individual learning, but documenting their experimentation and reflection provided an opportunity for others to learn from them. They never argued this point, which suggested that our teachers valued collaboration with their colleagues.

This notion was apparent and reinforced throughout the entire school year. From the beginning, the teachers indicated that it was useful to collaborate with other teachers to reflect on improving practices, with 17 out of 21 teachers on the survey responding that it was extremely useful. Feedback from network convenings frequently showed that teachers appreciated the time to work with department teams, and to talk with teachers from other schools. Interviewed teachers also expressed this sentiment. When asked what stood out for them about the network, the most common answer was collaboration with other teachers.

The Two Visions Intersect

Our team found one way to engage teachers in learning improvement science skills, even if they were not interested in the tools and processes, was by learning in context. While this was not an overt strategy when we started this network, it became apparent that teachers were more motivated to improve math instructional practice, rather than their ability to improve (i.e., improve their ability to improve). Put another way, most teachers did not demonstrate a desire to get better at driver diagrams or PDSAs cycles. They expressed a desire to be better at their instructional practices to help their students. One teacher expressed this sentiment by comparing the early large network meetings that did not include math instructional professional development to the later small network meetings that did. She explained:

> *Yeah, I guess the one negative thing from the beginning of the year, it was just the [large network] days that were like, full days, and not feeling like, it was moving very fast....But more recent things of, you know, seeing and working with more specific ways of adjusting to meet our students' needs. That has felt more helpful. And I know*

that there had to be, there had to be big picture conversation of like, "What's actually going on? Let's not just like, fix a problem the way that all the books say we should fix the problem." Um, it just felt like, it took a little while to get authentic.

This teacher was communicating that building evaluation capacity (problem-solving) in the context of adjusting to their students' needs felt more authentic than learning improvement science in the abstract with an undefined problem of practice. In doing so, she acknowledged the value of placing learning in the day-to-day teacher context, not just an educational context.

Although we had not formally drawn upon situated cognition learning theory during Year 1, it is apparent upon reflection that we were subscribing to its principles (Merriam & Bierema, 2014; Brown, Collins, & Duguid, 1989). Over the years, numerous people have told me they find improvement science to be overly abstract and "jargon-y." Situating the improvement science language and concepts into the everyday work of the teacher's classroom provided a way for teachers to construct meaning of these abstract concepts.

This constructing of meaning was evident at our April network convening. Teachers were asked to explain why they believed their latest change idea would improve the network's global aim (i.e., engage all students and meet their variety of learning needs). We requested they use the driver diagram to illustrate their thoughts. One-by-one, teachers from each of the schools stood up at their tables, while other team members held the driver diagram. They articulated how their idea was connected to a particular secondary driver, then connected to one of the primary drivers, which would then improve student learning. We (University hub) were thrilled, if not a little astonished, that they had effortlessly described their theory of improvement *and* used the language of improvement science (i.e., change idea, secondary driver, primary driver, aim). This successful moment was the result of the two visions intersecting.

Another instance of the two parallel learning tracks – building improvement science skills, and improving math instruction to engage all students and meet their variety of learning needs – fruitfully intersecting was the PDSA cycle. At the January network convening, we introduced Tom's number sense routines (e.g., Which One Doesn't Belong), and the PDSA process. We were deliberate about teaching them both in the same meeting. Although we did not specify that teachers should try a number sense routine as their change idea, the routines modeled the type of activity they should be experimenting with in their PDSA. For example, Which One Doesn't Belong was:

- Quick to implement: Teachers could experiment with it the next week without requiring a lot of planning time. It also only required about 15 – 20 minutes of class time. Teachers could complete a full PDSA cycle within a week since it could be tested during one class.
- Manageable: Teachers only needed slides with the four images or mathematical terms (which did require some planning time). Tom modeled how to lead the activity. He also demonstrated how to easily collect data for the PDSA cycle by writing down what was said, and who said it, during the activity. This was practical data collection: It informed instruction and assessed the level of student engagement in the activity.
- Meaningful: The activity was aligned to the network's aim – engaging all students in mathematical activities. It could also initiate meaningful reflection as to which students participated in the activity and why, and what assumptions teachers made.

Most teachers chose the Which One Doesn't Belong activity for their first PDSA, and we were amenable with that choice. Not only was the activity in line with Tom's vision, it also helped to scaffold learning the PDSA process. We felt this was crucial. Teachers were not charged with identifying an instructionally appropriate change idea to try in their classroom, and mastering how to implement it quickly, *and* learning a whole new process for evaluating that change. Our priority was for them to learn this process, rather than innovate change ideas.

The PDSA structure facilitated the math instructional learning, too. Essentially, the PDSA cycle is a form of adult experiential learning (Langley et al., 2009; Merriam & Bierema, 2014). Experiential learning refers to learning through life or workplace experiences, and the reflection upon those experiences (Merriam & Bierema, 2014; Dewey, 1963). Because many teachers opted to try Tom's number sense routines, the PDSA cycle provided a structure for putting that professional development into practice. The process itself required that teachers experience the number sense routine in practice (Do), and then reflect upon that experience (Study).

One teacher illustrated how PDSA cycles contributed to her instructional learning. In this example, the "tool" she is testing through the PDSA process is the Which One Doesn't Belong activity:

> *Every single time there [were] at least two things going on, it would be the actual tool that I was testing. And so, maybe I was testing Which One Doesn't Belong. But not only was I testing the tool, I was also testing the concept I was putting in the tool. And that's kind of why I was saying it'd be nice to have some examples, because sometimes if I didn't choose a great concept to place in the tool, then it would die. But then, if I did it in a way that made sense to the kids and made sense to what we were doing, like, the tool worked. So I think that's what it is. It made me think how to best use the tool. And if something didn't work, it was usually not the tool. It was usually what I put in it.*

From this example, it is also clear how the process not only led the teacher to consider the tool or activity, but also reconsider her decisions about how to use the tool. As discussed in Chapter 7, learning can also be transformative when the experience and reflection leads to the questioning of assumptions, and reconsidering how instruction occurs, rather than merely incremental improvement to the current instructional practices (Merriam & Bierema, 2014; Argyris & Schön, 1996).

Conclusion

Tom's notion of quality mathematical instruction and my plan for building improvement science capacity were more than complementary, they were synergistic. Even though we had not explicitly discussed how the two visions were intertwined during the first year, they guided our decisions every step of the way. His mathematical instruction expertise advanced our ability to build improvement science skills by providing context and an incentive to build evaluative capacity. This chapter illuminates the significance of establishing visions in both the subject area and improvement science capacity, and for the experts in both areas to deliberately work together. After walking out of a PDSA meeting at one of our schools, Tom and I joked that we were "in each other's head." Throughout the meeting, I had asked teachers "What did you learn about your students thinking?" and he had asked "What is your evidence?"

CHAPTER 6

THE POTENTIAL IRONY OF IMPROVEMENT SCIENCE IN EDUCATION

Introduction

According to Langley and his colleagues (2009), there is a Central Law of Improvement that states, "Every system is perfectly designed to deliver the results it produces" (p.79). This phenomenon makes practical sense. Results can only be as rigorous as the system. If you do not set up structures to do the work, the work will not occur. Herein lies the potential irony of improvement science in education. In my experience, I have seen and participated in improvement science initiatives where educators from multiple schools were assembled, taught improvement science concepts and tools, urged to apply them at their schools, and then released. The folks in charge of these convenings (including me) sent educators off without attending to the work structures in their schools, instead hoping with fingers crossed, that teachers and administrators would make the time and effort to follow through on their improvement science promises.

For me, the most obvious example of the unkept promise is the PDSA cycle. While learning about networked improvement communities, I attended a network convening of another network, which had been in existence for a couple years. During the meeting, I noticed there was no mention of the PDSA cycle. I asked the organizer about it, and she indicated they struggled to implement them. I then asked whether they worked with teachers at their schools. She replied they did not. This was reminiscent of my previous experience where it was challenging for principals to implement regular PDSA cycles. We hoped to inspire, motivate, and build capacity at central meetings, with little attention paid to their existing work structures back at the schools. In that case, we did not want to dictate how principals applied the process at their sites. We felt that was their purview, not ours. However, schools are hectic environments where teachers' and administrators' time and energy are forever being pulled in different directions. The project failed to gain traction, and resulted in no actionable learnings. Our system proved it was perfectly designed to achieve the insufficient results that it did.

Organizational Capacity

The previous chapter briefly discussed the purpose of embedding improvement science processes within schools as part of individual and organizational evaluation capacity building goals. This chapter expands upon the findings related to those ideas. Working regularly within schools provides teachers the opportunity to individually develop improvement science skills through consistency and practice. It also transfers the learning to the workplace, and builds organizational learning habits. Sustaining these habits requires the "development of systems, processes, policies, and plans that help embed evaluation work into the way the organization" works (Preskill & Boyle, 2008, p. 444).

In-School Work Structures

Establishing time and space. Our ability to regularly meet with schools and build systems and structures was related to their pre-existing organizational capacity, particularly in the form of leadership. From the outset, our University team shared our expectation of holding regular, on-site meetings with the teachers and administrators, (except in the case at Middleview.) But working at the schools required time and space, both contingent on the commitment of administrator leadership to provide them. That was one of the most significant roles administrators played.

In Year 1, we encountered differences among the schools' leadership that impacted our ability to build in-school work structures. I found this to be connected to the teachers' ability to develop their individual improvement science capacity. Table 8 shows how the number of in-school meetings corresponds with the number of completed PDSA cycles by school.

For example, the principal at Central exhibited a high level of commitment by carving out time during the school day for his three teachers, every two weeks. Teachers met during a conference period, even though they all had different class schedules and free periods. The principal found coverage for the teachers and alternated periods to lessen his teachers' burden. The administrators at Sawyer also prioritized our meeting time with schedule accommodations. Teachers were released early, one day a week for professional development time. That gave us that time to work with their teachers at least once a month.

The Roosevelt administrator also supported our work with her teachers, but their situation was unique. Because they already had improvement science experience and meeting structures for discussing PDSAs cycles we did not organize their within-school work structures. Instead, we attended their existing meetings as a participant, rather than a facilitator, so that we did not interfere with their existing processes. However, there was a facilitator (their lead teacher) and I did interject at times to coach the correct use of improvement science tools and methods.

Table 8

Number of PDSA Cycles Completed and In-School Meetings

School	Total Number Completed	Ave. Completed per Teacher	Number of Times We Met with Them
Roosevelt	22	3.67	5
Marshall	12	2.00	4
Middleview	7	1.75	1
Sawyer	13	3.25	8
Central	9	3.00	8

Note. Roosevelt already had pre-existing improvement science capacity and PDSA meeting structures.

As discussed previously in Chapter 4, we had difficulty scheduling regular meetings at Marshall. Their administrator wanted more flexibility with the process than other principals, and provided little assistance with Samantha's request to establish a set time to work with his teachers. Only through Samantha's persistence we were able to schedule meetings, but they were inconsistent and never became part of a regular schedule. The absence of a set meeting time adversely impacted the number of times that we met with them (Table 8).

Middleview was another exception, but only because we did not place expectations upon them due to our own assumptions regarding their ability to commit time and energy (as discussed in Chapter 4). We did not meet with the principal at the beginning of the year, thus, he never committed to carving out time and space for us. Once Middleview built their math team by mid-year, we attempted to organize time directly with the teachers. They were already having weekly, after-school meetings for another continuous improvement process championed by the district, so they tried to append our network to that recurring meeting. We attempted this several times, and each time, we were only given an average of 15 to 20 minutes, which proved to be insufficient. It was a challenge to make our brief time together meaningfully productive. We did, however, meet with them for one two-hour session as a "make up" for missing a network convening.

Facilitating collaboration conducive to improvement. Providing time and space for collaboration, while a crucial first step, does not ensure that the collaboration will be productive and/or promote learnings needed to solve complex problems of practice. At the very least, collaboration time should have a commonality of purpose and promote dialogue and discussion. At its best, collaboration advances team learning through the alignment of its members, whereby insights are generated and put into action (Senge, 2006). Through the initial fall survey, network members were asked what routines, meetings, or other structures did their school already have in place for teachers to collaborate towards improving their teaching practices. Almost all teachers (90%, 19 out of 21) indicated that they had collaboration time through some type of regular meeting (e.g., department meeting), common conference period, and/or professional development time. However, in that same survey, only about half of the teachers and administrators (48%, 10 out of 21) indicated that they met at least monthly, on average, with other teachers to discuss what helps students learn best. Additionally, less than a third of network teachers (28%, 5 out of 18) said that they asked another teacher for help or feedback on their own teaching practices at least once a month, on average. These data suggest that although teachers already met regularly, much their time was spent working together on topics other than improving student learning and/or their own practices.

Furthermore, I found that when they did meet regarding how to improve student learning and instructional practices, the structures were not always conducive to dialogue, discussion, and/or inquiry. An observation at Roosevelt illustrates this example. I attended one of their math department meetings where they discussed PDSA cycles. The lead teacher facilitated this meeting. She was extremely structured. Teachers were given a set amount of time (e.g., 10 minutes) to share what they tried in their classroom, and their results. Teachers sat in a circle and took turns speaking. After a while, I noticed that the teachers were not asking each other questions. I asked a few questions in an attempt to start a dialogue, but the lead teacher requested that we hold our questions until the end, after everyone had shared. I surmised that she was extremely concerned about keeping to the schedule and not going over the meeting time allotted. I was acutely aware that teachers' time is a finite commodity, so this was understandable, but it was also a conundrum that thwarted the ability to engage in productive inquiry and dialogue. By the end of that meeting, very little teacher-to-teacher questioning had occurred, rather it was teachers individually sharing before moving onto the next person.

Because of this and other similar experiences, we taught the teachers how to engage in inquiry and dialogue (as discussed in Chapter 4) in an effort to cultivate team learning. We encouraged them to use the protocol that we developed for network PDSA sessions in their in-school discussions, too. Our hope was not that they use a protocol every time they engaged in dialogue and inquiry, but would eventually develop a routine for asking meaningful questions, i.e., to push each other's thinking. While teachers did not always follow the protocol during school PDSA meetings, I did observe an increase in questioning, particularly at Roosevelt.

On the opposite end of the structure, I observed how collaboration time can be too unstructured to promote constructive dialogue. As a teacher from another school shared: "It's always a struggle to collaborate without turning it into a bitch session…I think we have a pretty positive math department, but at the same time, it does seem like, given time, that's the path we go down."

One of the benefits of improvement science is the provision for structured collaboration around a common purpose. Tools such as a fishbone diagram or a process map, organize conversations around potential causes of a problem. The driver diagram categorizes thinking around how to improve a problem. The PDSA cycle is a structured format for experimenting and reflecting upon results. These tools are helpful for guiding conversations, and keeping them on track. As one administrator stated, "I've appreciated just the different systems you've gone through to kind of organize teachers and their thinking, from an outside perspective." Because these tools are technical, an improvement science expert is beneficial to facilitate them, but an important secondary role became keeping the resulting discussions focused.

The benefit of an outside facilitator. Our facilitation at in-school meetings served multiple purposes: supporting the use of the improvement science tools, supporting productive collaboration, and providing an outside perspective. As mentioned, a key component of our improvement science capacity building vision was the need for an improvement science specialist to facilitate these meetings. (Roosevelt was the exception because they already had existing capacity and desire to lead these meetings.)

In interviews, I asked teachers and administrators whether they would have been comfortable facilitating the in-school meetings, when we developed fishbone diagrams and early pieces of the driver diagrams. Many teachers acknowledged that they did not have the capacity, time, or desire to lead those early meetings. One teacher said it would be like “the blind leading the blind.” Another teacher acknowledged the two issues of capacity and burdening a teacher with the responsibility.

> *I don't know if we know enough about it. We've been through it. I don't think we have time, to be perfectly honest. I mean we're running all the time. So that would be an extra burden for someone to go back and research and review, take the time to do this, and just to set it up.*

Even a teacher who felt that she had a solid understanding of the tools indicated that she would not want add another responsibility to her plate.

> *So, in that sense, it sounds like another responsibility. So, could I show up to a meeting and roll something out? Probably, but if I was supposed to be really thinking about it away from that time, I don't think that's really realistic.*

We also conducted a survey at the end of Year 1 to gauge teachers’ ability to lead and facilitate meetings, after learning improvement science concepts. The results are shown in Table 9. Many teachers felt that they could apply improvement ideas and tools without the help of a facilitator. However, what is especially notable here, is that almost none of the teachers (including many of the Roosevelt teachers) indicated that they could teach the concepts in order to facilitate other teachers within their school teams. The importance of this distinction cannot be overstated. Taken together with the interview results, a common notion arises. For the most part, teachers did not feel they have the capacity to teach and facilitate other teachers around the use of improvement science tools, nor did they want the responsibility. Even those who perceived that they have the capacity, acknowledged that it is more efficient and less burdensome to have an outside expert lead the meeting. This idea is further supported by other networks who have found an improvement science facilitator to be beneficial as well (Proger, Bhatt, Cirks, & Gurke, 2017).

Table 9

Teacher Self Ratings of Knowledge and Skills of Improvement Science Concepts, Terms, and Tools

How would you rate your level of knowledge and skills for the following concepts, terms, or tools?	I don't know what this is	I can recall this concept, term, or tool	I understand this and how it connects to our work	I can apply this with assistance from Univ. facilitator	I can apply this without assistance from Univ. facilitator	I can teach this to facilitate and lead our school team	n
Conducting root cause analysis (e.g., Fishbone Diagram, 5 Whys)	0	0	0	8	8	1	17
Using a process map	1	0	2	8	6	0	17
Developing a driver diagram	0	0	1	6	9	0	16
Developing a change idea aligned to the driver diagram	0	0	1	3	10	2	16
Conducting a Plan-Do-Study-Act (PDSA) cycle	0	0	0	3	12	2	17
Collecting PDSA data/evidence to inform whether the change idea resulted in an improvement	0	1	1	0	13	2	17
Developing meaningful process measures	1	1	2	3	9	1	17
Developing meaningful outcome measures	1	1	2	3	8	2	17

Note. Survey of network participants from the last network convening. May 2017.

Furthermore, the presence of network facilitators at the in-school meetings helped teachers feel supported throughout this process. As one teacher enthusiastically described, *That's one of the things I've been, like, telling my colleagues, especially the ones coming into Algebra 1. … I'm like, "Dude, they come and work with us during the meetings to prepare us for our next meeting," which is awesome because then it becomes productive time, you know?*

Facilitating the meetings also provided an opportunity for an outside perspective during these activities. Not only did it provide an instructional perspective outside of their current paradigm, as discussed in Chapter 5, but also it provided a more objective viewpoint that assisted in propelling the work forward. One teacher highlighted an example, "I mean, I think at one point in one of those, we had, like, four different Post-its for the same thing. And you guys were able to guide this. 'Maybe you guys could put this together.' 'Oh, cool!'"

More collaboration around instructional practices. The evidence suggests that teachers were engaging in more improvement-oriented collaboration and dialogue around their practices by the end of Year 1. The end-of-year survey indicated that a greater proportion of teachers were partaking in collaboration with other teachers around improving student learning, and their own teacher practices. Almost three out of four (71%, 12 out of 17) said that they met at least monthly, on average, with other teachers to discuss what helps students learn best. Now, more than half of the teachers (57%, 8 out of 14) indicated that they asked another teacher for help or feedback on their own teaching practices at least once a month on average.[4]

These results, combined with observation and interview data, suggest that regular in-school meetings structured around improvement activities such as PDSA cycles were leading to more discussions about student learning and instructional practices. As Tom noticed, the frequency with which teachers were meeting and reflecting was shifting conversations and thinking more towards how best to engage students in mathematical learning. He noted:

---

[4] Caution should be used if directly comparing to results of the first survey. Not all of the participants were the same in both surveys.

> *What's also really, really structured is the consistency by which teachers are given time to think about their practice, and come to the table to reflect about their practice, and to share some of the things that they've come to learn. That, I think, has been paramount to shifting the needle a little bit for some of the teachers with regards to their instruction and thinking about how the kids are engaging with mathematics. And I know that because I previously worked with folks at Middleview and Sawyer, and I know that in my time, definitely at Sawyer, it had been really challenging to get momentum, to build momentum. And I feel like the consistency by which we are meeting with them has been very much conducive to them, again, thinking about their practice, thinking about how they've been engaging kids, or how their students are engaging with the mathematics.*

Consistency

In the *Toyota Kata*, Rother (2010) explains that a continually improving organization is one that builds thinking and behavior patterns and systematic routines for adapting and improving. When employees practice these patterns and routines as part of their day-to-day work, they have the means and habits for improving desired outcomes. Rother (2010) acknowledges that the challenge is developing and maintaining these routines.

As with building any new behavior, I found that consistency and practice are key. By regularly working with the schools, we were not only trying to build improvement science skills, we were also trying to develop habits and routines (individual and organizational). During interviews, many teachers expressed that they learned some process, such as the driver diagram, by continually using it, or revisiting the tool. Another example is the protocol that we developed to assist teachers with engaging in productive inquiry and dialogue, as discussed previously. Meeting observations indicated that teachers were becoming more comfortable asking probing questions, and sharing their thoughts. While some of this comfort might have been related to practice using tools and the protocol, it could also be due to simply spending more time with their colleagues in this forum. Either way, it was the consistency of these meetings that contributed to behavior changes.

Regular meetings with schools also helped the University support network schools. Because we met with the teachers frequently, we built relationships with them, which then contributed to them sharing honest feedback with us. As discussed in earlier chapters, some teachers pushed back on conducting PDSA cycles during state testing or documenting reflections, or were upfront about why they did not complete the run chart pilot activity. We had these conversations when meeting with schools between network convenings, which allowed us to acknowledge and respond to existing challenges when planning the next network meeting. Even when concerns were not explicitly stated to us, we detected potential issues during our attendance. For example, as illustrated previously, it was during school meetings that I discerned that teachers at some schools were not asking each other questions, or time constraints (e.g., a 60 minute department meeting) were limiting productive inquiry and dialogue.

Rother (2010) offered a definition of continuous improvement that we, the hub, pursued in our work of leading the network that instructs "moving towards a desired state through an unclear territory by being sensitive to and responding to actual conditions on the ground" (p. 43). It was only through regular meetings with teachers that we were aware of actual conditions on the ground. Only then could we strive to continually improve our own processes for preparing schools to successfully participate in the network.

Consistency was beneficial to the network in other ways, too. It was through in-school meetings that we were able to develop a driver diagram that reflected each school's concerns (described in Chapter 4). We heard their discussions and compiled their causal analysis artifacts. Rather than trying to bring five schools together to reach agreement, we more efficiently served as the consensus makers.

Individual versus Organizational Capacity

The importance of building organizational capacity and consistent work structures cannot be overstated. However, while embedding work structures and systems in the schools contributed to teachers' ability to implement PDSA cycles and successfully participate in the network, their pre-existing individual capacity was a factor, too. This point is demonstrated by the variation in PDSA cycle completion among teachers within the same school.

There were noticeable differences between the teachers who completed the most PDSA cycles within their school, and those who completed the fewest. Interviews with teachers who completed fewer PDSA cycles suggested that they were more concerned with classroom management and discipline than the other teachers. These teachers also tended to be more teacher-directed in their instructional style. That is, they preferred leading the instruction in the classrooms rather than students directing their own learning through collaborative groups. On the other hand, teachers who completed more PDSA cycles reported that they were comfortable experimenting with new ideas in their classroom, and already *informally* reflected upon their lessons, in order to tweak and improve them. Thus, these teachers possessed some existing capacity similar to what was needed for successful PDSA cycles.

Network Structures

While this chapter emphasizes school organizational structures, it is important to note that network convening structures also contributed to the successful completion of PDSA cycles. For example, we ensured the consistent use of the driver diagram by incorporating it into meeting activities. These meetings (January to May convenings) were intentionally structured to provide improvement science and math instructional professional development and/or activities in the morning, and school teamwork time in the afternoon. It was during this afternoon time that teachers would plan their next PDSA cycle. As the year progressed, we began prioritizing this PDSA work time over other convening activities (if we did not have time for both.) We also started giving them time to complete their individual reflection section of the PSDA form. Through conversations with teachers, it became clear that teachers had difficulty finding the time to reflect and document their thoughts and learning after experimenting in their classroom. When teachers indicated that they were overwhelmed during state testing time, our offer to provide this extra work time assuaged some of their scheduling concerns.

Conclusion

These findings suggest that any hope of developing improvement science capacity of educators, teachers in particularly, first requires cooperative in-school work structures. Network leaders substantially benefit from the support of school administrators because they provide the key ingredients: time, space, and necessary supports for teachers. Without these, it is unrealistic to expect teachers to each find the time to prepare, and regularly gather in a room with other teachers to facilitate improvement science activities on their own. And to expect them to do so, without any pre-existing knowledge and skills, is impractical.

As with building and embedding any new behaviors in an organization, I found that consistency was as crucial as leadership support. Regular and frequent meetings with schools also benefited the hub by providing ongoing opportunities to assess challenges as they arise, acknowledge teacher perspectives and address their concerns. While this chapter highlights the benefits of establishing structures, it is also important to note that embedding these routines within schools is challenging, and it is still a work in progress for our network.

CHAPTER 7

NOT ALL LEARNING IS CREATED EQUAL

Introduction

The previous chapters discussed the network hub's goal of building and embedding improvement science skills by the end of Year 1, as demonstrated by the completion of PDSA cycles. However, PDSA cycles are only a means to an end. Ultimately, the purpose of improvement science (or any continuous learning process) within networked improvement communities is to generate collective learnings that can improve complex problems of practice. Therefore, it was also important to consider what learnings were generated during Year 1, and how they were generated. This chapter describes what type of learning occurred through the network, and discusses what contributed to those learnings.

Argyris and Schön's (1996) single-loop and double-loop learning provide a useful way for conceptualizing learning in networked improvement communities. Single-loop learnings are instrumental learnings that lead to improved performance without changing underlying values, norms, or strategies regarding current practices (Argyris and Schön, 1996). Double-loop learnings question underlying values, norms, or strategies, ultimately leading to changes in how and/or why certain practices are being done (Argyris and Schön, 1996). Both of these types of learning are valuable. Single-loop learning is useful for improving upon current practices while double-loop learning is helpful for reconsidering and transforming practice. Because networked improvement communities are predicated on the idea of solving complex problems of practice through collective learning, and these solutions may require new ways of thinking or behaving that go beyond merely improving current practices, it is important to also look to double-loop learning as a potential goal of networked improvement communities.

Consider the meeting with Sawyer's teachers discussed in Chapter 5. The teachers were working on the problem of students lacking a prerequisite skill, such as long division, necessary for a lesson. During their brainstorming, they devised a solution around teaching prerequisite skills during a warm-up, before the lesson. This type of learning through the collaborative dialogue of the teachers is an example of single-loop learning. However, while teaching long division during a warm-up might correct the issue, it does not address the fundamental problem: incoming students do not know long division. Therefore, the complex problem of practice – a lack of prerequisite skills –will likely persist. In this example, double-loop learning could result from inquiry around how they teach long division in the math classes, and questions about their own assumptions regarding how their students learn. Teachers would reconsider their own pedagogy and be open to new instructional approaches.

Double loop-learning occurs when the inquiry leads to changes in individual or organizational values of "theories-in-use," defined as patterns that are implicit in individual or organizational behaviors (Argyris and Schön, 1996). Theories-in-use can be compared to "espoused theory" which represents the strategies and values that individuals or organizations advance to explain their actions (Argyris and Schön, 1996). Fundamentally, espoused theory represents the notion of "what we say we do" which can be compared to theories-in-use that signifies "what we actually do" (Senge, 2006).

The ideas of espoused theory and theories-in-use are especially germane to the challenges faced in education. In my experience working with teachers, I have seen espoused theories and theories-in-use in conflict, sometimes unknowingly to the teacher. For example, in our network, one teacher espoused her beliefs that all of her students were capable of learning, and she seemed to truly believe it. Yet when she spoke about her students in the meetings, she often referred to them as the "low" students and the "high" students; thus unwittingly labeling students' ability to learn. In other instances, teachers were more overt about this conflict. They would agree with the school's espoused theory during meetings – that all students can learn math – but also blame their students for a lack of engagement in math, rather than acknowledging their own role. Double-loop learning occurs when teachers become aware of these conflicting values and assumptions, and resolve them in way that changes their underlying theories-of-use.

These ideas of single-loop and double-loop learning, and espoused theory and theories-in-use came to guide many of our network inquiry activities. While we wanted to foster both single- and double-loop learning, we realized that transforming teachers' instructional practices would absolutely require double-loop learning. Single-loop learning on its own would not be enough. Teachers needed to reconsider how and why they taught certain ways, and not just incrementally improve strategies that fit into their day-to-day paradigm of teaching. As we recognized the need for double-loop learning to solve our complex problem of practice – our current practices are not aligned with students' learning needs today – we attempted to foster it through inquiry structures and activities.

Learnings Generated During Year 1

The PDSA forms provided evidence of what teachers learned throughout the year. Combining these data with interviews, observations, and the end-of-year survey painted a picture of the type of learning that occurred during Year 1. Incremental learning occurred through the PDSA cycle: Teachers chose a change idea, such as Which One Doesn't Belong, made a prediction, collected data, assessed whether they met the prediction, and through reflection and inquiry, corrected the change idea until it achieved the results that they desired. This initial type of learning, one PDSA at a time, was more consistent with single-loop learning. However, through the course of math instructional professional development, PDSA cycles, inquiry and dialogue, and opportunities for reflection, it appeared that some double-loop learning was occurring, as well. In these cases, teachers appeared to be shifting their underlying values and assumptions about how their students learn math.

Illustrations of Single-Loop Learning

As part of the PDSA cycles, teachers were asked to reflect upon their experimentation of their change idea, including their results, what they learned, and what they will do next. For the purposes of this research, single-loop learnings would be knowledge generated that improved upon a practice or strategy, but did not alter underlying assumptions, values, and norms of teachers' practices. I assumed that teachers conducted their PDSA cycles within their existing instructional paradigm because the change ideas were inserted into their current ways of teaching, i.e., adding Which One Doesn't Belong as a warm-up activity. (As described more in the next section, multiple iterations of single-loop learning could lead to double-loop learning when teachers begin to question current practices through the knowledge gained in PDSA cycles, although it is not a necessary precursor.)

Not all teachers provided evidence of single-loop learning in their PDSA forms such as identifying actionable next steps to improve upon a strategy, or completing multiple iterations that improved upon a strategy or current practice. In these cases, for the purposes of this research, I would not classify single-loop learning as occurring because there was no evidence they generated learnings that could improve their instructional performance (regardless of whether it leads to double-loop learning.) *However*, it is quite possible that learnings were still generated and acted upon, but not documented. Of the 23 teachers, 16 demonstrated single-loop learnings through their PDSA forms. Differences among teachers tended to be related to the individual and organizational capacity (or lack thereof) needed to complete PDSA cycles, as discussed in the previous chapters.

There were many instances of single-loop learning in the network. Figures 10 and 11 provide examples that exemplify how this learning occurred through the PDSA process. In

Figure 10, the algebra teacher was building upon her first PDSA cycle (not shown) where she tried the Which One Doesn't Below activity during one of her periods. During that class, she incorporated the task as part of her lesson on two-variable association in scatterplots. She predicted that most of her students would engage in the discussion by raising their hands and sharing answers with the whole class. She found that even though more students were participating, compared to previous lessons, there was still unbalanced participation: Some students were not engaged at all, while other students were repeatedly contributing.

Through the second PDSA iteration, Figure 10, she improved upon the first cycle by having students within teams write down their answers, and then share by team. This time, she tried the activity in all of her algebra periods. Using the data she gathered in the first PSDA cycle, she made incremental adjustments regarding opportunities for students to engage in the activity; volunteering as part of the whole group discussion, versus writing first and sharing by team. She also paired the activity more closely with the lesson in this iteration. She indicated that she wanted to continue finding ways to incorporate as a landmark before/after the lesson.

This teacher's experience serves as illustration of valuable single-loop learning. She used data from the first PDSA cycle to modify her approach for providing students opportunities to engage in the activity. Every student engaged this time, at least through writing, and each team shared. She "corrected" the problem of unbalanced participation among her students. The learnings themselves still fit within her current instructional paradigm. It was an add-on to her existing lesson rather than altering her underlying values and norms of instruction. In her reflection though, she does express a desire to incorporate more, but at this point, she is only considering incorporating as an activity before/after her usual lessons.

PDSA Cycle Form

Change Idea	**Describe:** Which One Doesn't Belong for Exponential Functions **What is the change:** Building on the WODB structure from last time: improve wait time, all students thinking/writing, have teams share in turns before the same people can share again, & pre-select some student responses.
Predictions: What improvement do we think will happen?	Like last time, students will have fun and some will share who usually don't. This time, I aim to get even broader participation and more students' voices. I also want to highlight when students build on something a classmate shared.
Questions: What do we want to learn from this cycle?	Can we get 100% of our students engaged in thinking, writing, sharing (pair or to class), and listening to / restating / building on one another's ideas? Can they do it before any lessons on the topic? Does it help launch the topic?
Data: What information will we collect to answer our questions and test our prediction?	Students' responses both on their paper and shared out to the class poster Seating chart tally as a record of who in each team is writing, pair sharing, sharing to class, etc.
Results and Next Steps: What were the results? What did we learn? What will we do next?	This attempt was more successful than last time. I was able to use this to introduce a chapter as we first started looking at exponential functions. Every student wrote at least a couple on their page before sharing with groups/whole class, and each team shared out at least one idea. After this activity, they did the CPM lesson of exploring the exponential functions for different b values, and at the end of class we came back to this WODB slide and they said what new things they understood: for instance, what the equations for these 4 graphs could have been. I want to continue to find ways to tie this activity into the specific lesson they're about to do, as a landmark to visit before/after the lesson.

Figure 10. Second PDSA iteration for Central teacher, from February 16 – 23, 2018.

In the next example, a 6th grade teacher was also building upon a previous Which One Doesn’t Belong PDSA cycle (not shown). He experimented with the idea in a fraction-related lesson in the last 10 minutes of one of his classes. He predicted that the students would engage in the activity and demonstrate academic vocabulary related to relevant mathematical concepts. He found that while students were engaged in the activity and demonstrated knowledge around fractions, they struggled with reasoning and explaining their ideas. He felt that the use of visuals would further improve students' engagement in the activity, and help them articulate their reasoning. Additionally, he thought that visuals would help students better understand the relationship between fractions and whole numbers.

During his second PDSA cycle, Figure 11, he introduced the Which One Doesn’t Belong activity as part of a warm-up, in a different class period. Informed by the data in his first iteration, he modified his instruction. He incorporated visuals to help students “see” the relationship between multiplication and division in fractions. He found that his students were more engaged, and that the use of visuals helped them think about the problem, then articulate their reasoning. Through these cycles, he gained valuable knowledge (and evidence) that incorporating visuals into his lessons could engage the students and assist their learning. He indicated that he would continue to incorporate them into problem-solving activities, thus potentially improving his instructional practices. Depending upon the degree to which these learnings changed his instructional norms, this instance of single-loop learning could also incite double-loop learning.

PDSA Cycle Form

Change Idea	**Describe:** Which One Doesn't Belong **What is the change:** Students will "see" the relationship between fraction division and multiplication. I will do a second WODB with the class, modifying the first WODB with more visual examples of fractions.
Predictions: What improvement do we think will happen?	Students will be more engaged, with students that do not normally get engaged with the activity. Students will understand the relationship between fractions and whole parts.
Questions: What do we want to learn from this cycle?	What are my students thinking about fractions and what concepts are easily seen, and others which are not?
Data: What information will we collect to answer our questions and test our prediction?	I will record the responses on the board. Students will create their own for homework.
Results and Next Steps: What were the results? What did we learn? What will we do next?	Student engagement was more apparent during this version of the activity. The drawings and figures were able to get more of the students engaged and thinking about the problems. They were able to also find more creative answers and explain their reasoning better. I will use this info to help incorporate drawings into problem solving. Moving forward, I will now implement mixed numbers and operations through the use of the visual drawings.

Figure 11. Second PDSA iteration for Sawyer teacher, from February 6, 2018.

In addition to generating single-loop learnings through PDSA cycles, teachers also shared effective practices and resources with each other during network meetings, which could lead to improved instructional practices, as well. In one example during the April convening, teachers from multiple schools were engaging in inquiry and dialogue around their PDSA cycles. While one teacher shared a web-based graphing tool that she used with her students, I observed two other teachers, who were from different schools, immediately navigate to the website on their laptops, and ask her questions about how she applied the tool in the classroom. Building upon this excitement and eagerness to learn new ideas from other schools, we structured our last convening as a “showcase” whereby selected teachers demonstrated the different resources they used during their PDSA cycles. Again, it was apparent during observations, that teachers were enthused and engaged. The room was buzzing with on-task conversation and questions. One teacher shared in the survey, “I enjoyed today's session more than the previous one because there was a lot of designated time available to learn about different practices in greater detail and ask questions.”

Learnings that Were Characteristic of Double-Loop Learning

While I cannot be sure that instructional theories-of-use were changing, there was evidence that teachers were reconsidering *how* they teach, versus merely looking for ways to improve their current practices of teaching. Their own PDSA forms, words, and behaviors provided the best indication of these shifting values, norms, strategies, and assumptions around student learning. According to Argyris and Schön (1996), it is this questioning of underlying values, norms, or strategies that will ultimately lead to more transformational change in practices.

Double-loop learning, or learnings with characteristics of double-loop learning, seemed to occur through the combination of multiple means: math professional development, conducting PDSA cycles, and engaging in disciplined inquiry around the PDSA cycles. All of these experiences likely contributed to shifts in underlying instructional values and norms when it occurred.

In Tom's professional development sessions, he explained the philosophy behind the number sense routines and how to build upon student learning. He had teachers assume the role of students during these activities to give them the perspective of how their students learned. During PDSA cycles, teachers posited certain predictions, and reflected upon evidence that confirmed or disconfirmed these notions. They tried multiple iterations of an idea and were presented with data on how new strategies/practices were more beneficial to student learning.

They reconsidered certain practices, and why they do them. Even in cases when teachers did not complete a PDSA cycle, but still engaged in inquiry and dialogue with their colleagues, it spurred their own reflection about teaching practices.

The previous example in Figure 11 illustrates how double-loop learning could emerge from multiple PDSA iterations. A teacher from Roosevelt provided another example. He conducted four PDSA iterations around the general idea of group closure activities; those tasks that end a lesson. These activities can have multiple purposes, among which, to gauge student understanding, and summarize major points of a lesson. He conducted the four PDSA cycles over the course of approximately six to seven weeks.

A review of his PDSA forms showed how his thinking evolved through the process. For the first PDSA, he employed a team exit slip to assess student understanding. Upon reflection, he felt that the exit slip was easy to grade and fostered student discussion. Next, he experimented with another group closure activity where student groups wrote a summary of the day's lesson. He felt this activity also cultivated student discussion, but expressed concern that his students did not exhibit a solid mathematical vocabulary in their written summaries. He decided to provide sentence starters for the written summaries in his next PDSA round. He found that, again, his students had a good discussion, but very little academic mathematical language was being used. From here, he wrote in his PDSA form that he planned to add a list of vocabulary words with the sentence starters for the fourth PDSA round. But, then something changed. Instead of administering another iteration of the written summary group closure activity, he pivoted away from his initial idea, and instead, had the students write their own problems. In his PDSA form, he described the process that led to his realization:

> *I really wanted to know what students were thinking and how they were problem solving. The group exit slip did not provide that information for each student. The group summaries were too general and it didn't provide me any information about*

their understanding or how they were thinking. So I thought having them create their own problem would give me better insight on their thinking and understanding.

This teacher realized that exit tickets and written summaries did not provide him with enough information about student understanding. Somewhere along the process, there appeared to be a shift from wanting to know *what* his students knew, to *how* they think – a potential change in instructional values. My analysis cannot attribute this change to any particular event but it does represent a possible shift that was in line with Tom's vision of eliciting, understanding, and building upon students' thinking.

Reconsidering instructional practices and underlying norms and values, was particularly significant for teachers who only completed one or two PDSA cycles. These teachers seemed more reluctant to complete PDSA cycles and slower to embrace Tom's instructional vision, which was likely related to their teaching philosophies. Interviews with these teachers indicated that they tended to be more teacher-centered in their teaching styles, rather than being student-centered. As previously discussed, they were also more likely to express concern regarding discipline and classroom management, and may not have initially valued classroom experimentation as much as teachers who completed more PDSA cycles. Thus, these were the very teachers who needed to reconsider their assumptions about how students learn, and why they taught the way they did.

In one example, a teacher who had not yet demonstrated a change in practices, and was hesitant to try a change idea in her classroom, was beginning to voice a new sentiment regarding student-centered learning. She often referred to herself as a "traditional teacher," and began the year expressing uncertainty about the school's more collaborative curriculum. Towards the end of the year, she articulated how network activities helped her see the value of student-to-student discussions:

I mean, it [participating in the network] re-enforces what we're doing with this curriculum or the way the curriculum is set up. [It] is definitely a plus. Not just the content of math, but those soft skills about communication, and talking to each other, and putting that extra time.

In another case, a teacher who was very concerned about preventing his students from misbehaving and seemed reluctant to try a new idea in his classroom, was beginning to show a greater inclination towards student-directed discussion. He explained:

Well, I try to devote more time now to discussions, discussions amongst the students. We do these things, particularly with the warm-ups, and also when we're introducing a new concept. The idea of wait and think, think, pair, share. But just the idea of giving them a chance to talk about it and try if… Of course, if they really are talking about it. Then I just let it go for a while, and then, I use the "If you have one idea…" …Yeah, that's nice. The kids like it as something that they had not done before. But I mean, it's effective in that you can see what their ideas are, and how many of them you know, are thinking. I think just the idea of them again, discussing the concept or the problem, how to apply the concept, it's taking a little more time with it. It's helpful.

While this teacher's quote signified a change in strategy within his usual instructional paradigm (warm-ups to existing lessons), it also indicates a deeper shift in values regarding his teaching style and learning about student thinking. His administrator confirmed the shift that he, and other teachers, were making in their instructional practices. Importantly, this was a school where Tom was also working with teachers through the Math Project. Teachers were getting more frequent exposure to Tom's instructional views. The administrator reported:

So, they been doing grouping strategies, engagement strategies, much more mindfully and consistently this year, than they had in the past. I don't know if they realize that's because they've been talking about it more, and more exposed to it…I think he's just *really mindful about incorporating those strategies into his instruction, as well. And then grouping stuff. So, things that came out of the Math Project, came out of [the network], he's definitely implementing. They're doing it, as well.*

Some teachers also demonstrated less student-deficit thinking and language. This indicated that these teachers were changing their mindsets, or at the very least, questioning their underlying instructional values and norms. Several teachers began exhibiting more asset-based language during meetings. Others, who were initially very vocal in meetings about their students' lack of skills and knowledge, were now silent on the subject, suggesting that the group norm had shifted as well. Through his work with the teachers, including settings outside of the network, Tom was particularly cognizant of the language shifts:

That they might, might have been using more deficit language in the past, but now they're kind of thinking about like, "Okay, so I have a lot within my locus of control that can actually influence how the kids engage with math." Um, I think implicitly they're understanding that, and it's kind of shifting their thinking about the kids as, "They just don't know it, or they just disengage," to "Okay, there is something that they, they do know. "What the very least, that now they're learning to not be so blatant about, like, their bias against kids. It's shifting in that sense, too. I'm hearing less language of, you know, "These kids just don't do this or don't do that." Which again, even that in itself, is something that they're learning.

The Importance of Reflective Structures for Single- and Double-Loop Learning

While the two previous sections illustrate single-loop learning and potential double-loop learning, and identify processes by which they occur (e.g., professional development, PDSA cycles), a more thorough discussion about reflection on practice is warranted as a potential mechanism of these learnings. Through network activities, teachers engaged in internal and external reflection that resulted in individual and team learnings, and teachers commonly attributed their learnings to this reflection time.

As mentioned in Chapter 5, PDSA cycles offered a structure for experiential learning (Langley et al., 2009; Merriam & Bierema, 2014). More specifically, the converging of math instructional professional development with this PDSA format represented a reflective practice model of experiential learning. The reflective practice learning theory advances that learning occurs by reflecting upon or in our practical experiences (Merriam & Bierema, 2014). More than 30 years ago, Schön (1983) posited that many professionals, including teachers, were faced with complex issues – the real-world of practice was messy – and that learning in practice was necessary to be successful (Schön,1983; Merriam & Bierema, 2014). He believed that professional training alone was not enough to deal with this real-world complexity.

In our case, the PDSA cycle provided a format for trying new ideas, individual reflection, and dialogue and inquiry among colleagues. Teachers had a change idea, predicted the results, experimented in their classrooms, collected evidence, and compared their prediction with the actual results. Combining this format with Tom's number sense routines (or other change ideas) provided worthwhile content for them to put into practice, and then reflect upon. Learning commonly occurred through teachers' reflection upon their experiences experimenting in their classroom, which is referred to as reflection-on-action in this model of experiential learning (Merriam & Bierema, 2014; Schön,1983;).

Learning through reflective practice can also occur by considering one's espoused theory against their theories-in-use, as previously discussed (Merriam & Bierema, 2014). The reflective structure of the PDSA also provided a vehicle for examining conflicts between espoused theory and theories-in-use, whether implicitly or explicitly, particularly through reflection upon whether or not they met their prediction, and why. This idea of confirming or disconfirming assumptions through the PDSA cycle is discussed more in the next section.

It was clear through teacher interviews, observations, and the end-of-year survey, that the PDSA format facilitated both action and learning. One teacher, who was already very comfortable experimenting in her classroom before the network, acknowledged that this structure “reset” or reinvigorated her own practice, a sentiment that was echoed by other teachers as well. This teacher explained:

> *I feel like, I've had a lot of priorities like, reset, um, I don't know how to say this. There are things in teaching that, I like, have done, like, a couple of times and I know that it works great, but then you just kind of don't do it because you're, like, trying to, you know, get through with all these other things. It's, you know, you're juggling a bunch of balls and whichever one gets closest to the ground is the one you're catching right now. But, you know, it was, it's been a nice opportunity to really have time to reflect and work on, “OK, what, what are the things I want to actually change, not just go along with how I've been doing it?” And that's been really fun.*

Other teachers expressed a similar appreciation for experimenting and reflecting in their practice through the survey. Teachers were asked, “Has participating in the network caused you to reflect upon your own teaching practices? If so, how?” Almost half of the teachers (7 out of 18 teachers) indicated that participating in the network caused them to see the value of trying or implementing new instructional practices, which was the most common answer given. Other answers acknowledged that teachers were thinking more about how to engage students and/or meet their students’ needs (5 teachers). Specifically, one teacher expressed:

> *I have questioned why I do a lot of things in my practice to see if it is helping reach the end goal. I have reflected numerous times on the little changes that I make, or the new things I try, (even if it's not my PDSA) to see how effective the practice is. I feel like I've tried more things in my practice than I did the last two years I've been at this school.*

As the network hub, we intentionally designed activities to foster meaningful questioning and reflection about their practices and PDSA cycles. We also aspired to aid learning for teachers who had not completed their own PDSA cycles. Specifically, we:

- Provided opportunities for teachers to dialogue with colleagues within, and across, schools.
- Explicitly taught productive inquiry and dialogue skills (described in Chapter 4).
- Strategically grouped teachers to foster dialogue and inquiry across schools. We grouped teachers based on their personality and needs. We tried to combine teachers who we felt would connect and build relationships, provide valuable guidance, and/or who push each other's thinking.
- Developed a productive inquiry and dialogue protocol to steer these conversations. This protocol included four timed steps:
    1. Teacher shared the results from their PDSA cycle.
    2. Other teachers asked clarifying questions, and then more probing questions and/or suggestions to consider. Questions stems were provided.
    3. The initial teacher reflected upon what they heard from the others.
    4. Once all teachers had their turn, they identified bright spots (what worked well) and additional questions for consideration.

Intentionally Designing Activities to Foster Double-Loop Learning

Although we did not intentionally design Tom's number sense routines to foster double-loop learnings, upon our own reflection, we realized that these activities, combined with the PDSA format, promoted double-loop learning processes. This was found in their advancing a position that they could confirm or disconfirm with evidence, and coupling it with team inquiry (Argyris and Schön, 1996). For example, during the initial causal analysis conducted in the fall, many teachers articulated deficit thinking regarding their students. They expressed opinions that many students were not interested in math, did not possess basic mathematical skills, could not do basic reasoning, and so on. Tom's Which One Doesn't Belong and choral counting routines challenged these deficit assumptions.

Through the PDSA process, some teachers predicted that they would see an increase in participation and engagement, and in some cases, 100 percent participation. What is important to note here is that many of the teachers making these predictions were the *same ones* who attributed students' lack of interest in math to the students. This suggested that their espoused theory (PDSA prediction) might have been in conflict with their actual theories-in-use (instruction that assumes these students are not participating because of their own lack of interest, rather than the instructional practice, itself). Because the PDSA process required collecting observable evidence to assess whether they met their prediction, teachers in these examples collected data – students who participated in the activity, and what they said – and were often confronted with evidence that disconfirmed their initial theories-in-use. That is, the data suggested that more students were participating when the instructional activity engaged them in a different way. This was especially true when teachers iterated upon the initial idea and provided opportunities for students to engage in multiple ways, such as writing individually or in a small group. As already discussed in this chapter, some teachers were shifting their own thinking, language, and instruction as result of these activities.

It is also important to note here that in order for a teacher to self- and group-reflect upon this disconfirming evidence, it is extremely important that they are supported through either professional development and/or coaching, so they can successfully execute the activity. If a teacher cannot effectively implement the number sense routine, including knowing how to adjust when students need clarification or get stuck, they may be collecting evidence that confirms their student-deficit assumptions, i.e., attribute failure to their students, and not their own instructional gaps around the activity. Obviously, this would be detrimental to the learnings of the network.

Additionally, we recognized that PDSA inquiry and dialogue structures, whether within schools or network meetings, did not provide an abundance of time for teachers to engage in deeper discussions reflecting upon the potential conflict between their espoused theories and theories-in-use. Each teacher typically had about 10 to 15 minutes to share their PDSA results, and engage in inquiry and dialogue with their colleagues. Therefore, we decided that the network should promote these deeper discussions by intentionally creating the space and time.

At the April network convening, we designed an activity we called "circles of engagement." Teachers were grouped with teachers from other schools, and asked to label some of their students by engagement level (not engaged, somewhat engaged, very engaged) and consider why they assigned those labels. Then, they were asked to read a short article about middle school students' developmental needs and the classroom environment. University hub staff facilitated structured small group dialogue about the reading and teachers' own classroom experiences. Although, I have no evidence that this conversation resulted in double-loop learning, it spurred a lot of reflection and new dialogue around student experiences, and how teachers might better engage students. Upon the observation of two table discussions, it was clear these teachers were acknowledging that student engagement is a product of their own instructional practices. I heard teachers discuss how they could move students from the "not engaged" circle to the "very engaged circle" by using technology, giving students more varied opportunities to share their thinking, and more ownership and choice in their learning. Teachers were engaging in authentic group dialogue, which was apparent by their ability to advocate their own views, and acknowledge and question the perspectives of the article and other teachers.

Conclusion

The ultimate purpose of networked improvement communities is to generate collective learnings that can improve complex problems of practice. In education, these complex problems can be related to deficit-oriented instructional patterns, which may be tied to educators underlying values and norms. As such, improving persistent educational problems may require more than single-loop learning. All learning is not created equal. Network improvement communities should cultivate double-loop learning through PDSA cycles and thoughtful reflective and inquiry structures that provide opportunities for teachers to collect evidence that confirm or disconfirm their instructional assumptions.

CHAPTER 8

DISCUSSION

Introduction

Recent attention has turned to the promise of networked improvement communities to improve educational outcomes for students, as evidenced by the 2018 Gates Foundation initiative to fund Networks for School Improvement and the increase in relevant literature that has emerged over the past year. The potential of these networks is predicated on the idea that uniting communities of practitioners, researchers/evaluators, and subject knowledge experts will improve persistent educational challenges. By embracing diverse experiences from multiple schools and teachers, new knowledge will be generated to improve complex problems of practice. While networks can incorporate any continuous improvement framework, improvement science tends to be the most commonly established one due to its promotion by the Carnegie Foundation for the Advancement of Teaching.

But the shiny promise of networked improvement communities can be dulled by the reality of implementing improvement science in education. Networks that apply an improvement science framework require their members to learn and apply technical methods, tools, and terms, such as fishbone diagrams, process maps, driver diagrams, and PDSA cycles. Given their extensive daily duties, learning this new language and the technical tools may be difficult and daunting for teachers and administrators. Furthermore, the challenge of building improvement science capacity, both at the individual and organizational level, can be exacerbated by the diversity of multiple school contexts; no two schools are alike, and neither are their issues. Existing literature regarding network improvement communities provides scarce guidance on how to deal with these challenges and the inherent messiness of the real world.

This study sought to understand how a networked improvement community could prepare its schools to successfully participate in the network. Its purpose was to provide an empirical account of the intricacies of starting a new network and building the improvement science capacity of its members. My role in leading the network enabled me access to project documentation, artifacts, discussions, and notes, and conduct more formal observations, interviews, and surveys, to study how the network unfolded over the first year. By providing a descriptive narrative of network processes, decisions, and outcomes, my hope was that this research would provide guidance to others who aspire to start a networked improvement community. With this aim in mind, my research questions were:

1. How were schools prepared to successfully participate in the University networked improvement community?
    a. How did the hub (University) teach the schools improvement science?
    b. How did schools learn improvement science?
    c. How were the fundamental network components developed, i.e. theory of practice improvement, change ideas, measures?
2. What facilitated a school being prepared to successfully participate in a networked improvement community?
    a. What capacities, structures, and/or conditions were needed for schools to successfully implement improvement science?
    b. What capacities, structures, and/or conditions were needed for schools to generate learnings that contributed to the learnings of the network?

This research utilized multiple methods in an explanatory single case study of the network, with schools as embedded units. The case study was bounded by network-related activities in the first year. The network's end-of-Year 1 goal was to build individual and organizational improvement science capacity, as substantiated by completed PDSAs. Network members from all schools completed a total 63 PDSA cycles, with nearly all aligning to the network driver diagram. Due to this positive outcome, I employed process tracing analytical techniques in an effort to uncover what network processes contributed to teachers implementing improvement science and generating subsequent learnings, along with conducting cross-school (embedded unit) analyses to determine what capacities, structures, and/or conditions were related to implementation differences among schools.

Review of the Findings

Preparing Schools to Participate in the Network

The first research question explored the hub's role in establishing the network and teaching the schools improvement science. This question is primarily answered through a descriptive narrative of the hub's processes and decisions. This account shows that the University hub engaged in the following strategies to teach improvement science:

- Provided professional development presentations in improvement science methods and math instruction at network convenings.
- Provided work time at the network convenings for teachers to work with tools (e.g., fishbone, driver diagrams), and to plan PDSAs.
- Facilitated the application of improvement science tools at school meetings.
- Explicitly taught inquiry and dialogue skills to foster constructive PDSA conversations, within and across schools.
- Presented information about practical measures and tools to collect data at the network convening (albeit unsuccessfully).

While not explicitly stated, the narrative conveyed three important findings for how the hub established the networked improvement community and prepared its schools to successfully participate. Other relevant findings were also illustrated by this narrative and more explicitly stated under the second research question.

Expertise of network hub team. One of the first steps in founding a networked improvement community is composing a hub team. The findings suggest that the network hub team should have considerable expertise in three distinct areas: improvement science, subject knowledge, and building relationships with schools. It was this combination of expertise that enabled the University hub's ability to teach the schools improvement science.

This study demonstrates that a network improvement community benefits from a team member having previous improvement science capacity building experience in education. As the hub's improvement specialist, I brought experience in this practice, and was guided by my improvement science knowledge and understanding when making pertinent decisions about building improvement science capacity, developing fundamental network components (e.g., driver diagram), and when to adapt (or not adapt) improvement science concepts and methods. It was through my strong grounding in evaluation and improvement science, and past experience in these practices and in schools (and in teaching) that I understood the importance of prioritizing a capacity building goal before trying to improve the network's problem of practice. I was also committed to continually emphasizing our driver diagram as our network's theory of change, and conducting short PDSA cycles to accelerate the network's learning. I understood when to be flexible and when to stay rigorous to the process. Without my experience and acquired expertise, I might have made different and less successful choices, as I had in the past.

A networked improvement community hub will also benefit from having a team member with strong connections to the member schools, and familiar knowledge of how the schools operate. In other words, someone familiar with the political and organizational particularities of the schools, at least at a broad level. In this particular network, the school liaison played an important role in establishing a receptive atmosphere and setting up meetings at the schools, especially during the early days of the network. As a result, the schools progressed more quickly than the larger network through the initial phase of network initiation (i.e., understanding of the problem). Without these relationships, these meetings may not have been established so quickly, or so regularly (for most of the schools). The liaison's rapport with the schools also allowed the hub to be persistent without being perceived as bothersome. It was through these meetings that the hub also continued to build stronger relationships, and trust, which fostered honest teacher feedback, enabling the hub to more effectively respond to teacher needs. This was true even in the two schools where it was difficult to establish regular meetings.

The subject matter expert is the third invaluable element for teaching the schools improvement science and guiding the development of the network's theory of practice improvement. Without subject matter expertise – in this case, a math coach – it may take longer to employ this theory (i.e., driver diagram) and/or it may be difficult to establish meaningful changes to the system. In this study, it is quite likely that teachers' thinking would not have been pushed beyond their current classroom practices without the math coach's presence. In his role as a subject expert who sat outside of the current system, he helped schools understand alternate ways to engage students in mathematics. He knew how to engage teachers and give instructions for implementing the drivers in the network's driver diagram. He provided practical and concrete examples of what our drivers looked like in practice, and an entry point for teachers to experiment with ideas in their classroom. He knew a larger paradigm shift was required, and introduced a new system. The math coach also opened each network convening with professional development because we found that teachers related to it, enjoyed it and found it valuable.

This last point further suggests the synergistic nature of this finding. A team of complementing experts not only provides unique expertise in their respective areas, but also merges knowledge in a way that elevates the hub's ability to teach the schools improvement science. This finding suggests that networks should build teams with experts in all three areas, and work closely together throughout the process.

Two visions and learning in context. Throughout this study, there was evidence that teachers learned improvement science by tightly connecting concepts and tools to the teachers' instructional practice. Possessing expertise and clear visions in both the subject (math instruction) and profound knowledge (improvement science capacity) enables these connections. For example, the instructional number sense routines (e.g., Which One Doesn't Belong) provided manageable and meaningful content that could be experimented with in a quick PDSA cycle, and provided easily collectable practical data. Furthermore, learning improvement science in the math instructional context helped the teachers become more motivated about the process, as they tended to be more interested in improving their own instructional practices and collaborating with colleagues rather than learning the improvement science tools and processes themselves.

This finding suggests several important ideas for preparing schools to successfully participate in a networked improvement community. First, network hubs need to purposefully build the improvement science capacity of its schools and teachers. Due to my role in the network and my program evaluation background, and the fact that improvement science can be considered a participatory evaluation framework, the University hub conceptualized an evaluation capacity building model that distinguished individual and organizational capacity needs (Cousins & Earl, 1992; Christie, Inkelas, & Lemire, 2017; Preskill & Boyle, 2008; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012). Through the model, which also was the conceptual framework that guided this research, the hub identified activities needed to build individual teacher capacity, with a focus on building skills and behaviors, and those needed to embed capacity in the schools. By categorizing these activities, it was clearer what actions the hub needed to undertake to build capacity. This model guided Year 1 network activities.

Second, similar to the previous finding, it is important to establish two guiding visions – one to improve the subject area and one to build improvement science capacity. They are synergistic. Rather than operating on two parallel separate tracks, the notions of fostering engaging math instruction and teaching improvement science skills became interconnected and advanced teachers' network learning experiences. By framing the improvement science tools around the math instructional professional development, their learning was situated within the context of their classroom. Drawing upon principles from situated cognition learning theory, teachers were able to construct meaning and make sense of improvement science concepts that might have otherwise seemed disconnected and abstract (Merriam & Bierema, 2014; Brown, Collins, & Duguid, 1989).

Finding coherence and consensus. Another important point illustrated by the network's narrative, is the importance of the hub's role in fostering coherence across the diverse schools and finding consensus among them. The networked improvement communities' design of embracing diverse school experiences can contribute to challenges when trying to find a common problem of practice, aim, and driver diagram.

In this particular network, the range of each school's initial draft aim statements and the inability of one school to find consensus even within their own school, illustrated the challenge of finding coherence within a networked improvement community. The University's network included two middle schools, one junior high, one high school, and one K-12 school. They were spread across a sprawling, diverse city. Each school had unique personalities and needs. Even within one school, teachers represented multiple grade levels and courses, ranging from 6th grade math to 9th grade Algebra I. Thus, in order to prepare schools to successfully participate in the network, the hub needed to be strategic about bringing coherence to varied problems of practice and finding consensus on the aim and driver diagram.

While other literature (Bryk, Gomez, Grunow & LeMahieu, 2015) suggests considering one of two approaches – convene an initiation team of experts upfront to draft these fundamental components or convene the schools to draft these components together – the findings from this study suggest another viable approach. It is essentially a hybrid of the two, and was made possible by the hub's close working relationship with the schools.

These components can be developed by a core expert team *and* incorporate each school's perspective. This finding was a bit fortuitous because the University's hub had already started working with the schools and facilitating their dialogue around understanding the problem before it relaunched the network. The hub team knew the results of each school's causal analysis, and had engaged in dialogue around these issues with them. This presented the team with an opportunity to embrace the math coach's outside-the-system math instructional vision, combine it with each school's interest and needs, and quickly develop the network's fundamental components. This approach prevented a lengthy and potentially difficult process of finding consensus among the schools, and in the case of the driver diagram, alleviated the teacher's burden of having to learn how to develop a complex theory of change. This study suggests that teachers will not only embrace these components, but may also be grateful for simplifying the process.

Examining What Facilitates Successful School Participation

The second research question sought to further explain what capacities, structures, and/or conditions contributed to schools successfully implementing improvement science, and generating meaningful learnings. The network as a whole was deemed a success because of the improvement science capacity demonstrated by its members. The evidence of successfully building these skills and behaviors was the number of PDSA cycles implemented (63), and the learning generated through these processes. However, within the network there was variation by school and teachers. While this variation indicated that some schools and teachers were less successful than others, it also provided the opportunity to examine why there were differences and what facilitated success.

Improvement science implementation. There were three notable findings regarding improvement science implementation in schools. The first finding is further parsed in smaller findings related to the primary finding.

In-school work structures. One overarching and imperative study finding is the importance of establishing work structures *within* the schools whereby the teachers and hub team work together on improvement science activities. This study suggests that the number of completed PDSA cycles per school is positively related to the number of times the hub met with them. Based on my previous experience leading other improvement initiatives, it can be easy to take these in-school work structures for granted, assuming that administrators and teachers will make the time or utilize existing collaboration time, but these assumptions may be incorrect, and even unrealistic given their competing priorities. The three smaller findings, shown below, contributed to this primary finding regarding the importance of consistent in-school meetings with an improvement science facilitator.

*Time and space***.** When establishing the school work structures, administrators play a significant role. In the schools were administrators were more involved and supportive of the network, it was easier for the hub to establish consistent meetings with teachers. In the cases where administrators were either not fully involved, or did not value the process, it was more difficult to schedule regular meetings. Supportive administrators are needed to carve out the time and space for their teachers to meet. Without their cooperation and assistance, it will be difficult to for their teachers to implement improvement science activities.

Existing collaborative structures. Time and space alone are not enough. This finding indicates that schools tend to have pre-existing collaborative structures, however, they may not be conducive to productive team collaboration around improving practices. In this study, teachers indicated that they already met regularly, i.e., for example during department meetings or common conference periods, but they did not frequently discuss student learning nor solicit feedback from peers regarding their own teaching practices. Furthermore, the results of this study indicate that teachers may lack the necessary skills to facilitate constructive learning with their peers. Rather, they may have restrictive or loose meeting structures that deter productive dialogue and inquiry with one another, and/or may be reluctant to ask each other questions.

Benefit of an outside facilitator. Improvement science is technical and foreign to most teachers. Teachers who are learning improvement science need expertise from an outside facilitator to implement the tools and concepts. In this network, this need was evidenced on several levels: Teachers who lacked the capacity to engage in improvement science tools on their own, teachers who perceived they had the capacity to do their own learning, but not enough to teach and facilitate others, and teachers who perceived they had the teaching and facilitating capacity, but did not want the burden or responsibility. The presence of an outside improvement science facilitator also helped the teachers feel supported, while providing an objective perspective.

Taken together, with the evidence that teachers were engaging in more dialogue and inquiry about student learning and their practices by the end of the year, these smaller findings contribute to the broader finding. That is, the occurrence of consistent in-school meetings facilitated by an outside improvement science expert contributes to each school's ability to successfully implement improvement science. While some existing networks may have these structures in place – literature indicated that one network, the Building a Teaching Effectiveness Network (BTEN), has district facilitators who provide individualized school support – my experience and knowledge of other networks suggests that many do not (Russell, et al., 2017). Furthermore, existing literature does not detail the structures, processes, and challenges to consider when attempting to build school improvement science capacity. This study suggests that network hubs should attend to in-school structures by establishing clear meeting expectations with leadership, providing facilitation support to foster the use of technical tools, and promote collaboration focused on improving practices and student learning during these in-school meetings.

Consistency. There are two notable points concerning this finding. The first being related to the previous finding. Consistent use of improvement science methods contributes to teachers learning and applying them. This was evidenced in multiple ways: the regular use of the driver diagram at network convenings that contributed to teachers' outward articulation of how their change ideas aligned to it; the repeated PDSA planning time to work on short PDSA cycles that allowed teachers to regularly practice the process; and, the consistent engagement in inquiry and dialogue practices that helped teachers become adept at deeper teacher-to-teacher questioning.

This finding indicates that teachers need consistent practice when building new improvement science skills. Thus, they need opportunities to practice. As discussed in the last finding, it is unlikely that schools will have the time and space to cultivate these skills on their own nor the technical capacity to do so. Through school and network meeting structures, the hub should promote activities where teachers consistently revisit key improvement science tools and processes, such as the driver diagram and PDSA cycle.

The second point concerning the findings is also about consistency, but as it relates to the University hub. The consistency factor benefits network planning activities because the hub team regularly met with teachers, particularly in their schools, where teachers shared feedback between network convenings. Even when teachers did not explicitly express concerns, the hub team was able to observe potential issues, such as teachers' reluctance to ask each other questions during PDSA conversations. Through these feedback and observations, the hub team was able continually respond to actual conditions on the ground, and tailor network activities to the teachers' ongoing needs. Thus, this finding indicates that hub staff should consistently meet with teachers between network convenings in an effort to continuously improve their own processes for building the improvement science capacity of teachers and schools. Again, while it is possible that existing networks already consistently meet with members schools between convenings, my past experience and knowledge of other networks suggests that this is not a universal practice. Even if schools are receiving additional in-school meeting support, it may be district facilitators who are facilitating the meetings, rather than core hub staff (per the BTEN network description in Russell, et al., 2017).

Pre-existing individual capacity. Finally, while the previous findings in this section explain governable conditions and structures, it is also important to acknowledge another explanation regarding improvement science implementation. Pre-existing capacity differences among teachers can also contribute to variances in their completion of PDSA cycles. Teachers who tend to be less comfortable with the instructional ideas advanced by the driver diagram – teachers who are more teacher-directed in their instructional style, in this case – may be more hesitant to experiment with new ideas in their classrooms. Conversely, teachers who are already experimenting in their classroom, and informally reflecting upon their lessons, may be more comfortable completing PDSA cycles.

This finding suggests that individual and organizational capacity building efforts may be influenced by pre-existing individual capacity. Nevertheless, even those teachers who were more hesitant to try new ideas did complete at least one PDSA cycle. What remains to be seen for these teachers, however, is whether experimenting and reflecting in Year 1, albeit more slowly, will foster more experimentation in Year 2. This suggests a direction similar to those teachers who already had existing experimentation and reflection capacity in Year 1. Thus, this finding may be more of an expression of a developmental path rather an alternative explanation.

Generating learnings. The ultimate purpose of improvement science, or any continuous improvement process, is to generate meaningful and actionable learnings that lead to improvement. Therefore, my second research question also examines what learnings occurred in Year 1 of the network, and what contributed to their generation. Learnings, conceptualized as both single-and double-loop learnings in this research, were generated by most teachers in the network. To varying degrees, teachers demonstrated instrumental improvements in their current practices (single-loop) and shifts in their underlying instructional values and norms (characteristic of double-loop learnings). Teachers tended to attribute their learnings to their own reflections. Two principle findings fell under this research question. These findings are closely related as they both represent parts of an overarching reflection mechanism that facilitates learning.

Plan, Do, Study Act structure. This finding suggests that the PDSA format fostered learnings in multiple ways. This is not surprising considering improvement science advances PDSA cycles as its learning mechanism. However, what is notable about this finding are the various ways that the PDSA structure can manifest learnings in the education context.

First, through the PDSA format – identify a change idea, predict the results, experiment with it in their classroom, collect data, and reflection upon results – teachers generated knowledge about how to improve a particular practice. These learnings typically occurred one PDSA cycle at a time, and were more characteristic of single-loop learning as they were instrumental improvements to a change idea that teachers inserted into their current instructional paradigm (warm-up to the existing lesson) and did not seem to prompt a deeper reflection of instructional norms and values. The evidence suggests that these instances were more of a correction upon a change idea, rather than a deeper questioning of an instructional assumption or value underpinning the change idea. A teacher's ability to generate this type of learning was directly related to their capacity to conduct a PDSA cycle.

Second, multiple PDSA cycles, and more specifically iterations of the same change idea, seemed to stimulate learnings more characteristics of double-loop learning. When this occurred, the teacher reconsidered their strategies and whether they improved the specific issue, such as engaging students. Through the multiple iterations, these teachers strived to meet their prediction and that prompted more in-depth questioning of whether their change idea strategies improved upon previous iterations. This, in turn spurred further reflection upon why strategies *did no*t meet their prediction, thus, provoking a deeper contemplative questioning of instructional assumptions and norms underpinning the change idea strategies. Again, teachers' capacity for generating these learnings tended to be directly related to their ability to conduct a PDSA cycle.

Third, the PDSA format, as this particular network designed it, stimulated inquiry and dialogue around instructional practices. At the network convenings, teachers from multiple schools were grouped together to discuss the results of their PDSA cycle(s). This group reflection upon changes ideas, and ultimately instructional practices, along with the math professional development, likely contributed to a reconsideration of current practices even for teachers who only completed one or two PDSA cycles. Furthermore, teachers who had once overtly demonstrated student-deficit thinking and language (the same individuals who completed fewer PDSA cycles) were now either displaying a shift away from those mindsets, or choosing to keep quiet. Either way, the broad change suggested that these conversations were, at the very least, shifting group norms and values.

This finding highlights the vital importance of the PDSA cycle in networked improvement communities. This may seem insignificant because networks are designed to incorporate PDSA cycles, but my previous experience and knowledge of other networks demonstrates how challenging it can be to implement PDSA cycles in schools (Rohanna, 2017). Networks who fail to get traction with their PDSA cycles, or allow longer cycles that forestall multiple iterations, may not manifest the learnings necessary to solve complex problems of practice.

Professional development and reflective practice. This finding further unpacks this idea of network learnings and the crucial role that instructional professional development can play in fostering that learning. It is especially relevant to network improvement communities who aim to improve instructional practices and student learning.

When combined with professional development that teaches new instructional practices, the PDSA cycle offers an effectual structure for reflective practice. Reflective practice in itself, and more specifically reflection-on-action whereby teachers reflect upon practices they tried in their classrooms, is posited as a tool for teachers to solve complex problems (Schön, 1983; Merriam & Bierema, 2014). It is a process for learning in practice when more formal professional training may not prepare professionals (e.g., teachers) to deal with the real-world complexities (Schön, 1983).

This finding suggests that the pairing of instructional professional development with the PDSA format sparks a reflective practice mechanism that contributes to both single-loop and double-loop learnings. In this particular network, teaching the number sense routines, as practice for teachers to engage more students in mathematics, provided worthwhile content for teachers to implement and reflect upon. The PDSA prediction component, along with collecting data to assess that prediction, facilitates more informative reflection regarding whether the change idea results in an improvement and potentially how it could be further improved (depending upon the data collected in the PDSA cycle).

Furthermore, combining the reflective structure of the PDSA with an intentional instructional practice, such as the Which One Doesn't Belong routine, can cultivate double-loop learning. The PDSA structure provides a vehicle for examining conflicts between espoused theory and theories-in-use, whether implicitly or explicitly, as teachers reflect upon whether they met their prediction and why. In this network, evidence suggests that teachers, who initially felt some students would “never” be engaged in mathematics, were now reconsidering their own role in their students’ engagement. The PDSA cycle provided evidence that these students were, in fact, engaging in the activity. This gave teachers cause to reflect upon their underlying instructional theories-in-use. Networks can promote this development by purposely selecting instructional practices, as part of its professional development that can deliver evidence to disconfirm deficit assumptions about student learning. (e.g., “My students hate math.”) The degree to which this deeper reflection occurs is also related to teachers’ capacity for completing PDSA cycles, thus, further demonstrating the network’s need to purposefully develop individual and organizational improvement science capacity.

The idea of understanding the mechanisms that generate learning in networked improvement communities is a crucial one. Other relevant literature does not unpack what kind of learning occurs, or how it occurs. Yet, it is vital to understand these concepts because networked improvement communities are predicated on the idea of generating collective learnings to improve complex problems of practice (Engelbart, 2004). Furthermore, solving complex problems of practice in education likely requires both single-loop and double-loop learning. Persistent educational challenges can be due to educators’ instructional patterns that are tied to their own underlying values and norms (i.e., student deficit thinking). Changing these patterns may require confronting conflicts between espoused theory and theories-in-use, and the questioning of instructional assumptions about how students learn through the use of double-loop learning processes (Argyris and Schön, 1996).

Study Significance

The findings in this study paint a picture of how networked improvement communities can build the improvement science capacity of its member schools, which is a precursor to generating learnings that aim to solve complex problems of practice. As such, this study has important practical implications for networked improvement communities in education. The next section also advances a model for network hubs to consider when establishing new networks.

Establishing Networked Improvement Communities

One of the case study's most significant contributions lies in its detailed empirical account of establishing a networked improvement community, which can serve as a roadmap for others. Other literature offers practical guidance, but contains little empirical data or is missing a detailed account of the processes required (Bryk, Gomez, Grunow & LeMahieu, 2015; Russell, et al., 2017). Proger, Bhatt, Cirks, and Gurke (2017) note the scarcity of information about the *process* of forming a networked improvement community, and aimed to fill that void in their report entitled *Establishing and Sustaining Networked Improvement Communities: Lessons from Michigan and Minnesota*. However, they provided more explanation of *what* processes should occur, rather than a detailed unfolding of *how* they occur, with no guidance for actually building improvement science capacity. This literature also tended to focus more on the hub's role, and not the experiential component of the teachers and school participating in the network.

Notably, my findings and experiences are similar and/or complementary to those of Proger and her colleagues (2017). They found five key lessons in their process-focused study:

1. Build a cohesive team with members representing different types of expertise.
2. Reduce uncertainty by clarifying what participation entails.
3. Build engagement by aligning work with ongoing efforts.

4. Use tools and resources from improvement science to identify a problem that is important and specific enough to act upon.
5. Embed capacity building to develop additional expertise for using continuous improvement research to address problems of practice. (They also mentioned the benefit of an improvement science facilitator to build initial PDSA capacity.)

Their findings support the validity of my analysis, which is important, because of my dual role as researcher and network hub staff. However, as I mentioned, our studies vary in the amount of process detail. They provide little detail regarding how to apply their lessons in practice. This deficit is most apparent in their fifth lesson – embed capacity building to develop expertise for using continuous improvement research to address problems of practice – further highlighting the need for a detailed empirical narrative that portrays the decisions, processes, *and* challenges associated with building the improvement science capacity of network teachers.

Evaluation Capacity Building and Learning Theories

Building upon the idea that this research can serve as an empirical example for others, another study contribution is the incorporation of evaluation capacity building theory into network improvement communities. Existing literature on networked improvement communities does not provide this practical direction for building improvement science capacity. Evaluation capacity building literature provides an explicit framework that can be used by networks for identifying teacher and school capacity building goals. By using the Preskill and Boyle multidisciplinary capacity building model (2008), networks hubs can gain a greater understanding of what individual knowledge, skills, and attitude objectives to consider, along with what school organizational capacity is needed.

Furthermore, this study identifies underlying learning mechanisms for building the improvement science capacity of teachers, and for generating the necessary learnings for solving complex problems of practice. Existing network improvement community literature is devoid of this detail, potentially leaving a "black box" for those hoping to start networked improvement communities. Evaluation capacity building literature partially fills this void by directing those engaged in these efforts to consider adult and workplace learning theory (Preskill & Boyle, 2008), but it stops short of providing guidance on which particular theories to consider. This study expands upon existing literature by specifying situated cognition learning theory and reflective practice, and identifying the importance of incorporating meaningful instructional professional development to stimulate these mechanisms. Additionally, it highlights the vital importance of establishing productive in-school work structures in order to build capacity. Thus, providing a closer look into the black box for those hoping to launch new networks.

This study also unpacks the type of learning that can occur in networked improvement communities, and more specifically, PDSA cycles in education. As previously discussed, networked improvement communities are founded upon the idea of generating learnings to solve complex problems of practice. In education, persistent problems may be due to underlying assumptions about student learning and current instructional norms. Thus, networked improvement communities need to explicitly address these personal values and assumptions. By distinguishing double-loop learning from single-loop learning, networks can more purposely focus on these potentially powerful learnings, and strategies for fostering them. This includes incorporating meaning professional development activities that allow teachers to consider (confirm or disconfirm) their own theories-in-use.

New Capacity Building Model for Networked Improvement Communities

Based on my findings, I can further advance a model for preparing schools to successfully participate in a networked improvement community in the first year that draws upon empirical data, shown in Figure 12. This model is an important contribution to the field, because as previously described, there is no existing literature that details the process of how to build the improvement science capacity of educators participating in a networked improvement community. Furthermore, the model not only fills this gap, but it pinpoints underlying mechanisms of how improvement science capacity is built in the educational context and meaningful, necessary learnings are generated. Essentially, it is the opening of the black box.

Currently, network literature reveals two sets of broader tenets that, although fundamentally important, lack operational specificity. The first set developed by Bryk and his fellow authors (2015) establish six improvement principles:

1. Make the work problem-specific and user-centered.
2. Focus on variation in performance.
3. See the system that produces the current outcomes.
4. We cannot improve at scale what we cannot measure.
5. Use disciplined inquiry to drive improvement.
6. Accelerate learning through networked improvement communities.

Russell and her colleagues (2017) build upon these principles and further advance a framework for initiating network improvement communities. It offers five domains that they deem critical:

1. Learning and using improvement research methods.
2. Developing a theory of practice improvement.
3. Building a measurement and analytic infrastructure.
4. Leading, organizing, and operating the network.
5. Fostering the emergence of culture, norms, and identity consistent with network aims.

Many of these tenets are grounded in improvement science theory, and helpful for those launching new networks, but as already noted, are limited in their explanation of processes. Additionally, relying primarily on improvement science theory to guide network activities may be overly simplistic *in the educational context*. Improvement science was conceived from manufacturing and public healthcare improvement methods, both of which tend to be more process-oriented, aimed at production and costs, or patient safety and error prevention. But schools are very different places from factories and hospitals. They are often striving to change teachers' classroom practices, which can be very personal and deep-rooted.

The proposed model posits a new framework for networked improvement communities that is grounded in evaluation capacity building theory while embracing improvement science principles. From a theoretical perspective, there is no literature that merges the two concepts. From a practical perspective, it is advantageous to consider a model that applies evaluation capacity building theory to improvement science, because it explicitly distinguishes individual and organizational capacity needs and learning strategies to develop both.

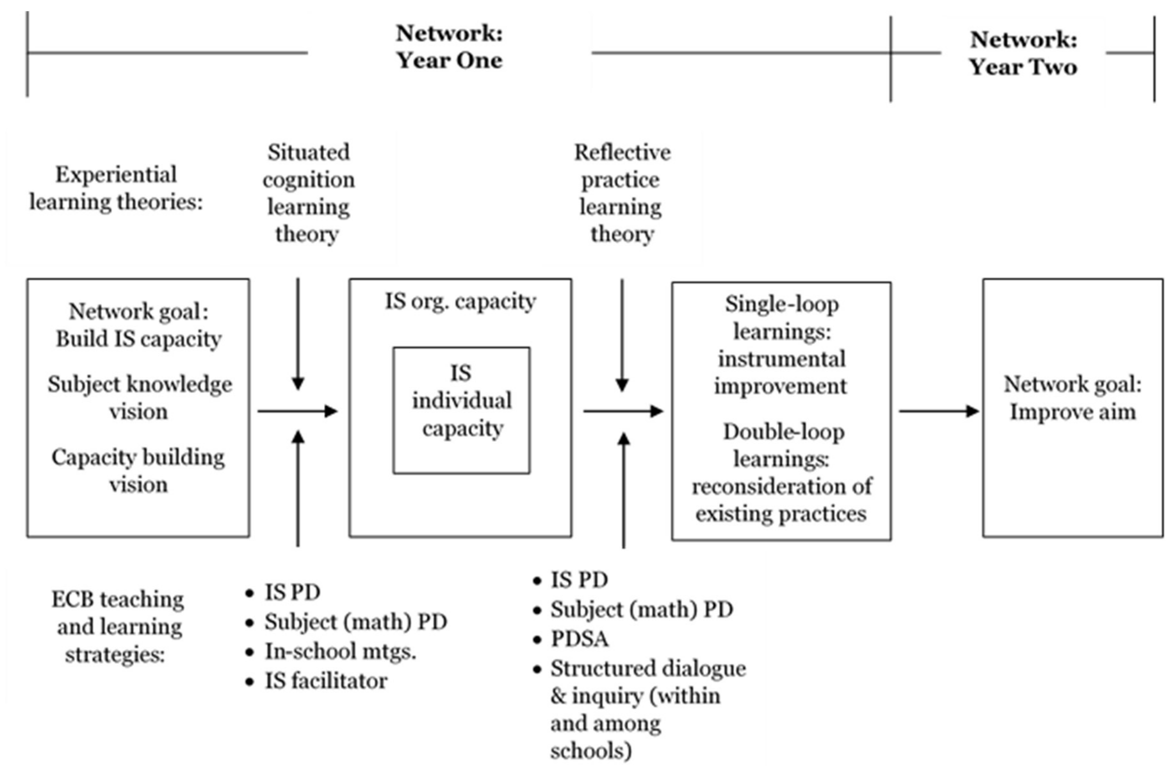

Figure 12. Year 1 improvement science capacity building model for networked improvement communities

The network's Year 1 goal (shown in the first box) is building improvement science capacity, which is served by identifying two visions: one regarding the subject knowledge, and one regarding the improvement science capacity. These visions are furthered by evaluation capacity building teaching and learning strategies (Preskill & Boyle, 2008), shown at the bottom of the model, and experiential learning theories, shown at the top of the model. In this model, situated cognition learning theory (learning in context) is a learning mechanism contributing to both organizational and individual capacity development (Box 2). This is transpiring through subject matter professional development, improvement science professional development, and in-school meetings facilitated by an improvement science specialist.

After the network hub builds improvement science capacity, the network can generate learnings, conceptualized as single-loop and double-loop learnings (Box 3). These learnings are generated through professional development, PDSA cycles, and structured inquiry and dialogue among teachers. Reflective practice (reflection on practice) is the underlying learning mechanism that drives the single-loop and double-loop learnings.

The model further posits that these *intermediate outcomes* – building improvement science capacity and generating learnings – need to occur before improving the network aim or problem of practice. This is theorized to occur in Year 2 of this model. Since this model is based on empirical data from Year 1, the box under Year 2 has not been further populated. Future research could study the factors and/or mechanisms contributing to success in Year 2.

Network hub staff can consider this model as a guide when preparing schools to successfully participate in networked improvement communities. Unlike existing literature, this model and narrative may provide a deeper understanding of how to generate these two intermediate outcomes, and draw a path to follow in Year 1. New network hubs can identify their own visions, and map out a comprehensive approach for building improvement science capacity and generating subsequent meaningful learnings, thus, gaining clarity on the mechanisms that were previously concealed in the black box. It should be noted that the findings from this case study, and thus this model, suggest one viable explanation with the presence of equifinality. That is, there could be other factors and mechanisms, or combinations of these that could lead to the same outcomes. This research only claims these findings as one potential path to derive these outcomes.

Limitations

This case study, like other research studies, has a few notable limitations. These limitations are discussed in this section.

While this study incorporates multiple methods, it primarily relies upon qualitative data. Qualitative data is more subject to researcher biases than other methods due to the researcher's role in collecting and analyzing interviews, observations, and artifacts through their own worldviews. Several steps were taken to increase the credibility of my analysis and findings. First, I gathered data from multiple sources, including survey data, and triangulated my analysis. Second, I followed process tracing best practices. Of note here, I systematically considered a wide range of alternative explanations for my findings, what evidence I would expect to see if my findings were true (or not true), and the potential biases of each source of evidence and weighted accordingly.

Another limitation of this study is that it is a single case study that was selected by convenience (i.e., the University network). Single case studies may be more limited in their generalizations than multiple-case studies. To address this generalizability limitation, I employed a nested case study design by embedding the schools, and conducted cross-school analyses to identify and explain variations within the one network case. Additionally, I incorporated theory to help explain the mechanisms that led to the successful outcome of the network (i.e., situated cognition and reflective practice experiential learning theories).

According to Yin (2014) and George and Bennett (2005), the use of these theories strengthens the ability to make generalizations from one case.

Finally, this study is also limited because it represents one cluster of schools, within a Southern California metropolitan area. These schools may not reflect the experiences of schools from other parts of the country. Furthermore, this network was an instructionally-based networked improvement community focused on building the improvement science capacity of teachers. This specific focus could limit the model's generalizability to other types of networked improvement communities that are less intent on building teacher capacity.

Directions for Future Research

The field of education would benefit from additional research related to networked improvement communities. More research is needed to guide new networks as they navigate the real-world challenges around supporting schools and building the evaluation capacity of educators. Future research could expand upon this present study in the following ways.

First, this study hypothesized a model for networked improvement communities to build improvement science capacity during Year 1. This model was developed from the empirical results of a single case study, and could be further evaluated through a future case study designed to test this theory. By doing so, the field could more directly examine how the experiential learning theories are explicitly put into practice and gain a more thorough understanding of how this approach could help networks.

Second, this networked improvement community was instructionally-based, and thus, the factors and mechanisms are specific to teachers and classroom practices. The field could benefit from considering other types of networked improvement communities that are less classroom-focused and experiment with broader goals such as improving districtwide attendance. Future research could examine what contributes to the successful preparation of members in these other types of networks, potentially resulting in the confirmation of a similar capacity building model or developing a new one.

Third, networked improvement community research would benefit from understanding processes and mechanism that extend into Year 2. Specifically, it would be useful to explore what evaluation capacity building learning and teaching strategies should drive Year 2, given that most, but possibly not all, members will have learned and practiced the basic improvement science methods and tools in Year 1. Maintaining and continually refining this capacity may require different strategies, plus, an examination of what processes and mechanisms contribute to improving the problem of practice after improvement science capacity is built. Future research could consider employing process-tracing techniques similar to those used in this case study.

Conclusion

From a theoretical perspective, this case study suggests that preparing schools to successfully participate in a networked improvement community necessitates more than improvement science theory. While improvement science provides tools and methods for understanding and improving problems, it does not provide knowledge of how to build the capacity for applying it. Evaluation capacity building theory fills this void. It provides a foundation for a new model, with a developmental Year 1 goal (build capacity) and improvement-oriented Year 2 goal (improve problem of practice), that explicitly incorporates experiential and organizational learning theories. Although not explicitly identified as a Year 1 goal, this model also specifies single-loop and double-loop learnings that will contribute to meeting the Year 2 goal.

From a practical perspective, this case study provides a detailed empirical account of *how* to establish a networked improvement community. Anyone hoping to start a networked improvement community, especially those with an instructional focus, would benefit from identifying a capacity building goal for their first year and working closely with a subject knowledge expert. While there may be other theoretical mechanisms or capacity building teaching and learning strategies that contribute to network success, this research provides a deeper examination of specific processes in a real-world context. Thus, it provides a roadmap for establishing successful networks, and helping others unlock the promise of networked improvement communities in education.

Appendix A

Interview Protocols

University Hub Staff Interview: Semi-structured Interview Protocol

Introduction: Thank you for meeting with me today. I'm hoping to learn more about you and your experience participating in the University networked improvement community and your perspective on how schools were prepared to successfully participate in the network. All of the information you share with me today is confidential. I'd like to encourage you to be extremely honest today. Please share what you feel comfortable sharing. You can choose not to answer a question if you prefer not to. Do you have any questions before we get started? [ASK THEM IF IT IS OKAY TO RECORD SO THAT YOU DO NOT HAVE TO TAKE DETAILED NOTES DURING THE INTERVIEW]

First, I'd like to learn more about your background.

- Tell me a little bit about your current position with the university?
- What is your experience working with schools?
- How did you end up working with the University networked improvement community?
- What is your role in the network?

Now I would like to turn the network, when you think about the network in general, what stands out to you? Tell me more.

Next, I want to get your perspective on how, and to what extent, the schools learned improvement science.

- To what extent, do you think the schools have learned the different components of improvement science? [PROBE FOR COMPONENTS]:
    - Root cause/causal analysis
    - Driver diagram or theory of change/improvement
    - Developing strategies/change ideas
    - Conducting PDSAs
    - Collecting and analyzing meaningful data
- How do you think they learned it [EACH COMPONENT]?
- What capacities, structures, and/or conditions helped to facilitate schools learning improvement science? [PROBE FOR CONTEXTUAL CONDITIONS TOO]
- What capacities, structures, and/or conditions do you think hindered schools learning improvement science? [PROBE FOR CONTEXTUAL CONDITIONS TOO]

I'd also like to get your perspective on the learning and sharing between and among schools.

- To what extent do you think teachers, **individually**, generated new knowledge or learnings during this process? (So not sharing yet, just learnings for themselves).
    - What capacities, structures, and/or conditions helped to facilitate those learnings? PROBE FOR CONTEXTUAL CONDITIONS TOO]

- To what extent do you think teachers learned from one another? Both within school and among schools?
    - What types of things do you think they learned from each other?
    - What capacities, structures, and/or conditions helped to facilitate those learnings? PROBE FOR CONTEXTUAL CONDITIONS TOO]
    - What capacities, structures, and/or conditions do you think hindered those learnings, if at all? [PROBE FOR CONTEXTUAL CONDITIONS TOO]

What else would you like to say about what schools learned through this process that we haven't already discussed?

Okay, we're almost done.

What do you think is the hub's role in our networked improvement community?

How can we best support schools in this process?

That was the last question. Thank you.

Teacher Interview: Semi-structured Interview Protocol

Introduction: Thank you for meeting with me today. I'm hoping to learn more about you and your experience participating in the University networked improvement community, including your experience with the larger network. All of the information you share with me today is confidential. I'd like to encourage you to be extremely honest today. Please share what you feel comfortable sharing. You can choose not to answer a question if you prefer not to. Do you have any questions before we get started? [ASK THEM IF IT IS OKAY TO RECORD SO THAT YOU DO NOT HAVE TO TAKE DETAILED NOTES DURING THE INTERVIEW]

First, I'd like to learn more about your background and your school.

- How long have you been teaching and where/what have you taught?
- Tell me about teaching style?
- Tell me a little bit about your school? How would you describe the students? The staff?

Next, I want to learn more about your experiences so far with the University networked improvement community, including the experience with the larger network.

- How did you decide to participate in it?
- When you think about your participation in the network, what experiences stand out to you? Tell me more.
- What aspects of improvement science do you feel like you've learned?
- How would you rate your understanding of the improvement science tools/methods that we've used so far? On a scale of 1 to 5, with 1 as no understanding and 5 as understand it very well.
    - Fishbone or causal analyses [ASK TO RATE, THEN WHAT CONTRIBUTED TO THE UNDERSTANDING]
    - Process maps [ASK TO RATE, THEN WHAT CONTRIBUTED TO THE UNDERSTANDING]
    - Driver diagram [ASK TO RATE, THEN WHAT CONTRIBUTED TO THE UNDERSTANDING]
    - PDSAs [ASK TO RATE, THEN WHAT CONTRIBUTED TO THE UNDERSTANDING]
- What else have you learned, if anything, from your participation in the network thus far? [STAY ON THIS QUESTION AND GET ALL OF THEIR LEARNINGS AND PROCESSES]
    - What **processes or supports** helped to facilitate that learning?
    - [IF NO] Why not? Was there anything that could have been done differently?
- To what extent is the PDSA process generating learnings for you? Can you give me an example(s) of what you're learning? [ONLY APPLIES TO THOSE WHO HAVE COMPLETED A PDSA CYCLE].

- Let's more talk about the PDSAs. I'd like to get a sense of your capacity around the different components of the PDSA. Before you started participating in the network what were your experiences, if any, with:
    - Experimenting with new ideas in the classroom
    - Working with data
    - Theory of change
    - Collaborating with other teachers
    - Documenting and reflecting

Now, let's talk about a little bit about the work at your school.

- Think back to the times where we met with you and worked on the fishbone diagram, and started developing a driver diagram, would you have been comfortable facilitating that meeting? Tell me more.

- How comfortable are you agreeing or disagreeing with colleagues in our school meetings?

- At your school, has anything changed since you started participating in the network? [PROMPT IF NEEDED: These could be your own practices, attitudes, collaboration, your colleagues, or something the math department does differently now.]

    - Has anything about the collaboration/interaction changed since participating in the network? Tell me more.

Has participating in the network caused you to reflect upon your own teaching? Tell me more. [IF APPLICABLE: WHERE/WHEN DID THOSE REFLECTIONS OCCUR?]

Is there anything else you would like to share?

That was the last question. Thank you.

Administrator Interview: Semi-structured Interview Protocol

Introduction: Thank you for meeting with me today. I'm hoping to learn more about you and your experience participating in the University networked improvement community, including your experience with the larger network. All of the information you share with me today is confidential. I'd like to encourage you to be extremely honest today. You can choose not to answer a question if you prefer not to. Do you have any questions before we get started? [ASK THEM IF IT IS OKAY TO RECORD SO THAT YOU DO NOT HAVE TO TAKE DETAILED NOTES DURING THE INTERVIEW]

First, I'd like to learn more about your background and your school.

- How long have you been a principal at this school?
- Can you tell me about how and why you became a principal?
- What subject did you teach before becoming a principal?
- Tell me a little bit about your school? How would you describe the students? The staff?

Tell me about your leadership model or style at this school.

Next, I want to learn more about your experiences so far with the University networked improvement community, including the experience with the larger network.

- How did you learn about the network? Why did you decide to participate in it?
- What is your vision for your school's participation in the network?
- When you think about your school's participation in the network, what stands out for you? Top of mind.

- Do you feel like you've learned anything from your participation thus far?
    - [IF YES] What have you learned? How did you learn it?
    - [IF NO] Why not? Was there anything that could have been done differently?

- Do you feel like your teachers learned anything from their participation thus far?
    - [IF YES] What have they learned? How do you think they learned it?
    - [IF NO] Why not? Was there anything that could have been done differently?

- Let's talk about the PDSAs. I'd like to get a sense of your school's experience the different components of the PDSA. Before you started participating in the network the school's experiences, if any, with:
    - Experimenting with new ideas in the classroom
    - Working with data
    - Theory of change
    - Collaborating with other teachers
    - Documenting and reflecting

Now, let's talk about a little bit about the work at your school.

- Think back to the times where we met with you and worked on the fishbone diagram, and started developing a driver diagram, would you have been comfortable and/or able to facilitate that meeting? Tell me more.

- At your school, has anything changed since you started participating in the network? [PROMPT IF NEEDED: These could be teacher practices, attitudes, collaboration, or something the math department does differently now.]
    - Has anything about the collaboration/interaction changed since participating in the network? Tell me more.
    - At your school, have any practices and/or structures changed since you started participating in the network?

Is there anything else you would like to share?

That was the last question. Thank you.

Appendix B

Network Meeting Observation Protocol

Observer: ______________________________	Observation Type:
Date/Time: Site/Setting:	
Purpose:	
Participants:	
Visual Map: (See picture)	
Describe the setting:	

Running record:

Appendix C

School Meeting Observation Protocol

Observer: ______________________________	Observation Type:
Date/Time:	Site/Setting:
Purpose:	
Participants:	

Was there a pre-meeting with principal? Yes No

If yes, please describe purpose:

What structures are missing or present?

What am I/we doing to facilitate the process?	What am I/we doing to teach process?

Individual attitudes towards this process?	Individual knowledge (missing or present)	Individual skills (missing or present)	Individual behaviors (missing or present)

Team culture	Team collaboration and dialogue (missing or present; effective or not effective)	Shared responsibility

Leadership	Other Thoughts/Themes

What is evidence of: attitudes?	What is evidence of: knowledge?	What is evidence of: skills?	What is evidence of: behaviors?	What is evidence of: active engagement?

How does school's organizational/contextual factors affect it's preparedness to successfully implement improvement science? (Anything that expands upon team and organizational?)

Anything around data collection and/or use?

Is any new knowledge being generated from the collaborative conversations? (Specifically the PDSA reflection.)

Appendix D

Network Surveys

Beginning of Year 1 Survey

Please complete the following survey. Your answers will help us assess your experience and knowledge of improvement methods, and how they are applied in a school context. We need your honest feedback. Please answer truthfully. There are no wrong answers. All of your individual answers are confidential.

Thank you for taking time to participate—your experience and opinions are extremely valuable!

Q1. Please select your school.

Q2. How long have you taught math or algebra at this school? (Teachers only)

This is my first year

1 -2 years

3 - 4 years

5 - 9 years

10 - 14 years

15 - 19 years

20 or more years

Q3. Before today, how familiar or unfamiliar were you with the following?

	1 Not at all familiar	2 A little familiar	3 Some-what familiar	4 Familiar	5 Extremely Familiar	I don't know what this is
Improvement science						
Professional learning communities (PLCs)						
Root cause analysis (e.g. 5 Whys)						
Driver diagrams						
Logic models						
Process maps						
Plan-Do-Study-Act (PDSA) cycles						
Systems thinking						
"System of Profound Knowledge"						

Q4. How confident or not confident are you that you can: (Teachers only)

	1 Not at all confident	2 A little confident	3 Somewhat Confident	4 Confident	5 Extremely Confident
Identify potential barriers to students' learning math/algebra concepts					
Make changes to your teaching or classroom practices that could improve student learning in math/algebra					
Test whether a change in your teaching or classroom practices resulted in an improvement of your practice					
Collect meaningful data to evaluate whether a change in your teaching or classroom practices resulted in an improvement of your practice					
Analyze data to evaluate whether a change in your teaching or classroom practices resulted in an improvement of your practice					

Q5. In your opinion, how useful or not useful is it to experiment with new teaching practices when it comes to your own teaching? (Teachers only)

0. Not useful at all Why?____

1. Slightly useful Why?____

2. Somewhat useful

3. Moderately useful

4. Extremely useful

Q6. In your opinion, how useful or not useful is it to <u>collaborate with other teachers to reflect</u> on improving practices when it comes to your own teaching? (Teachers only)

0. Not useful at all Why? ______________________________

1. Slightly useful Why? ______________________________

2. Somewhat useful

3. Moderately useful

4. Extremely useful

Q7. In the past school year (2016-17), how frequently or infrequently did the following occur?

	0 0 Times	1 1 to 2 Times	2 3 to 4 Times (About once every 3 months)	3 5 to 7 Times (About once every 2 months)	4 8 to10 Times (About once every month	5 More than 10 Times (More than once a month)	Not Sure or NA
Other teachers shared their effective teaching strategies with me							
I was encouraged by administrators to try new classroom practices or teaching strategies							
I received specific feedback from a coach or peer that helped me improve my teaching practice							
I received specific feedback from an administrator that helped me improve my teaching practice							
I worked side-by-side with administrators							
I met with other math teachers to discuss what helps students learn best							
I asked another teacher for help or feedback on my own teaching practices							

Q8. How much do you agree or disagree with the following statements?

	Strongly disagree	Disagree	Somewhat disagree	Somewhat agree	Agree	Strongly agree	NA
Teachers in my school take responsibility for improving the school							
Teachers in my school help each other do their best							
School administrators are supportive of trying new ideas even if they fail							
My school sets high standards for academic performance							
I can rely on other teachers for help when I need it							
I can rely on school administrators for help when I need it							
Teachers at my school treat me with respect							
Administrators at my school treat me with respect							
The school schedule allows sufficient time for teachers to collaborate							
New ideas are welcome at my school							

Q9. What routines, meetings, or other structures, if any, do <u>you</u> currently have in place to continually improve your own teaching practice? (Teachers only).

Q10. What routines, meetings, or other structures, if any, does <u>your school</u> currently have in place for teachers to <u>collaborate towards improving teaching practices</u>?

Q11. What structures would you recommend that the UNIVERSITY put in place to foster collaboration and/or shared learning <u>among school teams in the network</u>?

Q12. Any other comments about today or the network that you would like to share?

End of Year 1 Survey

Please complete the following survey. Your answers will help us assess your experience and knowledge of improvement methods, and how they are applied in a school context. We need your honest feedback. Please answer truthfully. There are no wrong answers. All of your individual answers are confidential.

Thank you for taking time to participate—your experience and opinions are extremely valuable!

Q1. Please select your school.

Q2. How long have you taught math or algebra at this school? (Teachers only)

- This is my first year
- 1 -2 years
- 3 - 4 years
- 5 - 9 years
- 10 - 14 years
- 15 - 19 years
- 20 or more years

Q3. How would you rate your knowledge and skills of the following?

	0 I don't know what this is	1 I can recall this concept, term, or tool	2 I under-stand this this concept, term, or tool and how it connects to our work	3 I can apply this concept, term, or tool with assistance from Univ. facilitator	4 I can apply this concept, term, or tool without assistance from Univ. facilitator	5 I can teach this concept, term, or tool to facilitate and lead our school team
Root cause analysis (e.g. Fishbone or 5 Whys)						
Process map						
Driver diagram						
Developing a change idea aligned to driver diagram						
Planning a Plan-Do-Study-Act (PDSA) cycle						
Implementing a change idea as part of PDSA cycle						
Collecting PDSA data to inform whether the change idea resulted in an improvement?						
Developing meaningful process measures to monitor the potential improvement of your own practices over time						
Developing meaningful outcome measures to monitor the potential improvement of student learning over time						

Q4. What have learned, if anything from your participation in the network this year? What processes and/or supports helped to facilitate that learning?

Q5. How confident or not confident are you that you can: (Teachers only)

	1 Not at all confident	2 A little confident	3 Somewhat Confident	4 Confident	5 Extremely Confident
Identify potential barriers to students' learning math/algebra concepts					
Make changes to your teaching or classroom practices that could improve student learning in math/algebra					
Test whether a change in your teaching or classroom practices resulted in an improvement of your practice					
Collect meaningful data to evaluate whether a change in your teaching or classroom practices resulted in an improvement of your practice					
Analyze data to evaluate whether a change in your teaching or classroom practices resulted in an improvement of your practice					

Q6. In the past school year (2017-18), how frequently or infrequently did the following occur?

	0 0 Times	1 1 to 2 Times	2 3 to 4 Times (About once every 3 months)	3 5 to 7 Times (About once every 2 months)	4 8 to10 Times (About once every month	5 More than 10 Times (More than once a month)	Not Sure or NA
Other teachers shared their effective teaching strategies with me							
I was encouraged by administrators to try new classroom practices or teaching strategies							
I received specific feedback from a coach or peer that helped me improve my teaching practice							
I received specific feedback from an administrator that helped me improve my teaching practice							
I worked side-by-side with administrators							
I met with other math teachers to discuss what helps students learn best							
I asked another teacher for help or feedback on my own teaching practices							

Q7. How much do you agree or disagree with the following statements? (School team = Those teachers from your school participating in the Univ. network with you.

	Strongly dis-agree	Disagree	Some-what disagree	Some-what agree	Agree	Strongly agree	Not sure	NA
All or most teachers in our school team took responsibility for improving their practices								
I can articulate my principal's goals for our school's participation in the Univ. network								
I spoke with my principal directly about my individual participation in the network and what supports I needed, if any								
My administrator provided the instructional support(s) that I needed to try change ideas in my classroom								
Other teachers from our school team treated me with respect during our Univ. network-related meetings at our school								
The school schedule allowed sufficient time for our school team to collaborate on network-related work.								
I felt like new ideas were welcome during our Univ. network-related meetings at our school.								

Q8. How comfortable or uncomfortable were you….? (School team = Those teachers from your school participating in the Univ. network with you.

	Extremely comfortable	Comfortable	Somewhat comfortable	Not comfort-able at all	NA
Asking other teachers questions at the Univ. network meetings.					
Offering suggestions to other teachers to push their thinking at the Univ. network meetings.					
Asking questions during our Univ. network-related meetings at your school					
Disagreeing with other school team teachers during our Univ. network-related meetings at your school					
Disagreeing with your administrator during our Univ. network-related meetings at your school					

Q9. How have your own practices changed, if at all, since participating in the network.

____

____

____

Q10. Any other comments about today or the network that you would like to share?

____

____

____

Appendix E

Interview Coding Framework

Network

- Network math vision
  - Math: Teachers connecting with network PD
- School knowledge
- Network meeting structures
  - Teachers as professionals
  - Meet off school campus
  - Share learnings
- Network Role
  - Hub's role
- Univ. learning IS
- Schools collaborating with University
  - University support
  - Facilitates building relationship among schools
  - Not telling them what to do
  - Listen to their needs
  - Importance of relationships
  - University drive the collaboration
- Network Challenges
  - Hard to miss day for meeting
  - Purpose unclear
  - Uneven participation
  - Different needs
  - Poor communication around network
  - Distracted at meetings
  - Go off on tangents
  - Smaller versus larger network
  - Practical vs theoretical

Organizational Capacity

- Structures vs no structures
- Pre-existing structures
- Lack of collaboration time
- Lack of common courses
- Leadership support
- Lack of leadership involvement
- Leadership style or philosophy
- School team context/culture

Individual Capacity

- Pre-existing math instruction capacity
  - Teacher vs student directed
  - Math teaching experience
  - Pre-existing math CONTENT knowledge
  - Pre-existing math INSTRUCTION knowledge
  - Traditional math teaching
- Teaching Style or Philosophy
  - Concerned about classroom management and discipline
  - Student deficit thinking
  - Type of Kid
- Pre-existing evaluation-related capacity
- Teacher Values and Attitudes
  - Network motivation
  - Value evaluation-related activities

        - Reinvigorated or "reset"
        - Stressing
        - Not value the network
    - Improvement science learned (What)
        - Process map
        - Language of IS
        - Collect and analyze meaningful data
        - Change ideas
        - PDSA
        - Driver Diagram
        - Fishbone/Root cause analysis

Improvement Science Learned

- Hinders improvement science learning
    - Attendance
    - Schools are hectic
    - Added work for teachers
    - Poor leadership
- Facilitates improvement science learning
    - Facilitation support
    - Set work time
    - Build in work time
    - Time and space
    - Set schedule or timeline
    - Network provided mtg time
    - Not enough time
- Embedded in-school work structures
- Structured collaborative format
    - Structured collaboration tools
    - Problem versus solutions
    - Understanding the problem
    - Value IS Tools
- Working with teachers from other schools
    - Relationships and trust
    - Validation from colleagues
    - "Not alone"
    - Enjoy working with other schools
    - Learn from other schools
    - Groupings among schools
- Clear and explicit expectations
- Strong leadership across partners
- Improvement science learned (How)
    - Teaching/coaching strategies
    - Making manageable
    - Self motivation
    - Working with other teachers (within schools)
    - Connection to math and classroom
    - Learn by doing
    - Consistency

Learnings Generated

- Moving away from deficit thinking
- New resources
- Evaluative thinking/process
    - More reflective
    - Processing information differently
- Reflecting on instructional practices
    - More than one way to do something

How kids learn math
More knowledgeable of curriculum
Consider other perspectives
Reconsider how/why did something

Adjusting instructional practices

New language
Math teaching practices
Task or tool

How generated

Math PD
Structured reflection
Parallel with math project
PDSA structures
Consistency

Appendix F

School Meeting Observation Coding Framework

What am I/we doing to facilitate the process?
What am I/we doing to teach process?
Individual attitudes towards this process?
- Not engaged in meeting
- Not wanting to authentically engage
- Not wanting to take day to go to network mtg.
- Impatient to action

Individual knowledge (missing or present)
- Lack of IS understanding
- Understand IS
- Has subject knowledge
- Missing subject knowledge

Individual skills/behaviors (missing or present)
Structures (missing or present)
Team culture
Team collaboration and dialogue (missing or present; effective
- Not asking questions (no inquiry)
- Effective dialogue
- Collaboration

Shared responsibility
Leadership
- Providing a support
- Facilitating
- Setting meeting structures
- Communication issue

What is evidence of: attitudes?
What is evidence of: knowledge?
What is evidence of: skills?
What is evidence of: behaviors?
What is evidence of: active engagement?
How does school's contextual factors affect it's preparedness
Collection of data
Use of data
Is any new knowledge being generated from the collaborative con
What is change idea? What is on their PDSA protocol?
Other emergent codes
- Need for Facilitator
- Different needs
- Trust
- Student deficit thinking
- Not completing PDSA form
- Competing with other priorities
- Skepticism of the process
- Other schools have same issues/interests
- They have questions
- Giving them say into the process
- Setting not conducive
- Going off topic or process
- Adjusting to their needs in the moment
- Lack of time concerns

REFERENCES

Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method and practice.* Boston, MA: Addison-Wesley.

Beach, D., & Pedersen, R.B. (2013). *Process-tracing methods: Foundations and guidelines*. Ann Arbor, MI: The University of Michigan Press.

Bennett, A., & Checkel, J.T. (2015). *Process tracing: From metaphor to analytic tool.* Cambridge, UK: Cambridge University Press.

Bill & Melinda Gates Foundation (2018). *Glossary: Request for proposal on networks for school improvement.* Retrieved from http://k12education.gatesfoundation.org/glossary-request-proposal-networks-school-improvement/#intermediary

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18* (1), 32-42.

Bryk, A., Camburn, E., & Seashore Louis, K. (1999). Professional community in Chicago elementary schools: Facilitating factors and organizational consequences. *Educational Administration Quarterly, 35 supplement*, 751-781.

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

California Department of Education. (2017a). K-12 Public school enrollment time series [Data file]. Retrieved from https://dq.cde.ca.gov/dataquest/

California Department of Education. (2017b). California assessment of student performance and progress: Smarter balanced assessment [Data file]. Retrieved from https://caaspp.cde.ca.gov/sb2017/Search

Carpenter, T. P., Fennema, E., Franke, M. L., Levi, L., & Empson, S. B. (1999). *Children's* *mathematics: Cognitively guided instruction.* Portsmouth, NH: Heinemann.

Christie, C. A., Inkelas, M., & Lemire, S. (Eds.). (2017). *Improvement Science in Evaluation: Methods and Uses. New Directions for Evaluation, 153.*

Cousins, B., & Earl, L.M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14* (4), 397-418.

Compton, D. W., Baizerman, M., & Stockdill, S. H. (2002). *Special issue: The art, craft, and science of evaluation capacity building. New Directions for Evaluation, 93.*

Deming, W.E. (1994). *The new economics for industry, government, education* (2nd ed.). Cambridge, MA: The MIT Press.

Dewey, J. (1963). *Experience and education*. New York: Macmillan.

Engelbart, D.C. (1992). Toward high-performance organizations: A strategic role for groupware. *Proceedings of GroupWare '92 Conference*. San Jose CA: Morgan Kaufmann.

Engelbart, D. (2004). *Augmenting Society's Collective IQ*. Retrieved from http://www.dougengelbart.org/about/vision-highlights.html

Fowler, Jr., F.J. (2009). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage.

George, A.L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge, MA: The MIT Press.

Gomez, L.M., Russell, J.L., Bryk, A.S., LeMahieu, P.G., & Mejia, E.M. (November 2016). The right network for the right problem. *Phi Delta Kappan*. Retrieved from http://www.kappanonline.org/right-network-right-problem/

Guba, E. G., & Lincoln, Y. S. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Program Evaluation*, *1986*(30), 73–84. doi:10.1002/ev.1427

Guggenheim, D., & Chilcott, L. (2011). *Waiting for "Superman"*. Hollywood, Calif: Paramount Home Entertainment.

Katz, S., & Earl, L. (2010). Learning about networked learning communities. *School Effectiveness and School Improvement*, *21:1*, 27-51. DOI: 10.1080/09243450903569718

Labin, S., Duffy, L., Meyers, D., Wandersman, A., & Lesesne, C. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation 33(3)*, 307-338.

Langley, G.J., Moen, R.D., Nolan, K.M., Nolan, T.W., Norman, C.L., & Provost, L.P. (2009). *The improvement guide: A practical approach to enhancing organizational performance* (2nd ed.). San Francisco, CA: Jossey-Bass.

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.

Louis, K. S. (2006). Changing the culture of schools: Professional community, organizational learning, and trust. *Journal of School Leadership, 16*, 477-489.

Louis, K. S., & Lee, M. (2016). Teachers' capacity for organizational learning: The effects of school culture and context. *School Effectiveness and School Improvement*, *27:4*, 534-556. DOI: 10.1080/09243453.2016.1189437

Merriam, S.B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Merriam, S.B., & Bierema, L.L. (2014). *Adult learning: Linking theory and practice*. San Francisco, CA: Jossey-Bass.

Moen, R.D., & Norman, C.L. (2010). Circling back: Clearing up myths about the Deming cycle and seeing how it keeps evolving. *Quality Progress, November 2010*. Retrieved from http://asq.org/quality-progress/2010/11/basic-quality/circling-back.html

National Council of Teachers of Mathematics (NCTM) (2014), *Principles to Actions: Ensuring Mathematical Success for All.* Reston, VA: The National Council of Teachers of Mathematics, Inc.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation 29:4,* 443-459.

Proger, A. R., Bhatt, M. P., Cirks, V., & Gurke, D. (2017). *Establishing and sustaining networked improvement communities: Lessons from Michigan and Minnesota*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

Rich, M. (2012, July 23). Enrollment off in big districts, Forcing Layoffs. *The New York Times*. Retrieved from https://www.nytimes.com/2012/07/24/education/largest-school-districts-see-steady-drop-in-enrollment.html

Ripley, Amanda. (2008, December). How to fix America's schools. *Time.*

Rohanna, K. (2017). Breaking the "adopt, attack, abandon" cycle: A case for improvement science in K-12 education. In C.A. Christie, M. Inkelas & S. Lemire (Eds.). *Improvement science in evaluation: Methods and uses*. *New directions for evaluation, 153*, 65-77.

Rose, M. (2016). Reassessing a redesign of community colleges. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/views/2016/06/23/essay-challenges-facing-guided-pathways-model-restructuring-two-year-colleges

Russakoff, D. (2014, May 19). Schooled. *The New Yorker*. Retrieved from http://www.newyorker.com/magazine/2014/05/19/schooled

Russell, J., Bryk, A., Dolle, J., Gomez, L., Lemahieu, P., & Grunow, A. (2017). A framework for the initiation of networked improvement communities. *Teachers College Record, v119 n5*. Retrieved from http://www.lrdc.pitt.edu/BOV/documents/NIC%20Initiation%20Framework_Russell_TCR_FINAL%20SUBMISSION.pdf

Saldaña, J. (2013). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage Publications Inc.

Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York, NY: Basic Books.

Seidman, I. (2013). *Interviewing as qualitative research*. New York, NY: Teachers College Press.

Senge, P.M. (2006). *The fifth discipline: The art & practice of the learning organization* (Rev. ed.). New York, NY: Doubleday.

Sowers, N. & Yamada, H. (2015). *Pathways Impact report: Three years of results from the community college pathways*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved from http://www.carnegiefoundation.org/resources/publications/pathways-impact-report-2013-14/

Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: A review of the literature. *Journal of Educational Change 7:4*, 221-258. DOI:10.1007/s10833-006-0001-8

Stoll, L., & Louis, K.S. (2007). *Professional learning communities: Divergence, depth, and dilemnas*. New York, NY: Open University Press.

Torres, R.T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation, 22*, (3), 387–395.

Walton, M. (1986). *The Deming management method.* New York, NY: Berkley Publishing.

White, C.M., Statile, A.M., Schoettker, M.S., Solan, L.G., Unaka, N.I., Vidwan, N., Warrick, S.D., Yau, C., & Connelly, B.L. (2012). Utilizing improvement science methods to improve physician compliance with proper hand hygiene. *Pediatrics*, *129*(4), e1042–e1050. http://doi.org/10.1542/peds.2011-1864.

Williams, C.E. (2015). *Utilizing improvement science to improve population health* [PowerPoint slides]. Retrieved from http://www.ihatoday.org/uploadDocs/1/williamsfull.pdf

Yin, R.K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage.