Evaluating the Impact of Statewide Supports to Reduce Within-School Achievement Gaps: A Mixed Methods Study of Focus Schools in Michigan

by

Monica P. Bhatt

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Educational Studies)
in the University of Michigan
2017

Doctoral Committee:

Professor David K. Cohen, Chair
Professor Susan M. Dynarski
Professor Ann Lin
Professor Donald J. Peurach

To my parents, Pankaj and Charu Bhatt,
and to countless immigrant parents like them,
whose unwavering faith in education inspired this journey.

And to my husband, Srinu Kusuma,
whose unwavering faith in me helped me complete it.

Acknowledgements

If this dissertation sees further, it is because it stands on the shoulders of giants. Thank you to my committee—David, Sue, Ann, and Don—who have collectively challenged, nurtured, and sculpted me into the scholar I am today. If there are faults or gaps in my knowledge or skills it is due to my own stubbornness. David, thanks to your teaching, I can no longer unsee the pervasive challenges to our educational system—and I count this is as a good thing. Perhaps more importantly, you have taught me by example how to do this work and still be the person I want to be. Thank you for your life lessons and for being my teacher, mentor, and friend. Thank you to Sue, who is the smartest person I know, who is an excellent teacher, who introduced me to the intellectually satisfying world of causal inference, and who has challenged me to get out of my comfort zone in ways that are difficult but rewarding. Thank you to Ann, who has literally opened windows to the world for me, whether at Republican primaries in Ypsilanti, Michigan or meeting with Sheiks in Qatar. Ann, I always learn something new when I speak with you or at least learn to look at the world in a different way; I will take that with me always. And if I forgot how to do so, rest assured that I will seek you out to refresh my memory over a cup of tea! Don is the reason that this dissertation exists, whether he realizes it or not. His kindness, generosity, support, and analytic thinking helped me navigate the river he told me about that fateful day at Sweetwater, and his connection to the world of continuous improvement has bolstered a career avenue for me in myriad ways.

There are a number of scholars with whom I have worked over the years whose influences on me and my work show up in this dissertation in ways big and small. They are: Lori Hill, Valerie Lee, Brian Jacob, Simona Goldin, Michal Kurlaender, Liz Gerber, Tom Smith, Laura Desimone, Claire Smrekar, and Jim Guthrie. Thank you to my friend-tor, Julie Kochanek, for her foresight that I might find a home in the intersection of improvement science and school reform and for the gift of leading the Michigan Focus NIC.

To the members of the Michigan Focus NIC, my interview subjects, and my colleagues at the Michigan Department of Education: thank you for your access, your trust in me, and your faith in the process. Thank you for enabling us to do meaningful work together.

Thank you to the Rackham Graduate School for the Rackham Merit Scholarship, the School of Education for its funding and support, and the American Institutes for Research for its funding and flexibility.

To my wonderful Ph.D. friends, thank you for your support, encouragement, camaraderie, commiseration, and couches when I came in from out of town: Sarah Fick, Annick Rougee, Anna Arias, Amy Pond, Timm Betz, Lagina Gause, Brittney Chesworth, Jean Mrachko, Stef Iwashyna (and her family), Florencia Gomez Zacarrelli, Becky Christensen, Andy and Michelle Kwok. A very heartfelt thank you to Emily House, who is my Ph.D. soulmate and the only one who really knows; Ginny Pruitt, who not only believed I could climb mountains but showed me how; and Rachel Basse, who held my hand through the finish line.

My family, both immediate and extended, epitomize love and support and never questioned my abilities to achieve this milestone, which is a gift in itself. Thank you Mom & Dad, Parul & TJ, Shivan, Yamuna Aunty & Srihari Uncle, Jason & Sravanti, and Meera & Will for reminding me what is most important in life.

Finally, thank you to my truly incredible husband Srinu, whose intelligence, work ethic, and compassion set a high bar for me on a daily basis and whose love and undeterred belief in me help me reach that bar on occasion.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

In this set of papers, I examine whether accountability-driven, state-led efforts to reduce within-school inequality, as measured by student achievement, have been successful. In doing so, I first examine the extant literature on school improvement to provide historical and policy context for school improvement efforts and then discuss why and how school improvement efforts with the specific aim of reducing within-school inequality may differ from those aimed at improving overall student achievement. In Chapter 2, I use a quasi-experimental regression discontinuity design to determine the causal impact of a bundle of services offered to schools in Michigan to narrow the within-school achievement gap in student achievement. In Chapter 3, I use a qualitative case comparison design to determine what organizational features of school supports were influential in the implementation of the treatment. In Chapter 4, I explore the potential of improvement science methods to achieve the aims of school improvement using a case study of a Networked Improvement Community across two mid-sized cities in Michigan.

Introduction

"As the mounting evidence on the weak effects of No Child Left Behind illustrates, it is extraordinarily difficult to design accountability systems that take into account the intense challenges of educating high concentrations of low-income children (Dee and Jacob 2011). There will be much to learn from the alternative accountability systems put in place by states that have been granted NCLB waivers."

– Greg Duncan and Richard Murnane, *Restoring Opportunity,* 2014

After the passage of No Child Left Behind (NCLB) in 2001, the congressional reauthorization of the Elementary and Secondary Education Act (ESEA) of 1965, the law came under increasing scrutiny for its prescriptive and punitive provisions. Consequently, in order to provide flexibility to states in achieving the goals set forth by NCLB legislation, the U.S. Department of Education offered ESEA flexibility waivers to states in exchange for state-developed plans that specified the way states would "improve educational outcomes for all students, close achievement gaps, increase equity, and improve the quality of instruction."[1] In September 2011, the Obama administration announced a formal process through which states were able to apply for a three-year ESEA waiver with a one-year extension provided to some states for the 2014-15 school year. As of April 2015, 45 states had applied for ESEA waivers and 43 states had been approved.[2]

[1] http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html

[2] Though the regulations set forth through this interim waiver application process will only continue through fall 2017 based on the most recent reauthorization of the ESEA, the Every Student Succeeds Act (ESSA) of 2015, the priorities reflected in the waivers have been codified in the new law, including the requirement to identify high achievement-gap schools and provide a statewide system of supports to ameliorate those gaps (Pub.L. 114–95).

Until the passage of the Every Student Succeeds Act in 2015, states were required to submit a state-developed accountability plan in exchange for leeway on NCLB accountability provisions. Among the many requirements that had to be addressed in the state-developed accountability plans, state recipients of ESEA waivers had to identify Priority Schools (lowest performing schools), Reward Schools (highest performing or highest progress schools), and Focus Schools (schools with the largest achievement gap). Schools that are identified as Priority or Focus Schools are given additional interventional support to improve overall student achievement and/or reduce the level of within-school equality. Together, this new system of school identification and supports constitutes the most recent iteration of the statewide accountability system.

In Michigan, per ESEA waiver requirements and guidance, Focus Schools are identified as the 10% of schools that have the largest gap between the highest performing 30% and lowest performing 30% of students, as measured by a composite of student achievement test scores by subject and grade. Schools that are designated as Focus Schools remain in a cohort of Focus Schools for four years, regardless of whether they are identified in subsequent years as falling in the 10% of schools with the largest within-school achievement gap. If schools are identified in subsequent years, the four-year cycle begins again. The first Focus School cohort was identified using Michigan's Top-to-Bottom Ranking methodology in the 2011-12 school year.[3] Focus School cohorts have been identified in each subsequent year, with the third and latest cohort identified for the 2013-14 school year. In 2014, 346 schools received the Focus designation; of

[3] The Top-to-Bottom list is part of Michigan's school accountability system, which ranks schools on their student performance in mathematics, reading, writing, science, and social studies and (for high schools) graduation rate data. School performance components include student achievement, improvement, and achievement gaps between the highest and lowest scoring 30% of students in each school. Additional information on the Top-to-Bottom ranking methodology can be found here: http://www.michigan.gov/mde/0,4615,7-140-22709_56562---,00.html.

these, 110 schools have been identified as Focus Schools for the first time, while 133 of these schools received the Focus designation in all three years (2011-12, 2012-13, and 2013-14).

As a consequence of being identified as a Focus School, schools are required to participate in a "data dialogue" facilitated by district personnel in which school and district staff review student- and school-level data and identify two areas of teaching and learning priorities that are likely to close the achievement gap. Schools must include these teaching priorities in their revised School Improvement Plan (SIP) and submit to monitoring by the district. District personnel must facilitate the data dialogue and also revise the District Improvement Plan (DIP) to specify how the district will support schools in implementing the identified teaching and learning priorities. District personnel in Title I districts are assisted by the state-funded Title I District Improvement Facilitator, who provides technical assistance and models the use of MISchool data to spur improvement.

Given the importance of the policy goals driving identification of Focus Schools (i.e., the reduction of within-school inequality), it is critical to examine whether the provision of supports to reduce within-school inequality are efficacious. Chapter 2 demonstrates that supports provided in Michigan likely did not reduce within-school inequality or raise average school achievement during the first two years. This policy, then, despite being nearly ubiquitous in U.S. states for the past five years, may not be having the effect written into its policy goals in states besides Michigan as well. However, there are two caveats to this finding. First, the study is underpowered and therefore may not be able to detect real effects at the school level. Second, average effects of any kind—negative, null, or positive—are likely masking heterogeneity in treatment effects at the school-level that may provide useful insight into what is required for such a policy to be efficacious and how those requirements can be replicated in different contexts. The

question then morphs from "Does this policy work?" to "How does it work or not work in various contexts?" This review of the literature on implementation discusses how this question has developed from more than five decades of research on policy and program implementation.

Chapter 3 indicates that the interventions detailed above comprise a loose bundle of solutions to the problem of within-school inequality. Taken together, the treatment is relatively weak in its prescription to ameliorate achievement gaps. Consequently, its success depends in large part on the capacity of the organizations and individuals tasked with executing the interventions. This local capacity was first identified in the implementation literature as critical, but lacking (see, e.g., Pressman & Wildavsky, 1973); more recent studies complicate this picture of a simplistic principal-agent relationship by examining the components of capacity to implement school improvement reforms and the ways those components play out in complex policy environments (see, e.g., Honig, 2006). This study shows that it is not only capacity but *coordinated capacity*—a systemic effort to harness individual and organizational capacity—that is important to achieving gains in student achievement for the bottom 30% of students.

Chapter 4 turns to the puzzle of how that coordinated capacity can be developed and harnessed to achieve the policy goals set forth in ESEA legislation. In this study, I examine one potential solution to this capacity problem, a Networked Improvement Community approach to school improvement in Focus Schools. A Networked Improvement Community connects individuals and organizations from various contexts to answer not only "What works?" but "What works, for whom, and in what context?" (Bryk, Gomez, Grunow, & LaMahieu, 2015). This approach, which stems from various disciplines including healthcare and systems engineering, uses a systematic, data-driven approach to inform practice. It has been touted as a potential solution to building and sustaining capacity in local entities, which has been shown to

be a mediator of policy success in large-scale educational reform (Smylie & Evans, 2006). Chapter 4 provides a case study of how such a community can be initiated and sustained in the context of federally-legislated, standards-based reforms and probes the extent to which such an approach can deliver on the premise of capacity building for school improvement.

Taken together, this set of studies asks and answers whether identification of Focus Schools is sufficient to reduce achievement gaps; how coordinated capacity at each level of governance—state, ISD, district, and school—influences policy implementation; and how a Networked Improvement Community approach may provide a potential solution to building this coordinated capacity within states. The findings have implications for the way schools and districts in particular structure improvement efforts, particularly under the auspices of the Every Student Succeeds Act (ESSA), the most recent authorization of the ESEA, which devolves considerable responsibility for implementation of policy goals to states.

References

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.

Duncan, G. J., & Murnane, R. J. (2014). *Restoring opportunity: The crisis of inequality and the challenge for American education*. Cambridge, MA: Harvard Education Press.

Honig, M. (Ed.). (2006). *New directions in education policy implementation: Confronting complexity*. Albany, NY: State University of New York Press.

No Child Left Behind Act of 2001, P.L. 107-119. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/pg1.html#sec1001

Smylie, M., & Evans, A. (2006). Social capital and the problem of implementation. In M. Honig (Ed.), *New directions in education policy implementation: Confronting complexity*. Albany, NY: State University of New York Press.

Pressman, J.L., & Wildavsky, A. (1973). *Implementation.* Berkeley, CA: University of California Press

Chapter 1

School Accountability and Improvement for Equitable Outcomes: Design, Implementation, and Evaluation

"Fair, flexible, and focused accountability and support systems are critical to continuously improving the academic achievement of all students, closing persistent achievement gaps, and improving equity."

— The U.S. Department of Education, ESEA Flexibility Waiver Announcement, 2012

The history of federal U.S. educational reform can be construed as simultaneously chasing two separate but related policy goals: equity and excellence (Cohen & Moffitt, 2009); that is, U.S. education policy has been written in a way that prioritizes fostering students' abilities to meet rigorous educational standards or ensuring all students have the opportunity or resources to do so, or both. The primary approach of policymakers for the past 25 years has been to implement a system of standards-based accountability to achieve both of these policy goals. Research shows that educational attainment and performance have improved over this time period (Kena et al., 2015). However, systematic achievement gaps in educational performance and attainment remain between high- and low-income students, students of differing racial and ethnic backgrounds, students with disabilities and those without, and native English speakers and English learners, among other subgroups. These achievement gaps are connected to growing wealth inequality within the U.S. and have adverse implications for outcomes through a child's life including, but not limited to, educational attainment, income, marriage, and intergenerational social mobility (Reardon, 2011; Duncan & Murnane, 2011; Duncan & Murnane, 2015;

Bradbury, Corak, Waldfogel, & Washbrook, 2015). This chapter explores the design, implementation, and evaluation of efforts to reduce these persistent achievement gaps to date and the implications this body of literature has for future reform efforts. Of primary importance is a fundamental question: Can the U.S. educational system be designed in such a way that ameliorates these achievement gaps, and if so, how?

The design of U.S. educational reform has centered on using levers such as standards and assessment to work around local control of U.S. schools (Cohen & Moffitt, 2009). Consequently, implementation of policies has been dependent on the capacity of individuals and organizations to conduct the work at hand and on the efficacy of the instruments that policies deploy (Cohen, Peurach, Glazer, Gates, & Goldin, 2014). Evaluations of accountability policies have shown that these efforts have demonstrated varied effects by grade and subject at best (see, e.g., Dee & Jacob, 2011; United States Department of Education, 2010) and inspire adverse policy effects such as cheating at worst (Jacob & Levitt, 2003). The story of educational reform to support equitable outcomes for students, then, is one in which whole-school reform is not systematically able to bring about desired changes in varied contexts in a desirable, replicable manner. This chapter explores reasons for this lack of efficacy and replicability and the potential for continuous improvement efforts to bring about those desired changes.

Policy Design: Standards-Based Accountability as a Policy Lever

Standards-based accountability systems were conceived of amidst increasing concern about the inequality of the American public school system and the improvement of education for all children (Cohen & Moffitt, 2009). The primary assumption of these systems is that manipulating structural and organizational elements can lead to improvements in student learning. This strategy rests on three-prongs: standards, testing, and accountability. Taken

together, these components represent the parameters of school improvement. This three- pronged approach seeks to impose structure on a system that is otherwise characterized by politically fragmented governance, deference to local authorities in decision making, and marked inequalities in resources (Cohen & Spillane, 1992).

Standards-based accountability systems operate as follows: First, schools are expected to communicate the measurable criteria set forth by standards to their students through teaching and learning (Figlio & Loeb, 2011). Then, students are assessed as to whether or not they have retained the content set forth in these standards. Finally, based on these assessments, schools and districts are held accountable with respect to gains on the assessment and/or a minimum level of proficiency set by the state (Smith & O'Day, 1990).

In this system, accountability can be explicit (i.e., through sanctions, school closing, or bonuses) or implicit (i.e., community pressures, local actors' decision making, real estate market responses) but is ultimately dependent on information dissemination as a mechanism for distributing rewards or sanctions (Figlio & Loeb, 2011). School improvement strategies and models are also devised and disseminated as mechanisms for improvement for schools in corrective action, i.e., those schools that failed vis-à-vis the accountability system set up by the state (Cohen & Moffitt, 2009).

Accountability regimes have been a primary policy instrument in U.S. education for over a decade, with the intellectual roots of standards- and assessments-based reforms beginning in the early 1990s (Smith & O'Day, 1990). First designed nationally as the Improving America's Schools Act of 1994 and Goals 2000, accountability systems rest on the assumption that information dissemination on measured outcomes and monitoring of explicit and/or implicit rewards and sanctions will lead to compliance by schools and districts (Figlio & Ladd, 2007).

However, there are four potential failures of such accountability systems in resulting desirable policy goals. First, there may be improvement on measured outcomes, but these measured outcomes may only represent a small subset of stakeholder objectives and therefore not lead to generalizable improvement in life outcomes such as graduation and postsecondary attainment. Second, schools and/or districts may engage in strategic behavior that subverts the aims of the accountability system (see, e.g., Jacob & Levitt, 2003). Third, schools and districts may not have the capacity to respond to such incentives (Figlio & Loeb, 2011). A fourth related issue may be that policies do not deploy instruments that could help implementers become more effective, such as coaching and training manuals.

Despite these potential failures, the standards-based reform movement has been successful in several ways. Most relate to the dramatic paradigm shift, which they represented in the purpose and organization of the American education system, specifically to achieve two separate but intertwined goals: excellence and equity. The emphasis of these reforms on accountability at every level of the system based on some previously agreed-upon state-level measures of proficiency signaled a change from the prior ad hoc, local manner in which schools had been judged. Furthermore, the development of a consensus on required content for particular subject areas, though often fraught with controversy, was also a departure from historical precedent, which largely had left such decision making up to teachers on a classroom basis. Finally, the focus on student learning outcomes as a primary criterion by which units within the school system should be judged shifted the onus of responsibility from students as recipients of knowledge and skills within schools to teachers and administrators as facilitators of learning for all students.

The Role of Capacity in Policy Implementation

Policy implementation in education might be thought of as the conceptual space in which the proverbial rubber of ambitious educational reforms meets the road of educational practice. Implementation often refers to a tangled web of institutional capacity, political will, organizational resources, governance structures, dissemination of information, human capital, talent, and individual cognitive sense-making by which policy goals are translated into practice; this then results in some set of desired or undesired outcomes that may or may not have been intentional. Over the past four decades, much attention has been given to the importance of policy implementation for educational outcomes with increasing specificity about what particular dimensions of implementation are crucial for the translation of policy goals into practice.

The literature to date delineates four waves of policy implementation research (Odden, 1991; Honig, 2006; Young & Lewis, 2015). The first wave defined the "problem of implementation" and introduced broad parameters for studying the problem, drawing primarily from the principal-agent theory of monitoring and compliance. In this wave, implementation of large-scale federal policies was seen largely as a failure, and the burden for this failure was placed on local practitioners who lacked, in broad strokes, either the will or the capacity (or both) to implement federal policies (Bardach, 1977). For example, in one of the first studies of the ESEA, Murphy (1971) discussed constraints as the extent to which federal policy could "redirect" local efforts. The second wave of researchers demonstrated that implementation was feasible. Researchers during this era focused on micro-processes, such as individual incentives and beliefs, which may help to explain variation in policy implementation and account for pockets of implementation (Berman & McLaughlin, 1979). Researchers' concerns during this era still centered on how to achieve fidelity of implementation—that is, whether and how well

policies and programs adhere to policy intentions and programmatic features over time (Elmore, 1983; Sabatier & Mazmanian, 1980). The third wave of implementation research has focused on integrating the macro world of policymakers with the micro world of individual implementers in order to determine what works in the implementation of an education policy (Honig, 2006). This strand of research has focused on "mutual adaptation" of policies—that is, not only do practitioners make sense of policies that are handed down to them, but the act of sensemaking turns practitioners into policy brokers; hence the act of policy interpretation exerts force on the macro level policy itself to affect change. If the third wave was focused on the question of "What combination of conditions and actions influences what works?" in the implementation of education policy, then the fourth wave is concerned with "What works, for whom, and under what circumstances?" (Cohen-Vogel et al., 2015). This fourth wave is the focus of the networked improvement discussed in Chapter 4 and provides the potential to build the coordinated capacity discussed in Chapter 3.

Wave One: The Principal-Agent Problem and the Role of Information and Monitoring

Early theories of implementation often cast implementation as a "problem" of the principal-agent relationship in which the agent—here, district administrators, school principals and teachers—did not comply or act in accordance with the explicit directives of the principal, or the policymaker (Pressman & Widalsky, 1973; Bardach, 1977). The result was often the thwarted aims of the initial policy or legislative principles, as specified by those making policy choices, attributed to agents on the ground—the implementers. During the first wave of implementation research, this principal-agent framing was used to define the broad parameters of the "implementation problem," which was largely deemed a problem of non-compliance on behalf of policy implementers based on the desired framework produced by policy architects

(McLaughlin, 1987). In an accountability context, this may be translated as a monitoring problem. If stakeholders—including parents and policymakers—cannot monitor the activities of schools then it follows that practitioners might act out of accordance with said stakeholders' wishes. In this characterization, information is the mechanism by which monitoring may occur with explicit and/or implicit rewards attached to compliance (Figlio & Loeb, 2011). In this characterization of implementation research, we could suppose that the core technology of teaching and learning was known and that individuals had perfect information; they just needed to be held accountable to act upon that information.

Wave Two: Street-Level Bureaucrats and Local Actors as Policymakers

Michael Lipsky shifted the parameters of this idea in his famous rendering of "street-level bureaucrats" in which he argued that public service bureaucrats have wide discretion in the way policy is implemented, and therefore, make many policy decisions in the course of implementation (1980). This view of agents not merely as implementers but as policymakers themselves with capabilities to make decisions that define the policy itself has expanded over the course of the past two decades. In this generation of research, factors that influence implementation were conceived to be size, intra-organizational relations, commitment, capacity, and institutional complexity (McLaughlin, 1987). As policy problems were variable in nature, it made sense, then, that the solutions might be variable as well. This variability depended on capacity, thought of as money, training and expertise or knowledge, and the will or motivation of individuals and institutions.

A cognitive strand of this research focuses on the way actors, specifically those individuals thought of as implementers by early implementation researchers—in this case district administrators, school leaders, and teachers—"make sense" of policies based on their prior

experiences, knowledge, and role within the institution (Spillane 2005; Coburn 2005). This line of research explores how motivation or will is influenced by individuals, and later, by institutional conditions. The focus here is on attitudes, motivation beliefs, and knowledge—their origin and the consequences of those characteristics in an implementation context. Change, then, is ultimately "a problem of the smallest unit" (McLaughlin, 1987). However, whether the money, training, expertise or knowledge embodied by capacity can be manipulated to then change attitudes, motivation, and will remains an open question. This line of research has been useful in describing the idiosyncratic variation in policy implementation that we often see in education research; the next step in this line of questioning was then to examine the patterns and institutional features that then might affect these attitudes and beliefs.

Wave Three: Micro-Level Processes as Nested in Political Contexts

The adaptation or interpretation of policy tenets by individual actors, as highlighted above, occurs within specific educational contexts and often results in policy interpretation in new and unforeseen ways; this reinvention of the policy in the local context, in turn, redefines the policy at the macro level through a feedback loop. In this way, policy is iterative. This idea is not new; Berman and McLaughlin's classic work on the Rand Change Agent Study introduces the idea of "mutual adaptation" between users—i.e., policy implementers—and institutions (1979). Similarly, Datnow and colleagues talk about the "co-construction" of policy across levels of governance (Datnow, Hubbard, & Mehan, 2002). These perspectives refute the notion of a principal-agent relationship in the technical-rational tradition in which implementers are seen as an impediment to policy goals (Bardach, 1977). Instead they are seen as potential key supporters in affecting change spurred by the policy design, particularly if they find the policy or program suited to their purposes.

Implementation in this understanding includes not only what occurs in the target organizations by targeted actors but the interplay of the politics of the policy between various levels of governance or institutions. This redefinition of policy in the implementation phase is markedly different from policy appropriation, where local contexts "appropriate" the parts of district, state, or federal law that offer the most congruence to their local ideologies, priorities, and goals. In co-construction, the expectation is that subordinate governance structures are partners in defining the parameters of a new reform initiative, such as college preparatory for all curricula. A similar argument can be made for institutions. The role of institutions and contexts are not only important because they "infuse the policy system with presumptions, preferences and prejudices that advantage some and disadvantage others" (Malen, 2006), but because they interface in a mutually adaptive manner to redefine the "'norms, rules, regulations, and definitions of the situation' that affect how actors think and behave" to enact reform (Rowan & Miskel, 1999).

Wave Four: Improvement Science as a Methodology for Implementation

It has been suggested that improvement science, also referred to as continuous improvement, may represent the fourth wave of implementation research (Cohen-Vogel, et al., 2014). There is a long tradition of using improvement science in diverse fields, such as engineering and healthcare, to address problems of practice (see, for example, Deming, 1993; Gawande, 2007). Although its primary origins can be traced to researchers' use of scientific inquiry to improve industrial manufacturing processes, the principal aim—to improve practice—has led to the rapid dispersion of this highly practical methodology across multiple disciplines (Langley et al., 2009). Perhaps unsurprisingly, then, there is a growing interest in using improvement science to address some of the most pressing problems of practice in education as

well (Bryk, Gomez, Grunow, & LeMahieu, 2015; Lewis, 2015; Sparks, 2013).

An important aspect of improvement science in education is the importance of collaborative research partnerships, defined as long-term working relationships between researchers and practitioners whose purpose is to study and improve upon problems of practice (Coburn, Penuel, & Geil, 2013; Kochanek, Scholz, & Garcia, 2015). One such partnership, a Network Improvement Community (NIC), has been touted as a way to better understand what works, for whom, and under what conditions (Bryk et al., 2015). NICs are networks of organizations not tied to specific geographic locations or entities that strive to leverage expertise and working relationships across contexts to address problems of practice through design, implementation, testing, and redesign. These improvement communities are characterized by four salient features (Bryk et al., 2015). They are

- focused on well-specified, common aims;
- guided by shared working theories of the systems and how they can be improved;
- disciplined to apply improvement science methods and measures to spur improvements in testable iterations, such as rapid PDSA (Plan-Do-Study-Act) cycles; and,
- organized to share and integrate practices and processes developed within their respective networks across other contexts

NICs use PDSA cycles to drive continuous improvement (Langley et al., 2009). The cycle highlights four principal questions guiding improvement research: (1) how to understand the present problem; (2) what aspect(s) of the problem should be addressed by the improvement process; (3) what changes need to be introduced; and (4) how to know the changes are improvements. This cycle was echoed later in Englebart's (2003) three-level activities in the

Networked Improvement Community. According to Englebart, institutions conduct primary business (A-level activity); work collaboratively with research units to collect data (B-level activity); and then develop solutions to problems and strategies for improvement (C-level activity). Englebart observed that as the three activities are being carried out, the institutions' understanding of the problem as well as both inner- and inter-institution capacity for improvement will be strengthened. More recently, researchers have applied this similar concept in improvement work in the field of education. This model for improvement has been utilized in addressing educational problems (e.g., Bryk, Gomez, & Grunow, 2011). The improvement model emphasizes three guiding questions, mirroring Langley's principal questions (2–4) and uses the PDSA cycle to test changes (Langley et al., 2009). The focus on practice-level questions and improvement indicate the potential for this approach to represent the next phase in implementation research—that is, it provides a potential method to answer the question "What works, for whom and under what circumstances?"

The Role of Capacity in Policy Implementation

The publication of James Coleman's *Equality of Educational Opportunity Survey* in 1966, which examined the relationship between school-based resources and student performance, began a decades-long debate about whether school-based resources "matter" for student performance. The connection between school-based resources and desirable student outcomes, previously assumed to be positive, was established as tenuous (see, e.g., Hanushek, 1989; Jencks, et al., 1972), though a widely regarded meta-analysis by Hedges, Laine, and Greenwald (1994) showed a modest effect of monetary inputs on students' standardized scores. Additionally, more recent methodological advances have allowed for the re-analysis of older data sources and settled some of the questions around what Cohen, Raudenbush, and Ball (2003) label "conventional

resources," such as teachers' formal qualifications, books, facilities, class size, and time. For example, though there has been controversy over whether reductions in class size benefits students in past studies, Dynarski, Hyman, and Schanzenbach (2013) use detailed administrative data on college-going to show moderate effects on postsecondary enrollment as a result of random assignment to smaller class sizes, with larger effects for minority and low-income students.

Concurrent with questions about whether inputs matter was a line of questioning that focused on organizational structures within and among departments, schools, and districts and the relationships between individuals that were associated with desirable student outcomes. From this work, education reformers garnered the importance of teachers' collective responsibility, academic press, and professional learning community to desirable student outcomes (see, e.g., Bryk, Lee, & Holland, 1993; McLaughlin & Talbert, 1999). The emphasis on the mobilization of human and social capital to policy implementation is also well documented in the policy implementation literature. For example, evidence from a case study analysis of standards-based reform in Massachusetts about the implementation of the Massachusetts Education Reform Act (MERA) suggests that "building an implementation regime requires that people, resources, and institutions be connected via formal and informal organizational networks" and that this dimension of capacity is distinct from conventional resources that the state might provide, such as financial and material supports (McDermott, 2004). In addition, this dimension of capacity might be what Cohen, Raudenbush and Ball (2003) term "personal resources," such as practitioners' will, skill, and knowledge. In this line of questioning, researchers are concerned with the ways individuals understand and interact with policy mandates, curricular materials, and the instructional challenges present in relating specific subject-specific content.

The extent to which conventional resources matter is still an important question; however, the positioning of the question has changed to probe the context(s) within which these conventional resources matter, the mechanisms by which they work, and the match between resources and desired goals. While it is pertinent to think about the way financial, material, human, social, organizational, and relational resources constitute capacity to accomplish a specific goal—e.g., eliminating achievement gaps at the school-level—it is also essential to think about the ways knowledge within schools and districts is developed and shared. These resources may be organized to help facilitate knowledge development about how to accomplish a particular goal, but usage ultimately depends on the individuals and organizations tasked with utilizing that knowledge. This is what Cohen and Moffitt refer to as "the dilemma of implementation" (Cohen & Moffitt, 2009)

Organizational theory has, in recent years, recognized that the ability of firms to continually succeed in a changing industry—to *innovate*—depends on the development of capabilities that move beyond conventional resources, such as financial capital and brick-and-mortar assets. This kind of organizational learning is largely a function of an organization's ability to seek out new information through socially constructed interpretations of interactions between the organization and the environment (Lave, 1993; Wenger, 1998) and then use that information, for example, to determine whether and how it should be incorporated into organizational structures (Weick, 1995). The ability of organizations to be able to learn in this way can depend on conditions for learning at the organizational level. For example, Cohen and Levinthal (1990) argue that "the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends is critical to its innovative capabilities" and that this absorptive capacity is largely a function of a firm's prior knowledge.

As has been demonstrated in prior sections, whole school reform as a way to target within-school inequality is a departure from prior expectations of American educational policy. As such, it is reasonable to anticipate that such work is difficult and that doing so would require development of innovative practices and strategies within schools and districts. It follows that schools and districts with more prior experience with such policies, personnel with a more sophisticated pedagogical arsenal at their disposal, and investments in organizational structures that might support such work would have greater capacity to innovate and meet the challenges of reducing within-school inequality. Conversely, a school without these prior experiences and capabilities would have lower levels of absorptive capacity to innovate in the context of this next iteration of standards-based reform.

This theory of absorptive capacity stems from research on individuals' cognitive structures where pre-existing knowledge aids in future learning. In this construction, "learning is cumulative, and learning performance is greatest when the object of learning is related to what is already known" (Cohen & Levinthal, 1990). Thus, not only is prior expertise in domains that may relate to future learning critical, but diversity of prior knowledge is important in arenas where there is "uncertainty about the knowledge domains from which potentially useful information may emerge" (Cohen & Levinthal, 1990). This cognitive lens of knowledge assimilation and utilization is prevalent in policy implementation research focusing on the role of individual actors situated within institutions (see, e.g., Spillane, 2005; Coburn, 2005).

Absorptive capacity may have different attributes at the individual and organizational levels, however, and it is important to note that an organization's absorptive capacity is not merely a sum of the absorptive capacities of its individual members. Absorptive capacity at the organizational level relies not only on substantive, technical knowledge, but also knowledge of

where to find other kinds of knowledge or capabilities within or outside of the organization. Additionally, the ideal knowledge structure for organizations involves a mix of shared, overlapping knowledge to facilitate communication and disparate, diverse knowledge bases to better equip organizations to deal with new domains or communicate with outside entities (Cohen & Levinthal, 1990).

Policy Evaluation: Mixed Methodologies as a Research Strategy

Since the "birth" of implementation research in the 1970s, policy implementation has been an important element of social reforms and a primary concern of policymakers and agency executives in enacting these reforms. The broad umbrella of policy implementation is often used to describe the processes by which policy is translated into practice; to understand the variability in educational outcomes that are found in the U.S. education system; and to explain why educational programs often yield null average treatment effects on desirable outcomes. In particular, understanding the persistent disconnect between the intended goals of policy design and the results of policy evaluations is imperative. By better understanding the processes and influences at work, we not only can retroactively identify the mechanisms by which policies function but also predict how future iterations of reform may behave.

Concurrently, there has been a growing emphasis in education research on improving rigor in research designs to provide empirical estimates that meet the standards for making causal claims. Methodologies used for causal inference include randomization as well as quasi-experimental designs such as regression discontinuity, instrumental variables, and fixed effects. The establishment of the Institute for Education Sciences (IES) in 2002 has facilitated funding and dissemination of such research through research reports, technical working groups, and an

online database called the "What Works Clearinghouse" (The Education Sciences Reform Act, P.L. 107-279).

One point to emphasize is that implementation studies originally stemmed from a desire to understand the impact of larger, federally-sponsored programs that tend to be more prescriptive and structured than broad, overarching policies that might have unintended effects beyond the stated goals of the initial legislation or local or state programs. In the case of Michigan Focus Schools, the treatment is a loose collection of suggested interventions, at best, many of which are dependent on local capacity to enact them. For example, one provision Focus Schools must adhere to is for central office administrators to enact a "data dialogue" with Focus School leadership to identify two principles of teaching and learning that contribute to achievement gaps within the school. The quality and content of the data dialogues is largely dependent on who is conducting the exercise and how they interpret state guidance for data use as well as the capacity of Focus School leadership to critically analyze their school level data (see Chapter 3 for a greater discussion of this point). In this way, variation in implementation is built into the policy design. It delegates improvement functions down from the Michigan Department of Education, which has been reduced in size from 2,000 individuals to less than 200 in the previous 10 years, to local entities including intermediate school districts, central office administrators, and Focus Schools. The question then becomes: Does this policy lead to the average effects that it seeks? Whatever the answer, there is bound to be some dispersion of effects. Hence further questions arise: What do local practitioners do when faced with a broad set of interventions to reduce within-school inequality, and what can we learn from the varied responses? Studies that help to answer these questions would help to improve our understanding of how and why policy effects are distributed amongst the population of affected institutions.

References

Bardach, E. (1977). *The implementation game: What happens after a bill becomes a law.* Cambridge, MA: The MIT Press.

Berman, P. & McLaughlin, M.W. (1979). *An exploratory study of school district adaptation*. Santa Monica, CA: The RAND Corporation.

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2015). *Too many children left behind: The U.S. achievement gap in comparative perspective.* New York: Russell Sage Foundation.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building Networked Improvement Communities in education. In M. Hallinan (Ed.), *Frontiers in sociology of education* (pp. 127–162). New York, NY: Springer.

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*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Bryk, A.S., Lee, V.E., & Holland, P.B. (1993). *Catholic schools and the common good.* Cambridge, MA: Harvard University Press.

Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013). *A strategy for leveraging research for educational improvement in school districts.* New York, NY: William T. Grant Foundation.

Coburn, C.E. (2005). Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Education Policy, 19*(3), 476–509.

Coburn, C.E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education, 77*(3), 211–244.

Cohen, D.K., Peurach, D.J., Glazer, J.L., Gates, K., & Goldin, S. (2014). *Improvement by Design: The promise of better schools*. Chicago, IL: The University of Chicago Press.

Cohen, D. K., & Moffitt, S. L. (2009). *The ordeal of equality: Did federal regulation fix the schools?* Cambridge, MA: Harvard University Press.

Cohen, D., Raudenbush, S., & Ball, D. (2003). Resources, instruction and research. *Educational Evaluation and Policy Analysis, 25*(2), 1–24.

Cohen, D., Raudenbush, S., & Ball, D. (2002). Resources, instruction and research. In F. Mosteller and R. Boruch (Eds.), *Evidence Matters: Randomized Trials in Education Research* (pp. 80–119). Washington, D.C.: Brookings Institution Press.

Cohen, D.K., & Spillane, J. (1992). Policy and practice: The relations between governance and instruction. In G. Grant (Ed.) *Review of Research in Education 18* (pp. 3–49). Washington, D.C.: American Educational Research Association.

Cohen, W.M., & Levinthal, D.A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*(1), 128–152.

Cohen-Vogel, L., Tichnor-Wagner, A., Allen, D., Harrison, C., Kainz, K., Rose Socol, A., & Xing, Q. (2015). Implementing educational innovations at scale: Remaking researchers into improvement scientists. *Educational Policy, 29*(1), 257–277.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., and York, R. L. (1966). *Equality of educational opportunity*. 2 volumes. Office of Education, U. S. Department of Health, Education, and Welfare. Washington, D.C.: U. S. Government Printing Office. OE-38001; Superintendent of Documents Catalog No. FS 5.238:-38001.

Dee, T.S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management, 30*(3):418–46.

Deming, W. E. (1993). *The new economics for industry, government & education.* Cambridge, MA: Massachusetts Institute of Technology Press.

Datnow, A., Hubbard, L., & Mehan, H. (2002). *Extending educational reform: From one school to many*. London, UK: Routledge Falmer.

Duncan, G., & Murane, R. (2014). *Restoring opportunity: The crisis of inequality and the challenge for American education.* Cambridge, MA: Harvard Education Press and the Russell Sage Foundation.

Duncan, G., & Murane, R. (2011). Introduction: The American dream, then and now. In Greg J. Duncan & Richard J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 3–26). New York: Russell Sage Foundation and Spencer Foundation.

Dynarski, S., Hyman, J., & Schanzenbach, D.W (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management, 32*(4), 692–717.

Education Sciences Reform Act. P.L. 107-279. Retrieved from http://www2.ed.gov/policy/rschstat/leg/PL107-279.pdf

Elmore, R. F. (1983). Complexity and control: What legislators and administrators can do about implementing public policy. In L. Shulman & G. Sykes (Eds.), *Handbook of teaching and policy* (pp. 342–369). New York, NY: Longman

Englebart, D. C. (2003, September). *Improving our ability to improve: A call for investment in a new future.* Paper presented at the IBM Symposium on the Coevolution of Technology-Business Innovations, Boulder, CO.

Figlio, D., & Loeb, S. (2011). School accountability. In E. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education*, *3*. Amsterdam: North Holland/Elsevier.

Gawande, A. (2007). *Better: A surgeon's notes on performance.* New York, NY: Picador.

Figlio, D., & Ladd, H. (2007). School accountability and student achievement. In H. Ladd, & E. Fiske (Eds.), *Handbook of research on education finance and policy*. New York and London: Routledge.

Goals 2000: Educate America Act. P.L. 103-227. Retrieved from http://www2.ed.gov/legislation/GOALS2000/TheAct/index.html

Hanushek, E. A. (1989). The impact of differential expenditures on school performance. *Educational Researcher, 18*(4), 45–65.

Hedges, L.V., Laine, R.D., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes (An exchange: Part I). *Educational Researcher, 23*(3), 5–14.

Honig, M. (Ed.). (2006). *New directions in education policy implementation: Confronting complexity*. Albany, NY: State University of New York Press.

Jacob, B.A., (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago public schools. *J Public Econ. 89*(5–6), 761–796.

Jacob, B., & Levitt, S. (2003). Catching cheating teachers: The results of an unusual experiment in implementing theory. In W. G. Gale & J. R. Pack, (Eds.), *Brookings-Wharton Papers on Urban Affairs*. Washington, D.C.: Brookings Institution Press, 185–209.

Jencks, C., & Phillips, M. (Eds.). (1998). *The black-white test score gap*. Washington, D.C.: Brookings Institution Press.

Jencks, C. D., Smith, M., Acland, H., Bane, M.J., Cohen, D., Gintis, H., Heyns, B., & Michelson, S. (1972). *Inequality*. New York, NY: Basic Books.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., Wilkinson-Flicker, S., Barmer, A., & Dunlop Velez, E. (2015). *The condition of education 2015* (NCES 2015–144). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from http://nces.ed.gov/pubsearch

Kochanek, J. R., Scholz, C., & Garcia, A. N. (2015). Mapping the collaborative research process. *Education Policy Analysis Archives, 23*(121). Retrieved from http://dx.doi.org/10.14507/epaa.v23.2031

Langley, G., Moen, R., Nolan, K., Nolan, T., Norman, C., & Provost, L. (2009). *The improvement guide* (2nd ed.). San Francisco, CA: Jossey-Bass.

Lave, J. (1993). The practice of learning. In S. Chaiklin and J. Lave (Eds.), *Understanding practice: Perspectives on activity and context*. Cambridge: Cambridge University Press.

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher, 44*(1), 54–61.

Lipsky, M. (1980). *Street-Level bureaucrats: Dilemmas of the individual in public services*. New York, NY: Russell Sage Foundation.

Malen, B. (2006). Revisiting policy implementation as a political phenomenon: The case of reconstitution policies. In Honig, M. (Ed.) *New directions in education policy implementation: Confronting complexity*. Albany, NY: State University of New York Press.

McDermott, K. (2004). Incentives, capacity, and implementation: Evidence from Massachusetts education reform. *Journal of Public Administration Research and Theory. 16*(1), 45–65

McLaughlin, M. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, *9*(2), 171–178.

McLaughlin, M., & Talbert, L. (1999). *Professional community and the work of high school teaching.* Chicago, IL: University of Chicago Press.

Murphy, J. (1971) Title I of ESEA: The politics of implementing federal education reform. *Harvard Educational Review, 41*(1), 35–63.

No Child Left Behind Act of 2001. P.L. 107-119. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/pg1.html#sec1001

Odden, A., Ed.(1991). *Education policy implementation.* Albany, NY: State University of New York Press.

Pressman, J.L., & Wildavsky, A. (1973). *Implementation.* Berkeley, CA: University of California Press.

Reardon, S.F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In Greg J. Duncan & Richard J. Murnane (Eds.). *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–116). New York, NY: Russell Sage Foundation and Spencer Foundation.

Rowan, B., & Miskel, C. G. (1999). Institutional theory and the study of educational organizations. In J. Murphy & K. S. Louis (Eds.), *Handbook of research in Educational Administration* (359–382). San Francisco, CA: Jossey-Bass.

Sabatier, P.A., & Mazmanian, D. (1980). The implementation of public policy: A framework of analysis. *Policy Studies Journal 8*(4): 538–60.

Smith, M.S., & O'Day, J. (1990). Systemic school reform. In S. H. Fuhrman & B. Malen (Eds.), *The Politics of Curriculum and Testing (Politics of Education Association Yearbook*, 1990) (pp. 233–267). London, UK: Taylor & Francis.

Sparks, S. D. (2013, October 2). 'Improvement science' seen as emerging tool in K–12 sphere. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2013/10/02/06improvement.h33.html

Spillane, J. P. (2005). Distributed leadership. *The Educational Forum*, *69*(2), 143–150.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*(3), 387–431.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2010). *State and Local Implementation of the* No Child Left Behind Act, *Volume IX—Accountability Under* NCLB*: Final Report*. Washington, D.C.: Rand.

Weick, K.E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: SAGE.

Wenger, E. (1998). *Communities of practice: Learning, meaning and identity.* New York, NY: Cambridge University Press.

Young, T., & Lewis, W.D. (2015) Educational policy implementation revisited. *Educational Policy 29*(1), 3–17.

Chapter 2

The Effects of Focus School Identification on Within-School Inequality: Evidence from Michigan

Abstract

This study uses a regression discontinuity design to examine the causal impact of being identified as a Focus School in 2011-12 on within-school inequality after two years. Under Michigan's approved ESEA waiver, Focus Schools are identified as the 10% of schools with the largest gap in student achievement between the top third and bottom third of students determined by a composite gap measure of student achievement on standardized tests. Results show potentially null effects for the effect of Focus School identification on within-school inequality using a gap composite as well as an overall school percentile index measures calculated by the state. However, these results should be interpreted with caution, as the study is underpowered. Furthermore, the construction of a within-school achievement gap measure may not be reflective of policy goals to address inequality in student achievement compared to alternative measures, such as examining disaggregated data by student subgroup. Given that 43 states are currently identifying and supporting Focus Schools through ESEA flexibility waivers, however, a potential null average effect could hide heterogeneity in treatment effects by state or methodology. These variations are worth examining in future research.

Introduction

Education reform in the United States over the past twenty years has been marked by the introduction of standards-based accountability, which overlays a framework of standards and assessments on a patchwork of local, state, and federal educational policy to drive improvements in student learning (Cohen & Moffitt, 2009). These ideas formed the basis for the 2001 reauthorization of the Elementary and Secondary Education Act (ESEA) of 1965, more commonly known as the No Child Left Behind (NCLB) Act, which held schools accountable for student proficiency by measuring a school's Adequate Yearly Progress (AYP) on statewide standardized assessments and implementing a series of school-based sanctions for poor performance. However, the effects of No Child Left Behind have been mixed. Some studies indicate that No Child Left Behind increased student achievement for fourth- and eighth-graders, particularly in mathematics and especially for Hispanic students compared to their White peers (Hanushek & Raymond, 2005; Jacob, 2005; Wong, Cook, & Steiner, 2011). Other studies show adverse effects of the policy, including exacerbated teacher turnover, higher incidents of dishonest behavior (e.g., cheating on standardized tests), and increased reclassification of students from low-income backgrounds as special needs to circumvent testing and reporting requirements (Neal & Schanzenbach, 2010; Figlio & Getzler, 2006; Clotfelter, Ladd, Vigdor, & Diaz, 2004; Jacob & Levitt, 2003). Furthermore, research indicates there has been wide variability between states in the efficacy of NCLB to close racial achievement gaps, particularly gaps in student achievement between Black students and their White peers (Dee & Jacob, 2011; Reardon, Greenberg, Kalogrides, Shores, & Valentino, 2013; Hanushek & Raymond, 2005).

To provide more flexibility to states in achieving the policy aims set forth in the legislation (i.e., all students proficient by 2020), the U.S. Department of Education allowed states

to apply for ESEA waivers beginning in 2011 that would release states from complying with NCLB provisions in exchange for compliance with interim progress measures to achieve the goals set forth in the NCLB law. Part of these waivers mandated that states must identify three categories of schools: Reward Schools, Priority Schools, and Focus Schools. Reward Schools are schools that are classified as the highest performing 5% of schools in the state; Priority Schools are the lowest performing 5% of schools in the state; and Focus Schools are defined as the 10% of schools with the largest achievement gaps. This last category of schools identified under ESEA waivers is arguably different from prior identification of schools under U.S. accountability policy as it specifically identifies *high-inequality schools* in lieu of *low-performing schools*. Consequently, it is of interest whether identifying schools with higher levels of within-school inequality leads to an increase in average achievement and/or a decrease in inequality.

Ongoing research suggests that there may be null or very small effects of Focus and Priority identification on student achievement or measures of within-school inequality.[1] However, more information is needed as to whether those null effects found in Louisiana, North Carolina, Kentucky, and Michigan (using student-level data) are reproduced in states across the nation, and if so, what the reason may be for the lack of effect of such a prevalent national policy. This study uses a regression discontinuity design to determine the causal impact of Focus identification in Michigan for the first cohort of schools identified in 2011-12. The rest of this chapter is organized as follows: Section II describes the policy context for Focus Schools in Michigan; Section III describes the data, sample, and measures used in the study; Section IV describes the identification and analytic strategies employed; Section V describes results for

[1] There are a series of papers coming out looking at the effects of ESEA waivers on student achievement. These papers were presented at the American Educational Finance and Policy annual conference on March 17, 2016 in Denver, Colorado, but have not yet been published. Authors include Hemelt & Jacob, Dee & Dizon-Doss; Doss & Dee; and Bonilla & Dee.

parametric and non-parametric estimation of treatment effects; Section VI discusses robustness and specification checks; and Section VII discusses limitations and implications for policy.

Focus Schools in Michigan

Identifying Focus Schools

In Michigan, per ESEA waiver requirements and guidance, Focus Schools are identified as the 10% of schools that have the largest gap between the highest performing 30% and lowest performing 30% of students, as measured by a composite of student achievement test scores by subject and grade. There are several different ways in which states could propose identifying the 10% of schools statewide with the highest achievement gap. For example, Tennessee's formula identifies gaps between subgroups of historically struggling students (e.g., racial/ethnic minorities, special education students, and economically disadvantaged students) and identifies schools that have the highest subgroup gaps and schools where an identified subgroup performs lower than a 10% proficiency rate for its students.

Michigan's construction of high achievement-gap schools as schools with large gaps between the top and bottom third of students was a departure from previous identification of achievement gaps at the student subgroup level. For example, past research had focused on the achievement gap between Black and White students (Jencks & Phillips, 1998; Fryer & Levitt, 2004; Reardon & Galindo, 2009) and, more recently, on the widening gap between economically disadvantaged students and their more affluent peers (Reardon, 2011; Duncan & Murnane, 2014). These data show that the Black-White achievement gap is approximately 0.7 standard deviations while the income achievement gap between students in the 90th percentile of family income and students in the 10th percentile of family income has been steadily growing and now stands at 1.25 standard deviations in reading and 1.4 standard deviations in mathematics

(Reardon, 2011). In these cases, a gap of 1 standard deviation roughly corresponds to three to six years of learning in middle or high school. In comparison to this tradition of examining standardized test score differences between student subgroups, the ESEA waivers require identification at the school-level. The distribution of a school-level achievement gap measure across schools has not been widely studied. One can assume, though, that the variance of this measure and its interpretation differs from that at the student-level. It is unclear what the variation in between-school variation is with respect to a measure of within-school inequality as it pertains to Focus Schools. Consequently, this study can be thought of as an exploratory study to examine the properties of a gap composite measure at the school-level. Qualitatively, however, such a measure may not address the specific goals of reducing unequal outcomes among students that the ESEA seeks to accomplish. Another possible outcome is that higher performing students become, on average, lower performing as a result of shifting resources to lower-performing students. To assess whether there is a change in overall school performance, I use a measure of the school performance index (SPI), to determine whether this substitution effect is real.

It is important to note that the initial accountability system for identifying and supporting Focus Schools developed by MDE changed over time to reflect evolving priorities, understanding, and capacity. For example, the first cohort of Focus Schools identified in 2012 using 2011-12 data were defined as the 10% of schools with the largest gap between the top 30% and bottom 30% of students. However, some schools protested identification under this system because even their bottom 30% of students performed relatively well compared to the state average. As a consequence, the cohort of Focus Schools identified two years later in 2013-14 excluded schools whose average school achievement was at or above the state average for

student performance. In another example, schools protested being identified annually because their bottom 30% of students was improving at a faster rate than the state average. As a result of these concerns, the 2013-14 definition of Focus Schools excluded identifying those schools whose students in the bottom 30% were improving at or above the state average rate of improvement.[2] These revisions in the Focus definition alleviated concerns that higher performing Focus Schools were diverting resources from schools and students with greater need. It also rewarded schools for improvement by excluding them from the accountability system.

Supporting Focus Schools

Schools that are designated as Focus Schools remain in a cohort of Focus Schools for four years, regardless of whether they are identified in subsequent years as falling in the 10% of schools with the largest within-school achievement gap. The first Focus School cohort was identified using Michigan's Top-to-Bottom Ranking methodology in the 2011-12 school year (N=358).[3] Focus School cohorts have been identified in each subsequent year, with the third and latest cohort identified for the 2013-14 school year.

The treatment, or intervention, for being identified as a Focus School in Michigan consists of the following bundle of services outlined in Michigan's federally approved waiver:

- Parents must be notified of a school's Focus status (Title-I schools only)

[2] Chapter 2 uses only the definition used to identify the 2012 cohort using 2011-12 data. Because definitions were not changed until 2013-14, I only examine effects of initial Focus identification in the first two years of implementation.

[3] The Top-to-Bottom list is part of Michigan's school accountability system which ranks schools on their student performance in mathematics, reading, writing, science, and social studies and (for high schools) graduation rate data. School performance components include student achievement, improvement, and achievement gaps between the highest and lowest scoring 30% of students in each school. Schools with at least 30 full academic year students with more than one year of data in at least two state-tested content areas are included in the rankings. Additional information on the Top-to-Bottom ranking methodology can be found here: http://www.michigan.gov/mde/0,4615,7-140-22709_56562---,00.html

- Schools must participate in a "data dialogue" facilitated by the district that identifies at least two principles of teaching and learning that can be implemented to reduce the achievement gap within the school. These principles of teaching and learning must be written into the School Improvement Plan (SIP).
- Schools must participate in the Superintendent's Dropout challenge, which identifies student at-risk for dropping out and provides them with interventions to attenuate those risks.
- Districts with Focus Schools must coordinate supports for Focus Schools by revising its District Improvement Plan (DIP), monitoring progress on SIP goals, and providing professional development about systemwide change.

In addition, Title I schools, which make up 51% of the 2011-12 Focus cohort, are assisted by a state-funded District Improvement Facilitator (DIF) and must set aside at least 10% of Title I funds for one of six options to attenuate the within-school achievement gap. These six options include professional development on Multi-Tiered Systems of Support (MTSS); weekly or daily time for teacher collaboration; administration of the Survey of Enacted Curriculum; needs assessment through the local ISD or AdvancEd Michigan; professional learning for students with disabilities assessment, MI-ACCESS; and culture or climate interventions. This bundle of interventions is required for all Focus Schools over four years, regardless of identification as a Focus School in subsequent years. It is important to note that the criteria for exiting Focus Status has been amended over the course of the past four years to eliminate schools that achieve above the state average or whose bottom 30% of students improve at or above the statewide rate of improvement. In addition, while the requirements for Focus Schools have not changed, the guidance provided on how data dialogues should be implemented and the technical assistance

that is provided to Focus Schools has changed over time. Consequently, this paper focuses on the effect of Focus identification in 2012 on 2013-14 outcomes, during which time the Focus exit criteria and intervention remained constant.

Data, Sample, and Measures.

Data for this study comes from administrative, publically available school-level data from the MISchool portal for SY 2011-12 and SY 2013-14, which includes all Michigan public schools in the state's Top-to-Bottom Ranking (N = 2866) for which there is a gap composite measure calculated in 2013-14. The total sample (N= 2648) includes 358 Focus Schools identified in 2011-12 but excludes schools identified as both Priority and Focus Schools and schools identified as Focus Schools as a result of a consistently low graduation rate. Data used from the MISchool database include (1) dummy indicators of Focus and Priority status, (2) compiled indices from Michigan's Top-to-Bottom School Ranking, and (3) student demographic data of the school population. State-level indices include school-level gap composite scores, school percentile index, and school percentile rank. Descriptions of how each of these measures are calculated and used in the analysis are described below. Student demographic data includes total school enrollment, race/ethnicity, economically disadvantaged status, special education status, and English learner status.

Focus and Priority Status (Treatment Indicator)

Publically available data released by the Center for Educational Performance and Information (CEPI) provides a dummy-indicator if a school was identified as a Focus or Priority School based on the state's Top-to-Bottom ranking using the gap composite score and school percentile index, respectively (see below for a definition of each measure). Focus Schools are identified as the 10% of schools with the highest gap composite score and/or a graduation rate

below 60% for more than three years. Priority Schools are identified by the state as the 5% of schools with the lowest school percentile index. For schools identified as both Focus and Priority (i.e., highly unequal and low-performing), Priority status took precedence; these schools are dropped from the sample (N = 41). Schools that were identified as Focus Schools due to their low graduation rate are also dropped to maintain a sharp discontinuity at the 10% cutoff for Focus identification (N = 2). The total sample (N= 2675) includes 358 Focus Schools identified in 2011-12 for which there is a gap composite measure calculated in 2013-14. Of these schools, 40% remained on the list in the 2013-14 school year (N=142). The gap composite score for schools identified again in 2014 was -1.33, while the average gap composite score for Focus Schools not identified again in 2014 was -0.24, indicating a more equal distribution of within-school achievement among those schools that had gotten off the Focus School list.

Gap Composite Score (Outcome)

The primary measure of within-school inequality is a gap composite score, which is a composite of differences in student achievement between the top 30% and bottom 30% of students in each subject tested in a school in a given year. In varying years, this measure serves as both the rating variable (2011-12) and an outcome measure (2013-14). The gap composite score is calculated by averaging student z-scores in the top and bottom 30% in a particular subject in a given school over a two-year period. The difference of this running average is calculated for each subject within a school and then standardized compared to the statewide distribution, giving a subject-school gap. Each subject-school gap calculated for a school is then averaged to produce an average school-level gap across all subjects. Schools are then rank-ordered by their gap composite score to identify the 10% of schools with the largest within-school gaps in student achievement.

School Percentile Index (Outcome)

A secondary measure of student performance within the school used is the 2013-14 school percentile index (SPI). The SPI is calculated by averaging student performance across all content areas (including graduation rate where applicable) in which the school received a school performance index z- score. For schools without a graduation rate index, SPI is calculated as the straight average of all z-scores calculated for the school. For schools with a graduation rate index, the school performance index on graduation rate must account for exactly 10 percent of the overall school performance index. This is accomplished by multiplying the straight average of all other z-scores calculated for the school by the value 0.9 and adding the graduation rate z-score multiplied by 0.1.

Total school enrollment. Total school enrollment is defined in the MISchool data as the total number of students identified on a count day in the fall and spring of a given school year.

Female enrollment. School-level proportions of female students are calculated by dividing total female enrollment by total student enrollment and total female enrollment in a given year.

Race/ Ethnicity. School-level proportions of White, Black, American Indian, Hispanic, and Asian students are calculated using total student enrollment by race/ethnicity divided by total student enrollment for a school in a given year.

Economically Disadvantaged status. A student's economically disadvantaged status is flagged for the entirety of a school year if the student is directly certified as such, eligible for supplemental nutrition such as free-or-reduced price lunch (FRL), or reported as homeless or migrant. The proportion of students who are classified as being economically disadvantaged is calculated by dividing by the total school enrollment.

Special Education status. The proportion of special education students within a school is calculated by dividing the number of students reported to have Individual Educational Plans (IEPs) with the total school enrollment.

English Learner status. The proportion of English Learners in a school is calculated by dividing the count of students participating in EL programs divided by the total school enrollment.

Table 1 provides descriptive information for school characteristics by Focus and non-Focus School. Focus Schools do not differ from non-Focus Schools by student demographics, though Focus Schools are, on average, higher performing and more unequal per the gap composite score used for identification.

Empirical Strategy

This study employs a sharp regression discontinuity design to ascertain the causal impact of being identified as a Focus school in 2011-12 on within-school inequality, as measured by a school's gap composite score. I calculate the intent-to-treat (ITT) estimate, which I demonstrate is the same as the treatment-on-the-treated (TOT) due to perfect compliance with being identified by the metric and being identified by the accountability system. However, the consequences of Focus identification—that is, the strength with which the bundle of interventions was implemented—may vary from school to school. This varying dosage does not track in the data but is documented in a nested qualitative case study examining implementation of supports in Chapter 3.

Identification.

The basic function for identifying treatment effects of being identified as a Focus School in Michigan is:

$$Y_{s,t+2} = BX_s + \beta_1(Focus)_{s,t} + \mu_s + \varepsilon_{s,t+2} \quad [1]$$

Where Y is a school's 2013-14 gap composite measure or school performance index, X is a vector of school characteristics; Focus is a binary indicator, which takes on a value of 1 if the school was identified as a Focus School in 2011-12 and 0 otherwise; μ are unobserved factors that are correlated with the outcome of interest; ε is an error term; and *s* and *t* are subscripts for school and time, respectively. This identification strategy takes advantage of the discontinuity that results from Focus status being dependent only on the gap composite measure, which creates a highly non-linear relationship between the gap composite measure and Focus status. Figure 1 depicts the probability of treatment based on the assignment variable, the gap composite measure from 2011-12; it shows that this policy has been implemented with fidelity in Michigan. Over 95% of schools identified as Focus Schools received Focus status. In approximately 5% of cases, schools were identified as both Priority and Focus Schools. In these schools, Priority status took precedence. These schools are not included in the analysis; therefore, Figure 1 demonstrates a sharp cutoff according to the selection rules for Focus identification.

There are at least two ways to conceptualize a regression discontinuity design, both of which rely on the exogeneity of treatment status at the cutoff. The first is to focus on the discontinuity itself and characterize the treatment effect as the relationship between treatment assignment (in this case a binary indicator of Focus status) and outcome variables (Hahn, Todd, & van der Klaauw, 2001). The second is to assume random assignment of schools to treatment and control conditions at the cutoff, also known as "local randomization" (Lee, 2008).

The first conception relies heavily on 1) accurately modeling the relationship between the assignment and outcome variables by ensuring correct model specificity and 2) dealing with treatment misallocation at the cutoff. Figure 2 shows what would happen if a linear specification

were overlaid on what the graphical evidence indicates should be modeled using a cubic function. This misspecification results in a significant positive effect using a linear model when in fact the cubic function indicates a null effect. The data analysis guards against these types of misspecifications by checking fit statistics and also comparing estimated results against non-parametric estimation techniques. With respect to the second consideration of treatment misallocation, in this study, treatment misallocation would result if a school manipulated student testing in a way that would reduce within-school inequality for each subject and also relative to other schools in the state. Given the complex nature of the manner in which the gap composite measure is calculated, this time of treatment misallocation is highly implausible. Another way in which treatment misallocation might occur is if schools identified within the 10% of schools with high levels of inequality were able to get off the Focus list; however, due to the public nature of the state's Top-to-Bottom Rankings and the public availability of data for replication, this is also highly implausible; furthermore, there is no evidence of such actions for 2011-12 Focus identification.

The local randomization approach can be thought of as a difference in means on either side of the cutoff. This second conceptualization of a regression discontinuity relies on the estimation of a local average treatment effect (LATE), which is the treatment effect at the cutoff but not generalizable beyond a narrow bandwidth of cases. While this approach reduces reliance on the functional form of the model, it necessitates low treatment misallocation and also a density of cases around the cutoff for estimation. In this approach, non-parametric estimation must pay careful attention to bandwidth and bin width selection—that is, the range of values included in the analysis for the running variable and the range of values from which the average of the running variable will be taken, respectively.

Estimation

This study uses both parametric and non-parametric techniques to estimate the treatment effects of Focus identification at the cutoff. The primary parametric regression used to estimate treatment effect is:

$$Y_{s,t+2} = a + BX_{st} + \beta_1(Focus)_{st} + g(Gap\ Composite)_{st} + \varepsilon_{s,t+2} \quad [2]$$

where $Y_{s,t+2}$ is a school's gap composite measure or SPI in 2013-14; a is a constant; X_{st} is a vector of school characteristics including total student enrollment, proportion of female students, proportion of students by race/ethnicity, proportion of economically disadvantaged students, proportion of special education students, and proportion of English learners; $(Focus)_{st}$ is a binary indicator of each school's Focus status in 2011-12 with 1 indicating inclusion in the treatment condition; β_1 is the causal effect of treatment on the outcome of interest; $g(Gap\ Composite)_{st}$ is a function of the assignment variable (gap composite score in 2011-12); and $\varepsilon_{s,t+2}$ is the error term. To check the robustness of effects to alternative model specifications, polynomial and interaction terms are included in the estimation of $g(Gap\ Composite)_{st}$. Results from the inclusion of linear, quadratic, and cubic terms are presented as are estimates from a truncated sample (±1), which allows cases closer to the cutoff to contribute greater weight to the regression estimate and reduces the possibility that outlying cases sway functional form.[4] To determine best model fit, a joint F-test and measures of fit (such as adjusted R-squared values) are used to determine optimal model specification. The optimal model specification for this study for each outcome is a cubic function.

[4] Estimates were also calculated for truncated samples of ±2 and ±3, yielding similar results. Estimates calculated for a truncated sample of ±0.5 had too few cases to estimate treatment effects per the empirically determined model specification.

Non-parametric analysis techniques are used as a robustness check on the functional form identified in the parametric analyses. In this approach, a bandwidth is selected within which a weighted regression is estimated along with the inclusion of a linear term, referred to as "local linear regression" (Hahn, Todd, & van der Klaauw, 2001). The challenge in this approach is to select the correct bandwidth. Narrower bandwidths yield more unbiased but less precise estimates due to the smaller number of cases that contribute to the treatment effect. Wider bandwidths are more precise but may yield biased results. A range of bandwidths is presented in the results that demonstrate this tradeoff between bias and efficiency, with a preferred focal bandwidth of 0.6, calculated following the "plug-in" procedure adapted for RD designs by Imbens and Kalyanaraman (2009) and DesJardins and McCall (2014). Non-parametric estimates present smoothed mean-differences of varying bandwidths around the cutoff with bootstrapped standard errors.

Results

Overall results show the effect of Focus identification has a potentially positive effect but likely null effect on within-school achievement, though this effect is inconsistent across parametric and non-parametric approaches, and a positive but statistically insignificant effect on SPI. Figure 2 shows a local linear regression graphical depiction of the discontinuity. We see that there appears to be a small negative jump at the discontinuity for Focus Schools, which are identified to the left of the cutoff. However, visual inspection of the discontinuity may lead to erroneous conclusions, so empirical results are presented to determine the true treatment effect. Table 2 shows parametric effects using linear, quadratic, cubic, and fourth-order polynomial and interaction terms. The preferred specification is a fourth-order polynomial, which is corroborated by values from joint F-tests following Lee and Lemieux (2010) and goodness-of-fit statistics

such as the AIC (Jacob, Zhu, Somers, & Bloom, 2012), which suggest that a fourth-order polynomial function is the best model specification for the data. Effects are shown to be primarily positive and statistically significant across model specifications. However, non-parametric estimates in Table 3 show null results at the optimal bandwidth of 0.6, half of that bandwidth at 0.3, and twice that bandwidth at 1.2. These results indicate null effects are more likely than the positive effects found in the parametric estimation approach (see, e.g., Gelman and Imbens, 2016).[5] Figure 7 and Table 5 show similar null results of Focus Status on overall SPI.

There are several threats to internal validity that must be considered in light of these null results. A first concern is that schools are able to manipulate treatment status despite being identified by the Top-to-Bottom Ranking methodology outlined by the state's department of education. There is no evidence that the schools identified through the Top-to-Bottom Ranking were able to negotiate Focus status after identification. The Michigan Department of Education has made the data publicly available to recreate rankings, and the methodology is transparently laid out to allow for replication. Figure 1 shows perfect compliance with respect to the probability of receiving treatment if identified as a Focus School, lending further credence to the idea that Focus status complied with rules set forth in Michigan's ESEA waiver. Furthermore, manipulation of student test scores prior to identification is unlikely in this case as 2011-12 data was used for identification and schools did not know at the time of test taking that scores would be used to identify Focus status. Still, if schools were able to manipulate treatment status, there would be visible sorting at the cutpoint in the data—that is, we may see a cluster of schools

[5] Gelman and Imbens (2016) suggest that controlling for higher-order polynomials in a regression discontinuity design leads to imprecises, volatile estimates. Consequently, I rely more heavily on the non-parametric results. All results should be interpreted with caution, however, given low power to detect effects.

distributed just above or below the cutoff. Results from a McCrary density test show that this sorting does not occur, providing further evidence that schools were not able to manipulate treatment status (Figure 5).

A second threat to validity would be if schools above and below the cutoff were different on characteristics that are related to the outcome of interest, such as gender, race/ethnicity, school size, free and reduced lunch status, or English learner or special education status. However, examining pre-treatment covariate characteristics at the cutoff demonstrates no significant differences on these measures (Figure 6). Unobserved characteristics that differ between identified and unidentified schools at the cutoff but are not discernible in the data may still prove to be confounding factors, such as school leadership or teacher experience. However, it is unlikely that the mechanism is biased given that the first Focus School cohort was identified in 2012 using 2011-12 data. At the time of data collection, schools and states were unaware that they would be identified for Focus Status and the complex nature of the identification mechanism had not yet been conceptualized.

A third concern is that there are other reforms occurring at the same time that either enhance or nullify potential effects of Focus identification on the school-level gap composite measure or overall student achievement in a school. This threat to internal validity could be examined using an outcome measure that is unrelated to the treatment to see whether the alternative, unrelated outcome also shows a null effect. However, identifying such an outcome in the data is problematic because of the whole-school nature of this particular accountability reform. Except for Focus identification, any reform would affect both Focus and non-Focus Schools. It is possible that any such reform may affect these types of schools differentially. If so, this remains a potential threat to internal validity.

A fourth threat to internal validity is that student composition has changed within Focus and non-Focus Schools between identification in 2012 and measuring outcomes of interest in 2013-14. For example, if high-performing students in identified Focus Schools left Focus Schools for non-Focus Schools, leaving behind lower-performing students, this may lead to a lower achievement gap in Focus Schools, but an increased achievement gap in non-Focus Schools. The reliance on school-level data in this study does not account for changes in student composition that might affect within-school inequality, particularly in a state where Focus Schools must notify parents of Focus status and inter- and intra-district choice exists. However, qualitative evidence suggests that barriers to moving schools are high, which may account for the fact that only 6.6% of Michigan students took advantage of Michigan's inter-district schools of choice program in the 2011-12 school year, though this number has been increasing steadily since 2003 (Cowen, Creed & Keesler, 2015). Principals have reported in the past that Focus status is not seen as consequential enough for parents to move students whereas Priority status may be (see Chapter 3 for a more in-depth discussion of this point). An examination of post-treatment covariate balance of demographic characteristics that may be related to the outcome of interest, such as gender, race/ethnicity, school size, free and reduced lunch status, English learner or special education status, also does not indicate significant effects or jumps at the discontinuity, which supports the idea that student composition was unchanged as a consequence of Focus status. Finally, Jacob and Hemelt (2016) use student-level data to produce similar estimates using student-level data in Michigan and do find evidence of significant changes in student composition.

In addition to issues arising from misspecified functional form, I run regression models using truncated samples to test the robustness of treatment estimates to the inclusion of schools

closer to the cutoff as well as farther away from the cutoff (Table 4). Perhaps not surprisingly, the magnitude of the treatment effect is stronger near the cutoff. However, the treatment estimates still show no statistical effect of Focus status on the within-school achievement gap, as the standard errors also increase closer to the cutoff. These estimates may suffer from being underpowered as well.

Finally, though we do not believe there was a true reduction in the within-school inequality index, we may worry the average level of achievement may have somehow suffered as resources shifted to help struggling students. Figure 7 shows the effects of Focus status on the school percentile index at the cutoff of the assignment variable. Table 5 shows that the point estimate for Focus status is positive (0.249) but not statistically significant at the 0.05 level, indicating that any resource diversion that occurred to struggling students was not detrimental to higher performing students on average and did not affect average performance at the school level.

Discussion

The results presented in this chapter suggest that identification of high-inequality schools under a statewide accountability framework is insufficient for producing significant effects on within-school inequality or overall school performance index. These results are consistent with similar studies examining the effect of ESEA waiver identification on average school achievement and within-school inequality (Hemelt & Jacob, forthcoming; Dee & Dizon-Doss forthcoming; Doss & Dee, forthcoming; Bonilla & Dee, forthcoming). However, they must be interpreted with the important caveat that the study is underpowered, and therefore treatment effects may exist that are undetectable using school-level data.

The theory of action of this policy seeks to empower local districts to support high achievement-gap schools in identifying students who are struggling and develop interventions that are appropriately targeted to that group of students. The flexibility embedded in these supports empowers localities to determine what is most appropriate for the struggling students who have been identified and then monitor implementation through the SIPs and DIPs. It also codifies collaboration between the school and district for data review to identify principles of teaching and learning. Finally, the intervention provides monetary resources for implementation for Title I schools and provides a list of possible options for spending those resources.

The findings of this chapter, however, indicate that this set of interventions was not powerful enough to significantly affect the achievement gap on average in Michigan Focus Schools. Analysis of previous education reforms may provide insight as to why this set of bundled services may be perceived as a weak intervention for the bottom 30% of students. First, the policy design puts the onus for improvement on the very localities that likely contributed to creating the achievement gap in the first place, whether intentional or not. The expectation that these same schools and districts can then work together to ameliorate this gap is just that—a presumption—and one that has not yielded fruit in past reform efforts (Cohen & Moffitt, 2009). Second, the lack of prescription of what to do with the monetary resources or what types of teaching and learning priorities to identify suggest that schools and districts must have the capacity to conduct such a data review, articulate pedagogical practices to implement, and support teachers in that implementation. Third, there must be organizational infrastructure in place to support these kinds of reform efforts. For example, if schools and districts are to collaborate in conducting a data dialogue, there must be sufficient time for central office administrators and school-level personnel to work together to conduct the data review and

identify principles of teaching and learning. Similarly, there should be time for teachers to collaborate in the implementation of those teaching and learning principles. Fourth, even if these capabilities existed, both at the individual as well as the organizational levels, there must be buy-in for the endeavor from both individuals and organizations and belief that the ways Focus Schools are identified are valid and reliable.

Another consideration in interpreting results is that I examine treatment effects after a two-year period. Given the diffuse nature of the bundle of interventions (see Chapter 3 for a greater discussion of this point), two years may not be sufficient to detect effects of the policy on student achievement outcomes. There is evidence that whole-school reforms take at least 3-8 years to be fully implemented and that significant learning about optimal options for implementation happens in the first few years of a policy (Fullan, 2001; Honig, 2006).

However, if the treatment effects that have been found in this study and similar studies nationwide are true, this research has several implications for policy design and implementation to reduce inequality. The first is to determine whether the policy goal itself is desirable; that is, does the policy penalize heterogeneous schools and privilege homogeneity with a student body? Second, the bundled set of interventions proposed for Focus Schools is a relatively weak intervention and relies heavily on educator capacity for implementation. As delineated in Chapter 2, an average null effect likely masks heterogeneity that is a result of implementation and other contextual factors. If the policy goal is desirable, perhaps focusing on implementation and capacity building is needed to strengthen the proposed interventions. Third, Focus Schools in Michigan are identified agnostic to grade-level and school characteristics (such as school size) that may influence Focus identification, which has implications for treatment effects.

Given that 43 states are currently under the requirements of ESEA waivers, which includes identifying Focus Schools, it is worth understanding how and if policy goals are being achieved through Focus identification. Even though Focus status will no longer exist under the newly authorized Every Student Succeeds Act (ESSA) of 2015, the latest reauthorization of the ESEA, within-school achievement gaps remain a focus of federal and state policy. Consequently, it is important to understand the causal effect of these policies on outcomes of interest, and in the case of null effects, determine what other types of policies or practices might achieve the policy aims set forth in the legislation.

References

Bonilla, S., & Dee, T.S. (2016). *The impact of NCLB waiver focus school reforms on achievement gaps: Evidence from Kentucky*. Working paper.

Clotfelter, C.T., Ladd, H.F., Vigdor, J.L. & Diaz, R.A. (2004). Do five school accountability systems make it more difficult for low performing schools to attract and retain high quality teachers? *Journal of Policy Analysis and Management*, *23,* 251–271.

Cohen, D.K., & Moffitt, S.L. (2009). *The ordeal of equality: Can federal regulation fix the schools?* Cambridge, MA: Harvard University Press.

Cowen, J., Creed, B., & Keesler, V. (2015). Dynamic participation in inter-district open enrollment: Evidence from Michigan 2005-2013. Working Paper #49, The Education Policy Center at Michigan State University, East Lansing, Michigan.

Dee, T.S., & Dizon-Doss, E. (2015, November*). School performance accountability and waiver reforms: Evidence from Louisiana*. Paper presented at the Association for Public Policy and Management Conference, Miami, FL.

Dee, T.S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management, 30*(3), 418–446,

DesJardins, S., & McCall, B. (2014). The impact of the Gates Millennium Scholars Program on college and post-college related choices of high ability, low-income minority students. *Economics of Education Review, 38*, 124–138,

Doss, C., & Dee, T.S. (2016). *Achievement gaps and triage: Evidence from NCLB waivers in North Carolina.* Working paper.

Figlio, D.N. & Getzler, L.S. (2006). Accountability, ability, and disability: Gaming the system? In Timothy J. Gronberg & Dennis W. Jansen (Eds.), *Advances in Applied Microeconomics*, Volume 14, pp. 35–49. Oxford, UK: Elsevier Science Press.

Fryer, Roland G., Jr. & Steven D. Levitt (2004). "Understanding the Black-White Test Score Gap in the First Two Years of School," *Review of Economics and Statistics 86*(2), 447–464.

Fullan, M. (2001). *The new meaning of educational change,* (3rd ed.). New York, NY: Teachers College Press.

Hahn, J., Todd, P., & van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression discontinuity design. *Econometrica, 69*(1), 201–209.

Hanushek, E.A., & Raymond, ME. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297–327.

Honig, M.I. (Ed.). (2006). *New directions in education policy implementation: Confronting complexity.* Albany, NY: The State University of New York Press.

Imbens, G. W., & Kalyanaraman, K. (2011). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, *79*, 933–959.

Jacob, B.A., (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics 89*(5–6), 761–96.

Jacob, B.A., & Hemelt, S. (2016). Differentiated accountability and education production: Evidence from NCLB waivers. Working Paper.

Jacob, B.A., & Levitt, S. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics, 118*(3), 843–877.

Jacob, R. T., Zhu, P., Somers, M.A., & Bloom, H.S. (2012). *A practical guide to regression discontinuity*. MDRC Working Paper. New York, NY: MDRC.

Jencks, C., & Phillips, M. (Eds). (1998). *The black-white test score gap*. Washington, D.C.: Brookings Institution Press.

Lee, D.S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics, 142,* 675–97.

Lee, D., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48* (2), 281–355.

Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics, 92*(2), 263-283.

Reardon, S.F. (2011). The widening socioeconomic status achievement gap: new evidence and possible explanations. In R. J. Murnane & G. J. Duncan (Eds.), *Whither opportunity? rising inequality, schools, and children's life chances*. New York, NY: Russell Sage Foundation.

Reardon, S.F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal 46*(3), 853–91

Reardon, S.F., Greenberg, E.H., Kalogrides, D., Shores, K.A., & Valentino, R.A. (2013). *Left behind? The effect of No Child Left Behind on academic achievement gaps*. Retrieved from http://cepa.stanford.edu/content/left-behind-effect-no-child-left-behind-academic-achievement-gaps

Wong, M., Cook, T. D., & Steiner, P. M. (2011). No Child Left Behind: An interim evaluation of its effects on learning using two interrupted time series each with its own non-equivalent comparison series (WP-09-11). Institute for Policy Research Working Paper Series. Evanston, IL: Northwestern University.

Figure 1. Probability of Focus status in 2011-12

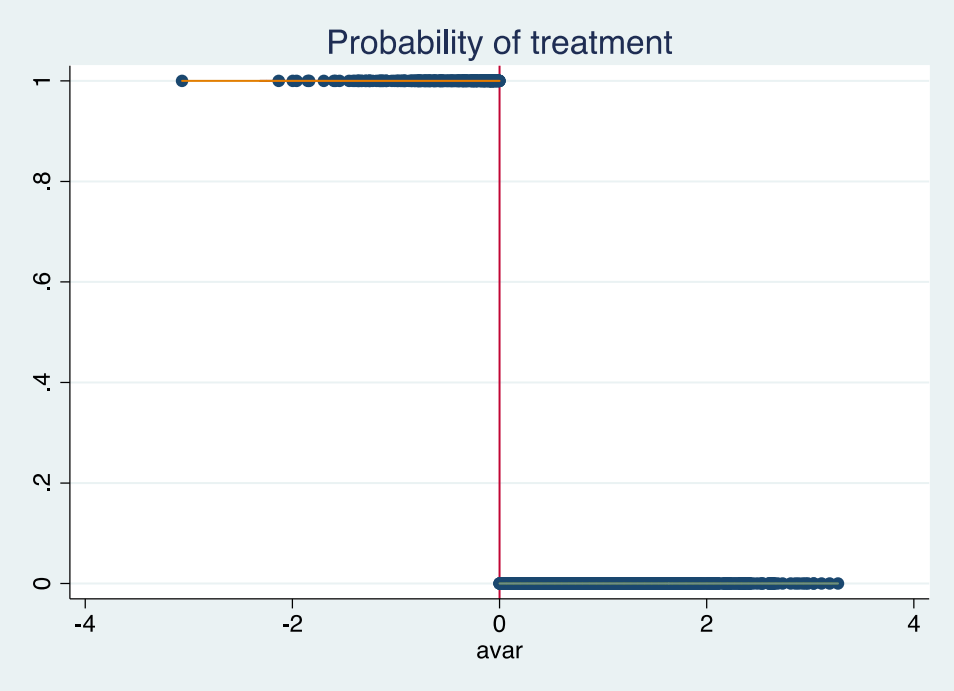


Figure 2. Misspecified linear model overlaid on local linear averages (width=0.1); coefficient = 0.273; p = 0.05.

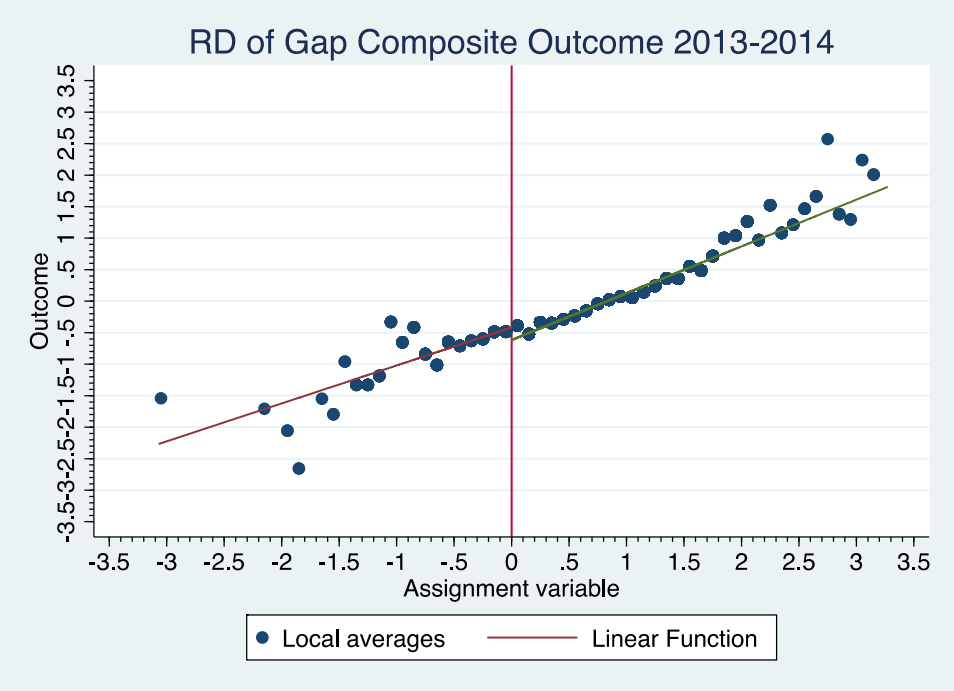


Table 1. Differences in Means by Focus School Status in 2011

	Total (N=2,581)	Non-Focus School in 2011 (N=2241)	Focus School in 2011 (N=340)	Diff.
2011 Gap Composite Score	-0.08	0.11	-1.21	1.33***
	-0.69	-0.53	-0.47	
2013 Gap Composite Score	-0.09	0.01	-0.69	0.73***
	-0.76	-0.71	-0.78	
2011 Overall School Index Score	0.12	0.10	0.21	-0.041
	-0.73	-0.71	-0.83	
2013 Overall School Index Score	0.07	0.03	0.33	-0.25***
	-0.82	-0.79	-0.92	
2011 School Percentile Rank	52.85	52.40	55.66	-0.81
	-26.66	-26.28	-28.84	
2013 School Percentile Rank	51.67	50.43	59.51	-7.88***
	-27.74	-27.29	-29.31	
Total Enrollment	438.00	439.90	426.00	-3.34
	-333.70	-340.50	-287.90	
% Female	0.48	0.48	0.48	-0.0015
	-0.08	-0.08	-0.06	
% Native American	0.01	0.01	0.01	-0.000417
	-0.03	-0.03	-0.04	
% Asian American	0.04	0.04	0.04	0.0276
	-0.13	-0.13	-0.09	
% Black	0.19	0.19	0.20	-0.0119
	-0.29	-0.29	-0.30	
% Latino	0.07	0.07	0.07	0.00303
	-0.11	-0.11	-0.11	
% Economically Disadvantaged	0.54	0.54	0.54	-0.00716
	-0.24	-0.24	-0.23	
% Special Education	0.13	0.13	0.15	-0.0162
	-0.15	-0.15	-0.17	
% English Learners	0.04	0.04	0.04	-0.00293
	-0.11	-0.10	-0.12	

* p<0.05, ** p<0.01, *** p<0.001

Figure 3. Graphical analysis of regression discontinuity of School Gap Composite Index using local linear regression (width = 0.1)

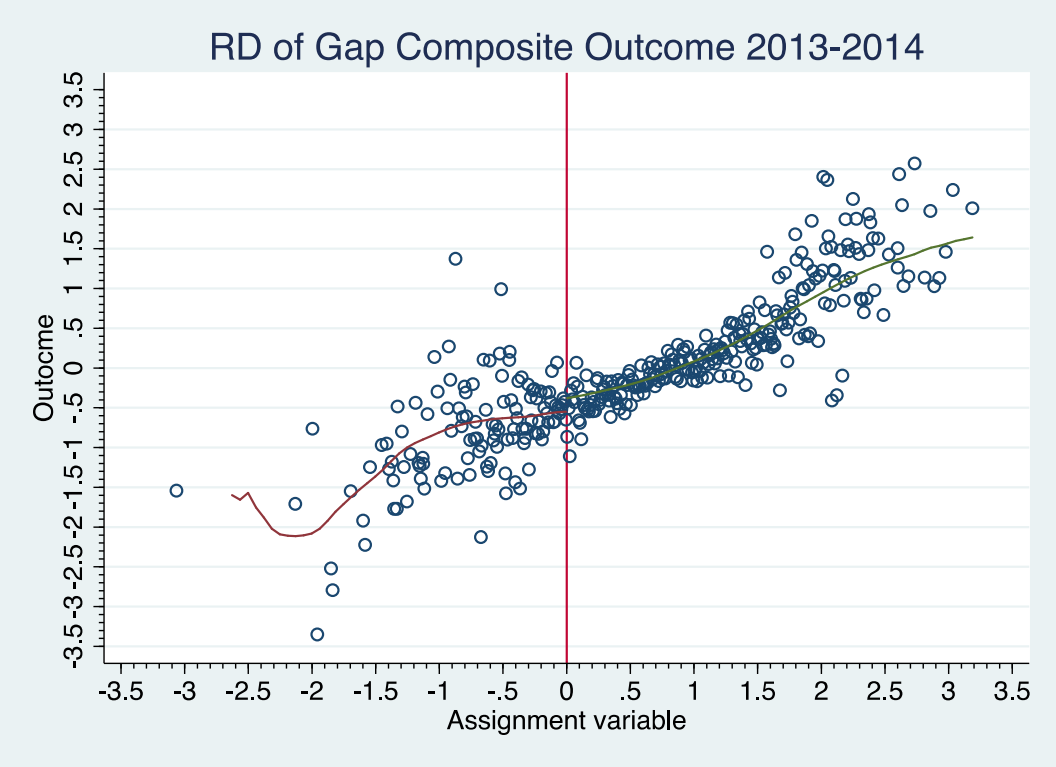


Figure 4. Fourth order polynomial overlaid on local linear averages (Coefficient = 0.081; p = 0.01)

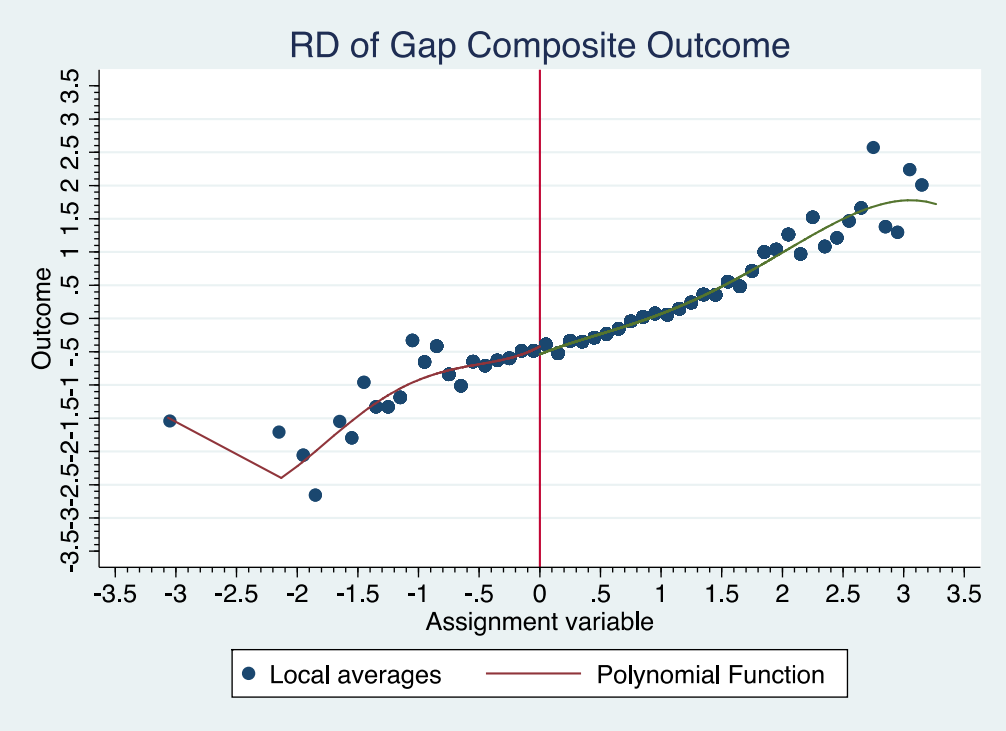


Figure 5. Graphical depiction of sorting at the cutoff (McCrary density test)

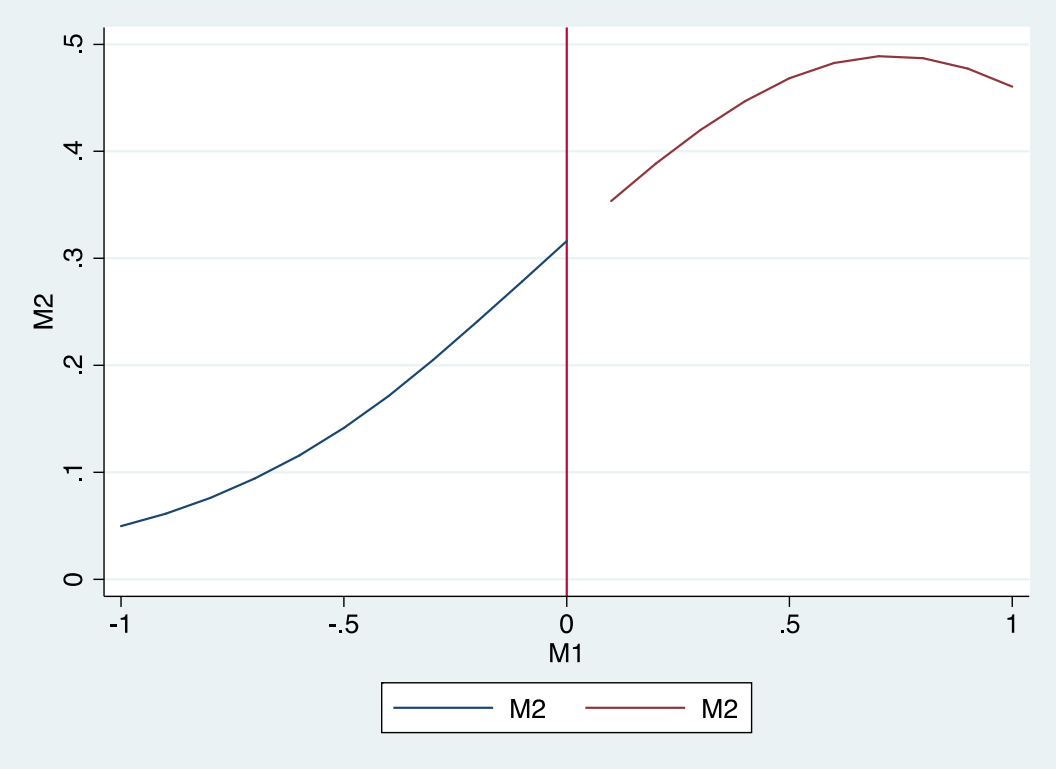
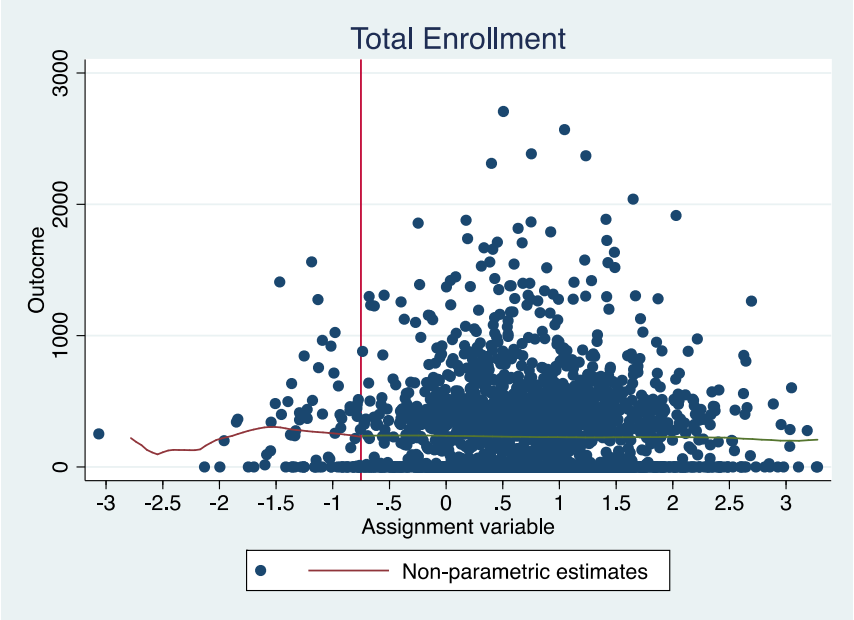
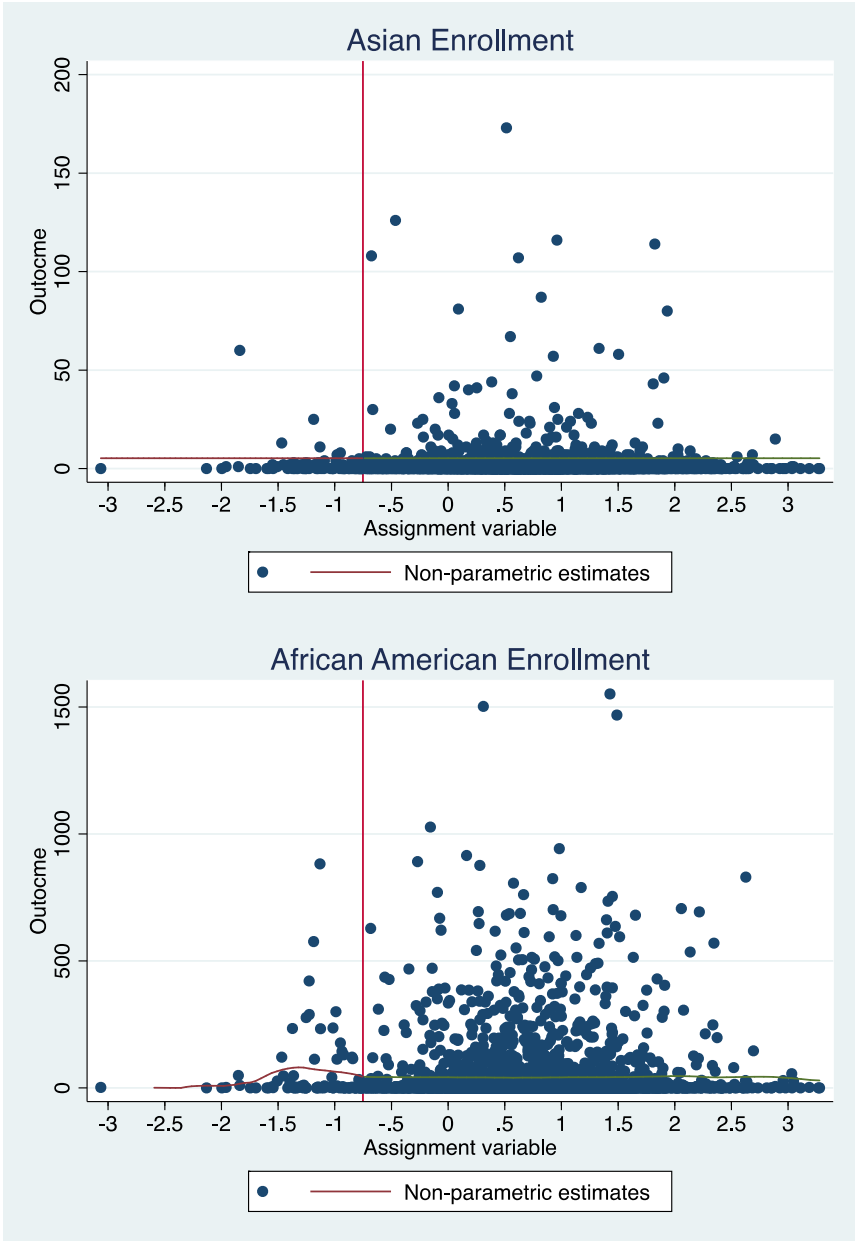
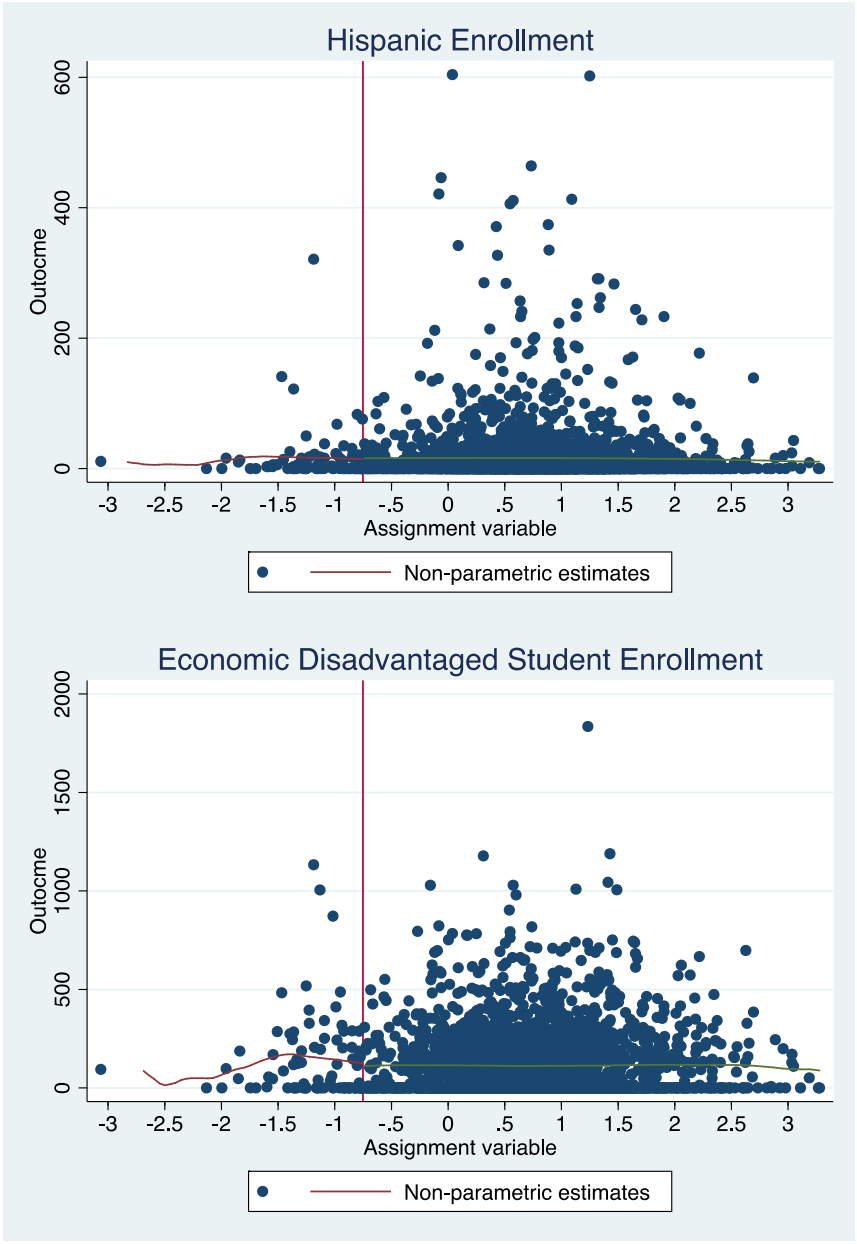


Figure 6. Discontinuity of pre-treatment covariates on outcome (p>0.10)




Female Enrollment
Outocme
0
500
1000
1500
-3
-2.5
-2
-1.5
-1
-.5
0
.5
1
1.5
2
2.5
3
Assignment variable
Non-parametric estimates
American Indian Enrollment
Outocme
0
50
100
150
200
-3
-2.5
-2
-1.5
-1
-.5
0
.5
1
1.5
2
2.5
3
Assignment variable
Non-parametric estimates


Asian Enrollment
Outocme
0
50
100
150
200
-3
-2.5
-2
-1.5
-1
-.5
0
.5
1
1.5
2
2.5
3
Assignment variable
Non-parametric estimates
African American Enrollment
Outocme
0
500
1000
1500
-3
-2.5
-2
-1.5
-1
-.5
0
.5
1
1.5
2
2.5
3
Assignment variable
Non-parametric estimates


Hispanic Enrollment
Outocme
0
200
400
600
-3
-2.5
-2
-1.5
-1
-.5
0
.5
1
1.5
2
2.5
3
Assignment variable
Non-parametric estimates
Economic Disadvantaged Student Enrollment
Outocme
0
500
1000
1500
2000
-3
-2.5
-2
-1.5
-1
-.5
0
.5
1
1.5
2
2.5
3
Assignment variable
Non-parametric estimates

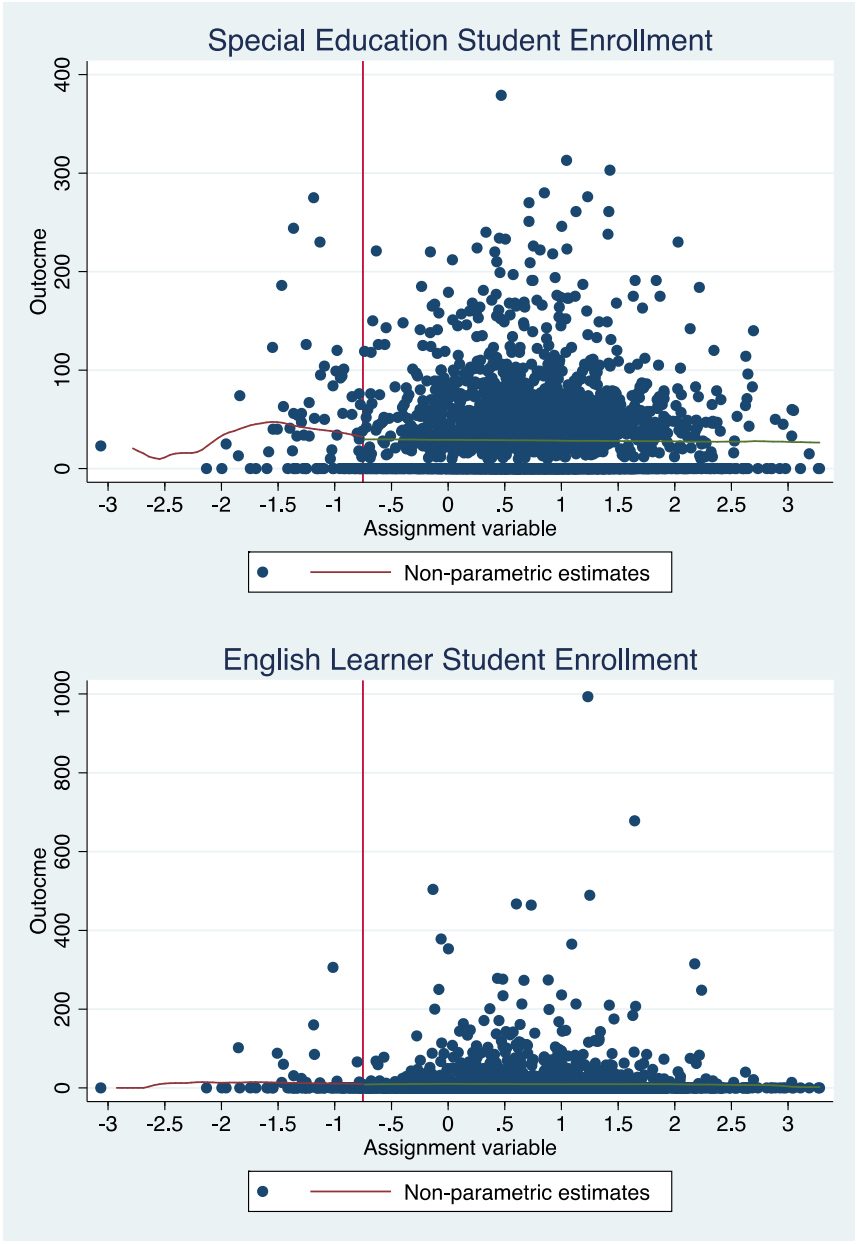

Special Education Student Enrollment
Outocme
0
100
200
300
400
-3
-2.5
-2
-1.5
-1
-.5
0
.5
1
1.5
2
2.5
3
Assignment variable
Non-parametric estimates
English Learner Student Enrollment
Outocme
0
200
400
600
800
1000
-3
-2.5
-2
-1.5
-1
-.5
0
.5
1
1.5
2
2.5
3
Assignment variable
Non-parametric estimates

Figure 7. Graphical analysis of regression discontinuity of School Performance Index, 2013 using local linear regression (width = 0.1)

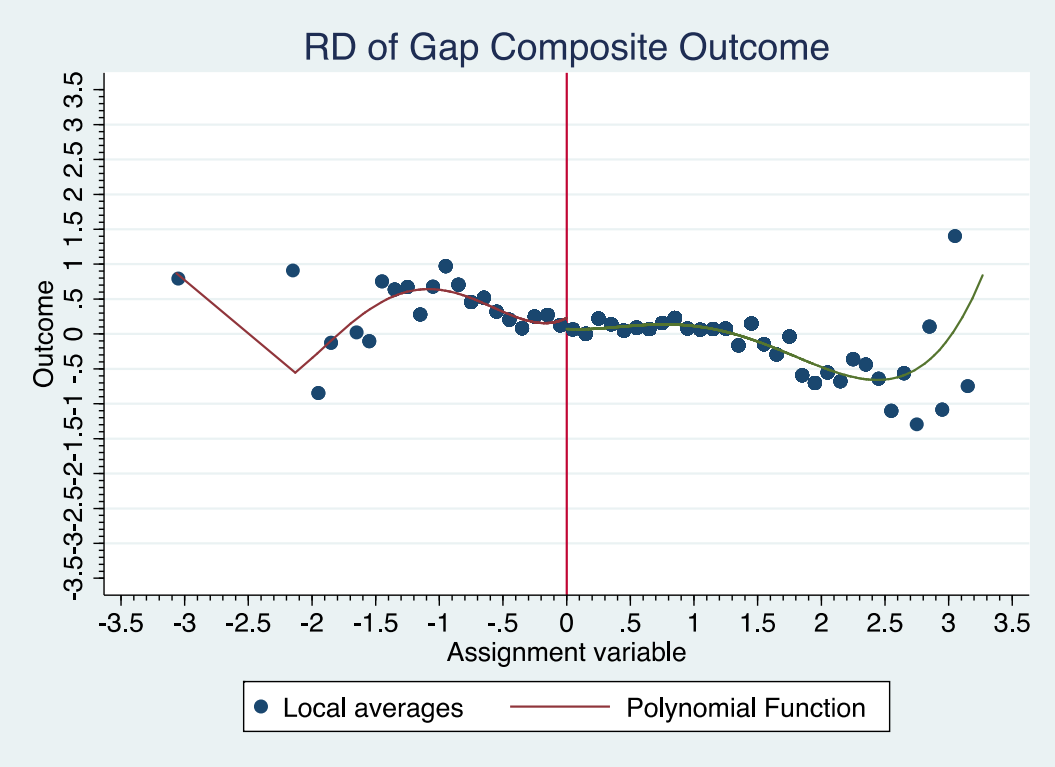


Table 2. OLS and Parametric Effects of Focus Identification on 2013-14 Gap Composite Measure

	Linear	Quadratic	Cubic	Fourth-Order Polynomial	Fourth-Order Polynomial w/ Covariates
Focus Status	0.199***	0.072	-0.035	0.119	0.293*
	-0.05	-0.07	-0.08	-0.1	-0.14
Gap Composite Score (GCS)	0.743***	0.466***	0.341*	0.738*	1.260**
	-0.02	-0.07	-0.16	-0.29	-0.4
GCS*Focus	-0.14	0.048	-0.415	0.274	0.908
	-0.08	-0.18	-0.36	-0.68	-0.96
GCS^2		0.128***	0.255	-0.415	-1.308*
		-0.03	-0.14	-0.44	-0.6
GCS^2*Focus Status		-0.181	-1.030**	1.959	5.860**
		-0.1	-0.38	-1.26	-1.79
GCS^3			-0.033	0.359	0.887**
			-0.04	-0.25	-0.33
GCS^3*Focus Status			-0.168	0.97	2.644*
			-0.1	-0.79	-1.12
GCS^4				-0.072	-0.164**
					-0.06
GCS^4*Focus Status				0.365*	0.893***
					-0.21
	-	-	-		
Constant	0.617***	0.510***	0.481***	-0.540***	-0.375*
	-0.02	-0.04	-0.05	-0.06	-0.15
Covariates?	N	N	N	N	Y
Adj R-squared	0.36	0.37	0.37	0.37	0.4
AIC	4671.67	4658.24	4657.05	4654.18	2423.04
BIC	4695.1	4693.38	4703.89	4712.73	2521.98
N. of cases	2581	2581	2581	2581	1349

* p<0.05, ** p<0.01, * p<0.001

Table 3. Non-Parametric Estimates of Effect of Focus Identification on 2013-14 Gap Composite Measure

Bandwidth	Treatment Estimate	Standard Error
0.3	-0.04	0.26
0.6	-0.05	0.37
1.2	-0.08	0.22

Note: 0.6 is the optimal bandwidth using the "plug-in" method to determine optimal bandwidth (see, e.g., Imbens & Kalyanaraman, 2009). Epanechnikov kernel function are used for non-parametric estimates. * p<0.05, ** p<0.01, * p<0.001

Table 4. Estimation of Treatment Effects with Truncated Samples

	Polynomial	Gap Composite within 1	Gap Composite within 2	Gap Composite within 3
Focus Status	0.119	0.367	0.082	0.105
	-0.1	-0.22	-0.13	-0.11
Gap Composite Score (GCS)	0.738*	13.993*	-0.161	0.58
	-0.29	-6.26	-0.9	-0.42
GCS*Focus	0.274	-13.169*	1.801	0.432
	-0.68	-6.37	-1.22	-0.75
GCS^2	-0.415	-137.328*	1.995	0.009
	-0.44	-55.64	-2.78	-0.8
GCS^2*Focus Status	1.959	138.024*	1.35	1.536
	-1.26	-55.74	-3.43	-1.42
GCS^3	0.359	374.287**	-1.868	-0.025
	-0.25	-142.23	-3.25	-0.57
GCS^3*Focus Status	0.97	374.085**	4.846	1.354
	-0.79	-142.27	-3.65	-0.94
GCS^4	-0.072	0	0.571	0.037
	-0.05	(.)	-1.27	-0.14
GCS^4*Focus Status	0.365*	-0.14	0.179	0.256
	-0.15	-1.07	-1.34	-0.2
Constant	-0.540***	-0.795***	-0.460***	-0.526***
	-0.06	-0.2	-0.09	-0.07
Adj R-squared	0.37	0.09	0.21	0.33
AIC	4654.18	1142.83	3618.44	4585.79
BIC	4712.73	1182.09	3674.86	4644.21
N. of cases	2581	580	2085	2544

* p<0.05, ** p<0.01, * p<0.001

Table 5. OLS and Parametric Effects of Focus Identification on 2013-14 School Percentile Index

	Linear	Quadratic	Cubic	Fourth-Order Polynomial	Fourth-Order Polynomial w/ Covariates
Focus Status	-0.041	0.029	0.126	0.159	0.249
	-0.07	-0.09	-0.11	-0.14	-0.19
Gap Composite Score (GCS)	-0.219***	0.157	0.813***	-0.11	0.091
	-0.03	-0.1	-0.21	-0.4	-0.54
GCS*Focus	-0.003	-0.739**	-1.786***	1.056	1.093
	-0.1	-0.25	-0.5	-0.94	-1.3
GCS^2		-0.174***	0.833***	0.721	0.507
		-0.04	-0.2	-0.6	-0.81
GCS^2*Focus Status		-0.042	0.138	2.68	4.299
		-0.13	-0.53	-1.72	-2.42
GCS^3			0.174***	-0.737*	-0.619
			-0.05	-0.34	-0.45
GCS^3*Focus Status			-0.307*	3.306**	4.406**
			-0.14	-1.07	-1.52
GCS^4				0.168**	0.147
				-0.06	-0.08
GCS^4*Focus Status				0.348	0.624*
				-0.2	-0.29
Constant	0.231***	0.085	-0.065	0.071	0.214
	-0.03	-0.05	-0.07	-0.08	-0.2
Covariates?	N	N	N	N	Y
Adj R-squared	0.03	0.03	0.04	0.04	0.05
AIC	6303.72	6287.896	279.01	6268.42	3237.55
BIC	6327.14	6323.036	325.85	6326.98	3336.48
N. of cases	2581	2581	2581	2581	1349

* p<0.05, ** p<0.01, * p<0.001

Chapter 3

Examining the Implementation of a Statewide System of Supports for Focus Schools in Michigan

> If the design puzzle was to structure conditions among schools, designs, organizations, and environments in order to increase the potential for effectiveness, then the implementation puzzle was to manage interactions within and among these four domains over time in order to realize that potential. . . . . .
>
> — Cohen, et al., *Improvement by Design*, p. 61 (2014)

> I think being identified as a Focus School increases the sense of urgency for a [school] building to say, "Oh, what the heck. We've been identified. What do we need to do?"
>
> — District-level interview respondent

Abstract

This chapter provides evidence of the differential ways in which supports for Focus Schools were implemented during the 2012-14 school years. Using data from approximately 60 hours of observations in schools, districts, ISDs, and state-led meetings over the course of the 2013-14 school year; nine semi-structured interviews with school, district, ISD and state-level personnel; one focus group of 10 school, district, ISD, and state-level personnel; and documentation on Focus School supports and interventions provided by the Michigan Department of Education, this study examines factors that influence implementation of district-provided supports to reduce within-school inequality. The accountability system to identify and support Focus Schools was largely developed in a way that gave power to districts to implement reform despite a historically weak track record of doing so in the past. Given leeway to respond

to this open design, school districts employed one of several strategies. Of the two districts studied here, one took a more coordinated approach while the other deferred responsibility to support its Focus Schools to ISD personnel. The two districts are more urban and have more capacity than other districts in the state, however, indicating that districts without historical support from leadership and personnel with knowledge to implement complex initiatives may use strategies not discussed in this study.

Study results indicate that while there was initial hesitation about Focus identification, many school and district personnel eventually saw Focus status as an opportunity to improve educational opportunities for struggling students. Second, results indicate that identification itself created pressure for compliance, but that identification plus support was perceived to yield better results with respect to reducing within-school achievement gaps. Third, coordinated capacity—that is, a systematic approach to the intervention that couples capacity at different governance levels in service of the same goal—may mediate the extent to which supports were implemented in such a way to strengthen policy effects. Taken together, these findings have implications for the ways accountability policy is implemented. Specifically, educational systems in the U.S. need a mechanism or tool to develop this coordinated capacity. Chapter Four explores one potential way this coordinated capacity may be developed through Networked Improvement Communities utilizing improvement science methods.

Introduction

In 2011, the U.S. Department of Education (ED) announced an opportunity for states to apply for waivers from select provisions of the Elementary and Secondary Education Act (ESEA), specifically those enacted in the 2002 reauthorization more commonly known as No Child Left Behind (NCLB). In exchange for flexibility from federal regulations, ED required the

state education agency (SEA) to demonstrate its ability to continue to work towards the goals set forth in the ESEA. These new requirements included the identification of Focus Schools, or high achievement-gap schools, and the provision of supports to reduce inequality. The plan for this and other ESEA waiver requirements was to be detailed in each state's waiver application and submitted to ED for review and approval.

The window for ESEA waiver applications from the time President Obama announced the waivers in September 23, 2011 to submission in Michigan was approximately five months: Michigan submitted its first waiver proposal on February 28, 2012 and received notification of its approval on July 19, 2012. In this short time frame, state department of education personnel undertook the task of reading and understanding the requirements set forth in the waiver, coordinating a response across internal departments, and writing and submitting a waiver application that conveyed this plan to application reviewers from ED. In Michigan, as in the other 43 states with approved ESEA waivers, the proposed method of Focus identification was a departure from previous methods of identifying inequalities in student achievement.[1] Furthermore, schools in the Focus set did not have previous experience with being identified as high achievement-gap schools. As a consequence, many practitioners and administrators had to adapt old strategies and practices or invent new ones to incorporate these new provisions into their daily routines.

This study examines the way individuals and organizations responded to Focus School identification in Michigan during the 2012-13and 2013-14 school years given its quickly developed and implemented origins. I use the case of Focus School practitioners and support providers in two mid-sized districts in Michigan to understand variation in response to the

[1] See Chapter 2 for a more in-depth discussion of this point.

policy, paying particular attention to the role played by organizational capacity as a mediating factor. The way educators and educational organizations responded to this policy provides an opportunity to examine if and how accountability policies to reduce within-school achievement gaps influence the daily work of schools, districts, intermediate school districts, and state personnel.

Background

The State Role in Focus Identification and Support

Pressure to reform aspects of No Child Left Behind had been mounting prior to the announcement of ESEA waivers for state departments of education. Conventional wisdom dictated that the proposals for reform should be incorporated into the reauthorization of the ESEA, a Congressional process that occurs once every five years. However, efforts to reauthorize the law were not gaining traction under a deadlocked Congress, which meant that the 2002 iteration of the legislation, NCLB, continued to be the law of the land despite containing what some critics saw as outdated or outmoded, overly punitive stipulations. ESEA waivers were issued in part because of this Congressional gridlock coupled with mounting pressure to enact change.

Given the desire for states to have some reprieve from NCLB regulations, such as developing and tracking annual measureable objectives (AMO), reception to the ESEA waivers on behalf of state personnel in Michigan was initially positive. Over time, though, the waivers caused consternation among state personnel as information was released in waves regarding parameters of the new regulations and timetables for submission and approval. State personnel described this time as one of fast-paced change, requiring SEAs to be more nimble in response to federal regulations than they had to be in the previous decade since the roll out of NCLB.

It is important to note that the parameters of Focus identification and support in Michigan, as in many other states, occurred under a constrained set of circumstances, including limited SEA resources, personnel, and capacity; lack of developed core technology for complex tasks at the state or local levels; and a strong, persistent tradition of local control. Taken together, these factors led to an imperfect system of accountability for Focus Schools. Practitioners and administrators at the school, district, and ISD levels then had to react to this system at the same time that state personnel requested and incorporated feedback to the identification strategy and supports provided.

Limited SEA resources, personnel, and capacity. Since the 1980s, the state role in education has been expanding from one of program administration to compliance, and most recently, to support (Weiss & McGuinn, 2016). This expanded role has not been accompanied by a commensurate increase in funding, staffing, or expertise (Boyle, LeFloch, & Therriault, 2008). In a survey of SEA personnel in the 2011-12 school year, the majority of states reported at least a 10% *decrease* in funding due to budget cuts resulting from the 2008 economic downturn (Kober & Rentner, 2012). Furthermore, 24 of 38 states surveyed reported taking actions to reduce or re-allocate SEA personnel as a result of these budget cuts, though there is some evidence that states are re-allocating remaining personnel to areas critical to federally legislated educational reforms, such as statewide longitudinal data systems and supporting low-performing schools (Kober & Rentner, 2012).

Michigan specific data corroborates this trend, both with respect to funding levels as well as the number of SEA personnel. In Michigan, appropriated 2012 levels of spending on education were reported to remain below 2008 spending levels (Oliff & Leachman, 2011).[2] This

[2] http://www.cbpp.org/sites/default/files/atoms/files/9-1-11sfp.pdf

overall decline in statewide spending on education was also reflected specifically in spending for school and district improvement. In a review of ten states' expenditures on school and district improvement, Michigan had the lowest per school expenditure at $9,172 per school, less than half the ten-state average of $23, 912 and almost one-fifth of Connecticut's expenditure on school and district improvement of $44, 885 per school (Joachim & Murphy, 2013).[3] A 2004 report to the state board of education voluntarily commissioned by the Michigan Department of Education with the Council of Chief State School Officers to examine NCLB implementation reported then-current SEA staffing levels at 394 full-time employees (FTEs), down from a peak of 2,622 FTEs in 1980.[4] While the optimal size of an SEA is unclear, it is certainly true that as the role of the SEA has become more comprehensive and complex, the number of SEA FTEs has declined considerably and continues along this trajectory: in 2012, the first year of Focus identification, MDE was reported to have 387 FTEs.[5]

It could be true that decreased funding and declining personnel numbers have shed inefficiencies in the statewide bureaucracy and led to a more productive agency. However, interviews with and surveys of SEA personnel indicate a perception of limited capacity to adequately address the additional requirements placed on them by federal legislation. Over 40% of 38 states surveyed reported not having adequate staff expertise to support low-performing schools (Kober & Rentner, 2012). Efforts to recruit and retain expertise can be a challenging endeavor for state departments of education due to hiring restrictions placed by state appropriation legislation, lower salaries than surrounding industries for civil servants, and lack of a pipeline from which to nurture and cull talent (Weiss & McGuinn, 2013). This problem is

[3] http://www.crpe.org/sites/default/files/pub_capacity%20challenge_dec13_0.pdf
[4] https://www.michigan.gov/documents/ITEM_A_89910_7.pdf
[5] http://www.michigan.gov/documents/mde/Item_G_HR_Report_April_2016_519996_7.pdf

exacerbated by the fact that the core technology of supporting low-performing schools is still under development and often not easily implemented in a complex policy environment.

Lack of core technology for complex tasks. This reduced number of staff personnel coincided with being asked to perform increasingly complex tasks. For example, under NCLB, states were required to identify schools that were not making adequate yearly progress (AYP) on a number of different metrics and enact a tiered system of consequences if schools did not comply. If schools were persistently identified as low performing, they were required to adopt one of four turnaround school models to improve student achievement. These tasks, while difficult in a number of ways, such as developing a statewide longitudinal data system to identify schools and complying with newly mandated federal reporting guidelines, were at least prescriptive in the sense that they required states to follow a formulaic vision of identification and support. The ESEA waivers, by comparison, awarded states flexibility to develop a formula whereby three types of schools—Priority, Reward, and Focus—would be identified, in addition to detailing how Priority and Focus Schools would be supported once identified.

As has been discussed before, Focus Schools in particular were a novel designation in that high achievement schools had not been identified by states prior to the ESEA waivers. The core technology of how a school should reduce within-school inequality or how states should support schools in that endeavor had not yet been created, or at best, was very much in the development stage. SEAs had no "off-the-shelf" intervention they could deploy to address the ESEA provisions and therefore had to create such an intervention based on prior practice and knowledge and professional capacity. However, SEAs had diminished personnel to accomplish these more complex tasks, many of which, as in the example of identifying and supporting Focus Schools, they had not done before.

Strong tradition of local control. Along with reduced capacity to perform complex and novel tasks, SEAs had to contend with another issue in its proposal to design and adopt a new accountability system under the ESEA flex waivers: a legally binding tradition of local control. It has been widely documented that the U.S. education system is one that strongly values local control at the school, community, and district levels and resists efforts to consolidate power at the state or federal levels (Cohen & Moffitt, 2009). The historical consequences of local control include a disconnect between pedagogical training, curricula, and professional development for educators and increased susceptibility to local political whims, among others (Cohen & Bhatt, 2012).

While local control is an impediment to designing statewide systems of supports across the country, it is a particularly strong design challenge in Michigan. It has been documented that Michigan has one of the most highly decentralized educational systems in the country (McLaughlin & Talbert, 2001). More than a mere historical or cultural legacy, this tradition of local control has been codified into law through the Headlee Amendment of 1978. This amendment to the Michigan Constitution, approved by Michigan voters, created an important stipulation in the relationship between state and local actors. The one most pertinent to the provision of statewide programming in education is Section 29, which prohibits "unfunded mandates," or the ability of states to require actions of local entities without fully funding the effort. Focus Status through the ESEA waiver required a 10% building set-side for Title I Focus Schools. Beyond this allotment of federal funds, however, the state's fiscal woes and declining revenues made funding new policy and program initiatives difficult. In the case of Focus Schools, this meant that any requirements asked of Title I Focus Schools must be modest and fit

within a small operating budget and that any requirements asked of non-Title I Focus Schools did not have legal consequences for lack of compliance.

Designing a System of Accountability for Focus Schools in Michigan

The reality of SEA's ability to accomplish the complex task of designing an accountability system with respect to personnel and organizational capacity had two consequences. First, any design of an accountability system in Michigan proposed under the ESEA Flex Waivers had to hew to these constraints, leading to a system with a limited state role. Consequently, MDE proposed a system whereby the state used the SLDS to identify Focus Schools but devolved the provision of support to districts, beyond a blueprint of suggested parameters. The design of this type of "hands off" accountability system is not unique to Focus School policy and has precedent in education (see, e.g., the roll out of Title I itself in Cohen & Moffitt, 2009). As we see, however, this ubiquity does not imply that a successful solution set to this design-problem (or lack thereof) exists nor is such a design fatal. The consequences of this type of policy design on implementation are discussed throughout the study.

Second, the initial accountability system for identifying and supporting Focus Schools was necessarily imperfect and changed over time to reflect evolving priorities, understanding, and capacity. For example, the first cohort of Focus Schools identified in 2012 using 2011-12 data were defined as the 10% of schools with the largest gap between the top 30% and bottom 30% of students. As mentioned in the previous chapter, some schools protested identification under this system because even their bottom 30% of students performed relatively well compared to the state average. As a consequence, the cohort of Focus Schools identified two years later in 2013-14 excluded schools whose average school achievement was at or above the state average for student performance. In another example, schools protested being identified

annually because their bottom 30% of students was improving at a faster rate than the state average. As a result of these concerns, the 2013-14 definition of Focus Schools excluded identifying those schools whose students in the bottom 30% were improving at or above the state average rate of improvement.[6] These revisions in Focus definition alleviated concerns that higher performing Focus Schools were diverting resources from schools and students with greater need while rewarding schools for improving by excluding them from the accountability system. The revisions also reflect the constant evolution of SEA capacity to interpret ESEA provisions and apply them to the Michigan context.

The District Role in Focus Identification and Support

Michigan's model for providing supports to Focus Schools is concentrated at the district-level, which is in contrast to other states' approaches (e.g., in Minnesota supports provided to Priority and Focus Schools flow through the six regional centers of excellence that bypass districts and work directly with schools). There is some evidence that supports for struggling schools at the district level are more efficacious than efforts for school turnaround that are conducted at the school level, particularly with respect to building instructional capacity, aligning teaching and coaching to new curricula and standards, and supporting principals' efforts (Rhim & Redding, 2014; Zavadsky, 2012; Marsh, et al., 2005). However, the extent to which this model has been successful under an accountability framework is variable (see, e.g., Mintrop & Truijllo, 2007). Furthermore, whether these variable district effects are similar for accountability policies that aim to help schools improve within-school achievement gaps is not widely documented.

[6] Chapter 1 uses only the definition used to identify the 2012 cohort using 2011–12 data. Because definitions were not changed until 2013–14, I only examine effects of initial Focus identification in the first two years of implementation.

A 2013 report put out by the Center for Reinventing Public Education categorizes Michigan's approach towards supporting low-performing schools within the parameters of the ESEA waivers as parallel strategies: 1) a statewide takeover or "all-in" strategy represented by the Education Achievement Authority (EAA) in Detroit and 2) a "results without rancor" strategy to support Focus Schools by developing relationships between MDE and local education agencies (Murphy & Rainey, 2012).[7] The former is a strategy now employed through the School Reform Office for the lowest performing schools, or Priority Schools, and offers a comparative vision for how statewide supports might be organized.[8] Studies of an intensive, statewide approach to "fixing" low-performing schools, such as that of the Tennessee Achievement School District, have not posted dramatic gains in student achievement and often suffer from many of the same capacity issues that districts face, "including expertise, experience, money, and strong leadership" (Glazer & Egan, 2016).[9] Focus Schools, in contrast, do not report to MDE, nor are they perceived as being "taken over" by the state. Instead, the MDE model seeks to build expertise at the district level in order to facilitate the reduction of within-school achievement gaps. As Murphy and Rainey (2013) point out, this model is optimistic in its assumption that infrastructure can be built at the local level despite historical evidence of weak capacity or unwillingness to do so in the past. They conclude, "The external political threats to [the Results Without Rancor model] stem from the fact that things stay the same more than they change.

[7] Murphy and Rainey (2012) also discuss a third model of SEA-LEA relationships termed the "Bounded Disequilibrium" model in which the SEA plays more of a directive role in structuring incentives and disincentives for the local education agency. The authors conclude Michigan is not currently pursuing this strategy.
[8] In 2010, the Michigan legislature passed MCL 380.1280c, which created a statewide accountability system reporting to the newly formed School Reform Office (SRO) for the lowest-performing 5% of schools in the state. In 2015, Governor Rick Snyder issued Executive Order 2015-9, which transferred the SRO from MDE to the governor's office. Consequently, the SRO manages Priority Schools in Michigan while Focus Schools are supported by MDE.
[9] Since 2003, districts run by a state entity to improve low-performing schools have operated in four states: Michigan, Nevada, Louisiana, and Tennessee (Glazer & Egan, 2016).

Since the theory of action does little to directly disrupt the status quo, it is less likely that opposition to the reforms will emerge."

Focus Supports: A Flexible Blueprint for Change

Focus schools identified in 2012 received a bundled set of services that constitute the intervention discussed in Chapter 2. These include:

- Participation in the Superintendent's Dropout Challenge, where schools identify at least 10-15 students at any grade level who are at risk of dropping out and provide them with focused supports
- Participation in data dialogues led by the district to identify at least two teaching and learning priorities on which to focus change efforts
- Revision of the School Improvement Plan (SIP) to incorporate the teaching and learning principles identified during the data dialogue

In addition, Title-I schools were required to:

- Identify local school boards of Focus Status
- Identify parents of Focus Status
- Use the district set-aside of Title I funds (at least 10% of previous Focus School annual budget) for pre-approved uses, including providing a multi-tiered system of supports for the lowest performing students; providing weekly or daily time for teacher collaboration; administering the Survey of Enacted Curriculum; conducting a needs assessment; supporting professional learning for MI-ACCESS; initiating culture and climate interventions.

While Focus Schools are identified at the school level, supports are funneled through districts. Consequently, districts with Focus Schools were required to:

- Conduct data dialogues with Focus Schools to identify at least two teaching and learning priorities to support improvement for the bottom 30% of students
- Revise its District Improvement Plan (DIP) to delineate how it will support Focus Schools in the district
- Provide technical assistance to Focus Schools to improve their reform efforts
- Monitor Focus Schools' School Improvement Plans (SIP)
- Set aside at least 10% of a Title I Focus School's previous annual operating budget using Title I money.

Taken together, this bundled set of interventions aims to provide a multi-tiered system of support (MTSS) to students in high achievement-gap schools. While these interventions are presented as a discrete list of requirements, the fact that they were all administered through one programming body, MIExcel, meant that Focus status became a funnel for supports that flowed from the district to the school with cohesive guidance and opportunities for organizational learning flowing from the state through the ISD. This theory of action authorizes local districts to support high achievement-gap schools in identifying students who are struggling and develop interventions that are appropriately targeted to that group of students. The flexibility embedded in these supports allows localities to determine what is most appropriate for the struggling students who have been identified, and then monitor implementation through the SIPs and DIPs. It also codifies collaboration between the school and district for data review to identify principles of teaching and learning. Finally, the intervention provides monetary resources for implementation for Title I schools and provides a list of possible options for spending those resources.

As discussed earlier, however, these parameters for how districts should provide supports to Focus Schools grant a lot of leeway that allows local entities to respond in various forms. This hands-off approach is intentional. The devolution of authority to local entities is a policy design that is particularly useful when the core technology and capabilities are weak, as a state entity has little direction to validly give and local entities would have little capability for uptake. Consequently, this set of bundled services may be a weak intervention for the bottom 30% of students. First, the policy design puts the onus for improvement on the very localities that likely contributed to creating the achievement gap in the first place, whether intentional or not. The expectation that these same schools and districts can then work together to ameliorate this gap is just that—a presumption—and one that has not yielded fruit in past reform efforts (Cohen and Moffitt, 2009). Second, the lack of prescription of what to do with the monetary resources or what types of teaching and learning priorities to identify suggest that schools and districts must have the capacity to conduct such a data review, articulate pedagogical practices to implement, and support teachers in that implementation. Third, there must be organizational infrastructure in place to support these kinds of reform efforts. For example, if schools and districts are to collaborate in conducting a data dialogue, there must be sufficient time for central office administrators and school-level personnel to work together to conduct the data review and identify principles of teaching and learning. Similarly, there should be time for teachers to collaborate in the implementation of those teaching and learning principles. Fourth, even if these capabilities existed, both at the individual as well as the organizational levels, there must be buy-in for the endeavor from both individuals and organizations and a belief that the way Focus Schools are identified is valid and reliable.

This study suggests that while buy-in for Focus identification has grown over time, there is variation in the extent to which districts are able to develop coordinated capacity to support Focus Schools in reducing within-school achievement gaps. In addition, efforts to build learning organizations that use data as an instrument to collaborate with other learning organizations yielded more coordinated capacity and higher perceptions of successful implementation of Focus supports. The following sections provide a framework and research questions to define the parameters of the study, discuss the data used to draw conclusions, and review each finding and its supporting evidence.

Research Questions

Given that 43 states are implementing such policies across the nation, it is worth investigating what factors contribute to the distribution of effects around the potentially null average discussed in Chapter 2.[10] There is a plethora of research documenting the difficulty of bringing about changes in educational practice, indicating that policy implementation often thwarts or subverts policy aims (Pressman & Wildavsky, 1973), or from a less technocratic perspective, adapts policy aims and goals to local contexts accounting for resource and capacity constraints (Honig, 2003). Another option is that policies fail to provide effective instruments for execution of policy goals.[11]

In *Improvement by Design*, authors Cohen, Peurach, Glazer, Gates, and Goldin (2014) outline a series of "puzzles" presented by organizations that sought to develop educational systems. Their work with three unique educational systems—America's Choice, Success for All, and the Accelerated Schools Project—led to the conclusion that building educational

[10] See, e.g., an editorial post by Brian Jacob for Brookings on "Harnessing the Value of Failure," available at http://www.brookings.edu/research/papers/2015/12/03-harnessing-value-failure-jacob

[11] For a complete discussion of the literature on implementation research, please see Chapter 3.

infrastructure to improve low-performing schools allowed these organizations to overcome previously identified weaknesses in the U.S. educational system, such as weak capability and local control. However, it also led to the creation of new challenges such as more coordination, management, resources, and time. Therefore, what the authors refer to as an "unprecedented" approach to building educational systems created questions about organizational capacity and learning to implement developed tools such as curriculum, professional development, and supporting materials to affect student outcomes.

One might argue that building a statewide system of support (SSOS) for Priority and Focus Schools under ESEA was similar to building an educational system in that the SSOS in a given state was required to follow the principles of "improvement by design" with the caveat that the "design," as demonstrated previously, is an intentionally hands-off approach that devolves authority to local entities. In this case, however, the goal was not to improve the lowest performing schools but instead to reduce within-school inequality at the school level. Indeed, the authors argue that this is the primary strategy employed by SSOS program managers who employ a multi-tiered system of support (MTSS) to improve student outcomes. As one interview respondent put it:

> [T]he Department of Education has been pretty clear that what they want people to do is a multi-tiered system of support. They want them to use the school improvement process. They've put together state-level training that's consistent across the whole state. This is the process, these are the steps we want you to use. There's a common template for school improvement that's evaluated, and they've said we really want your overarching school reform to include a multi-tiered system support…. Really the whole state is moving in this—just it looks different in each district depending upon the resources and the population. (Interview, district-level personnel)

Under this framework, "improvement by design" is a strategy that is concerned with design, implementation, improvement, and sustainability of whole school reforms to improve

student outcomes. The series of studies encompassed in this dissertation is concerned with the first three with an eye towards the fourth. This chapter specifically deals with the implementation puzzle, which is framed as the way schools, policy or program design, environments, and organizations interact to enact the goals set forth by the policy or program design.

This study uses a qualitative case study design to explore the following questions:

1. Did the policy design elicit a desired response from schools and districts? If so, how?
2. What supports were provided to Focus Schools? How were these supports provided?
3. Were there differences in the ways supports were provided or implemented based on district or school characteristics? If so, what were they, and how did they influence implementation?
4. What organizational factors, such as school size, leadership, urbanicity, or Title I status, influenced the implementation of supports offered to Focus Schools?

Data and Methodology

To address these questions, I use data from field notes from 60 hours of observations over the course of the 2013-14 school year; nine semi-structured interviews with school, district, ISD and state-level personnel; a 10-person focus group; and documentation on Focus School supports and interventions provided by the Michigan Department of Education. I use a nested qualitative case study design to identify the themes across Focus Schools. According to Yin (2003), a case study design should be considered when (a) the focus of the study is to answer "how" and "why" questions; (b) you cannot manipulate the behavior of those involved in the study; (c) you want to cover contextual conditions because you believe they are relevant to the phenomenon under study; or (d) the boundaries are not clear between the phenomenon and context. In the Focus School context, the purpose of the study is to learn how supports for Focus Schools are

implemented in the context of concurrent reform efforts at the school-district and state levels. To understand factors that influence implementation of supports, understanding contextual conditions is essential.

To identify interview respondents, I first worked with the Michigan Department of Education data office to create a sampling frame of 15 schools in six districts within three ISDs. This frame varied public and charter status; urbanicity; school and district size; Title I status; and geographic location within the state. I then worked with MDE to recruit respondents from each ISD, district, and school, respectively. The final sample included two of the three original ISDs, two districts, and four schools. Several individuals opted not to participate due to competing interests and ongoing initiatives in their contexts.

Once interview respondents had agreed to participate in the study, interview data were collected using semi-structured interview protocols (see Appendix A). Interviews were conducted with teachers, school principals, district facilitators, ISD administrators, and state-level support teams. The hierarchical nature of identification and supports for Focus Schools make understanding the supports provided at each level of administration important to understanding how the phenomenon of implementation takes place. Interview protocols ask individuals about their Focus identification in their respective working contexts, role in supporting Focus schools, perception of facilitators, and barriers to reducing within-school achievement gaps in Michigan. Each interview lasted approximately 60 minutes and adhered to privacy and ethical considerations approved by the Institutional Review Board (IRB). Interviews were transcribed using a transcription service.

Field notes from school observations and from quarterly meetings of the MIExcel program to ISD-level support staff were also included in the analysis and employed a wide

variety of participant observation techniques (see, e.g., Emerson, Fretz, & Shaw, 2011). Documents were collected at events, through interview subjects, and via the websites of the Michigan Department of Education, ISDs, and districts in Michigan. For each source of data, I use open coding techniques rooted in ethnomethodology to analyze the data followed by hierarchical thematic coding (Emerson, Fretz & Shaw, 2011). Data were coded by hand using Excel to organize codes and then confirmed using NVivo 2011 software.

District Context

The schools and districts sampled in this study are concentrated in two mid-sized urban areas in Michigan: Riverside Public Schools and Lakewood School District.[12] Both districts serve approximately 11,000-12,000 students, a similar size to approximately 20% of all school districts in the U.S. This district size represents an important but often understudied tier of schools that have many schools in need of improvement but not large enough to garner the attention of external providers and research firms for support and evaluation.[13] On average, both districts are lower performing than the statewide and national average on metrics such as student achievement scores and graduation rates. In addition, both Riverside and Lakewood serve higher numbers of economically disadvantaged students, racial and ethnic minorities, and special education students than other districts in the state. While these characteristics separate Riverside and Lakewood from other districts in the state, they are representative of districts statewide and nationwide with schools in need of improvement, which are, by definition, lower performing and oftentimes serving diverse student populations. Both districts spend approximately $7,000 per

[12] District names have been changed to protect participant's identities. All data has been taken from publically available information through Michigan's MISchool Data portal available at: http://www.mischooldata.org.
[13] For example, the National Center for Education Statistics offers data on the 100 largest school districts in the U.S. at http://nces.ed.gov/ccd/pub_100_largest.asp The districts represented in this annual study are the most studied districts in the U.S.

student on instruction, which allows for a comparison in approaches to supporting high achievement-gap schools that is not skewed by uneven resources spent on teaching and learning.

Despite these similarities between the districts, there are important differences. Riverside is a slightly lower performing district than Lakewood, with an average graduation rate of 61%. On any given test and in any given grade, Riverside student performance on MSTEP, the statewide standardized assessment, hovers between 20-30% proficient students at best. Riverside also serves a very economically disadvantaged population (71%) and a higher number of special education students (15.8%). Black students make up the largest fraction of the student population (40%), while 20% are Hispanic, 5% are Asian, and approximately 25% are White. The student-teacher ratio at Riverside is approximately 23:1. Riverside's struggling student population has meant that the district has more Priority Schools than Focus Schools and spends the majority of its resources on supporting low-performing schools. Indeed, individuals within the district reported being less concerned with Focus status than Priority status given the urgency inspired by the latter.

In comparison, Lakewood's average graduation rate is 71%, while the state average four-year graduation rate is 80%. Its achievement scores in mathematics and ELA indicate higher percentages of proficient students across discipline and grade levels at 40% proficient, which is closer to the statewide average of 41% in mathematics and 46% in ELA. Lakewood serves a lower percentage of students who are economically disadvantaged (61%) and who receive special education supports (11%), though these proportions of students are still higher than the statewide average. The student-teacher ratio at Riverside is approximately 21:1, which is comparable to but slightly lower than that of Riverside. While Lakewood is, on average, a higher performing district than Riverside, it serves a diverse population of students and has had a larger

number of Focus Schools identified in comparison to Riverside. This difference may account for Lakewood's more coordinated and systemic approach to supporting Focus Schools in comparison to Riverside's approach of delegating Focus support responsibility to the ISD or dealing with each school on a case-by-case basis.

Another important distinction between the two districts is that while Riverside's population has been declining steadily in proportion to the declining statewide population, Lakewood School District's population has increased by almost 25% over the past decade. Lakewood's progressive investment in its educational infrastructure and industries were not affected by the 2008 economic downturn and may account for this differing trend. This qualitative difference in district approaches also may account for the differences in their approaches to supporting Focus Schools, as shown later in this chapter.

Findings

The "implementation puzzle" described in *Improvement by Design* discusses that the complexity of what might formerly have been called "weak implementation" is really the result of contingent interactions among schools, policy or program design, environmental factors, and organizations that surround and support the school, including but not limited to districts, ISDs, and states. The following table organizes the findings from this study using this framework suggested by Cohen et al. (2014). In the following section, I delineate each of these findings systematically and discuss implications of each for educational systems—particularly states—implementing reforms to improve outcomes for struggling students using whole-school reforms.

Table 6. Organizational Framework for Focus Implementation in Michigan

Interaction	Finding
School + Design	• Initial buy-in low due to: 1) Perception of gap composite score as unreliable metric; 2) Redistribution of Title-I funds from neediest schools • Buy-in improved due to: 1) Consistent re-identification of schools on annual basis; 2) Awareness of needs of struggling students in higher performing schools
School + Design + Environment	• Accountability was sufficient to spur compliance, even without monetary support. • Accountability plus monetary support was seen as important to reducing within-school achievement gaps in Focus Schools.
School + Design + Environment + Organization	• In both cases—accountability as well as accountability plus monetary support—coordinated capacity, or a systemic approach to the intervention that couples capacity at different governance levels in service of the same goal, was seen as a critical mediating factor in successful implementation. • Data use can serve as a common template for coordinating and building capacity.

Interactions: Schools and Policy Design

Buy-In for Focus Identification

Interview respondents reported having dual reactions to Focus identification that at times seemed at odds. These reactions also differed based on the level of governance. For example, initially, there was a sense among schools and districts—the levels of governance that had real implications of identification for their daily practice—that Focus Schools were identified because of a measurement fluke. In other words, there was not a lot of buy-in that the metric used to identify Focus Schools—the Top-to-Bottom Ranking—was a reliable way to measure within-school achievement gaps. Part of this pushback stemmed from the identification of schools that had never before been identified by the state accountability system because on average they were performing well enough to escape notice of state turnaround efforts for the lowest performing schools. Teachers also reported initial frustration on being identified, though this stance softened over time:

> It was a frustration for the teachers I think at first, finding out that we're a Focus School. There's a lot of pressure put on teaching with high-stakes assessments. But once they found out more and they realized this is something that we can deal with, and we have data that we can use, and we have a plan in place, and we can come up with something that would work, that was fine. It's always frustrating because I know the teachers work really hard trying to bring all the students up. But they'll do what needs to be done to try to get off the Focus School status. (Interview, school-level personnel)

One concern of determining the effect of Focus identification on within-school inequality is that Michigan has interdistrict and intradistrict choice. Title I Focus Schools were required to notify parents about Focus status. This could have meant that parents of students in Focus Schools removed their children, causing changes in student composition that affected causal estimates. However, principals and district leaders reported not hearing about any parents who removed

their students based on Focus Status. Said one ISD-level interview respondent: “They were just thankful their kid wasn’t in a Priority School.” Another principal corroborated this sentiment:

> From parents, I heard relatively little or nothing from our letters. I maybe had one or two comments, I would always present it each year at our parent-teacher organization or PTO. Those are all very involved parents. They're aware of what our teachers are doing. They're aware of our population and the diversity in our population. So I've had very little pushback or comments from parents to know how to answer that. (Interview, school-level personnel)

At the state and ISD levels, however, there was a sense that identification was an opportunity to leverage resources to try to support struggling learners who may not otherwise have received supports from traditional Title I funding or district support. Eventually, this perception filtered down to some schools and districts as well, as schools were continually identified on an annual basis even as safeguards were built into the methodology to exclude schools whose bottom 30% of students performed higher than the state average or schools that were able to demonstrate the bottom 30% of students in their schools were improving at or above the state average rate.[14] Consequently, in the second year, Focus Schools began to take the distinction more seriously, due to consistent re-identification. As one district interviewee put it:

> [When Focus designation first came out], it was kind of discussed and it was like, "Why are we being punished for diversity?" The schools that were identified [in our district] were our schools that had the highest scores. So they had a high percentage of students that are from middle-class families, where both parents had gone to college. And they also lived in neighborhoods—because we did a little redistricting so that we have a mix of students. And then we looked at all the other schools in the state that got a lot of Focus designation, and it was the same thing. Schools with diversity are being targeted. And so that was the discussion at first—kind of irritation. (Interview, district-level personnel).

The buy-in of the metric of Focus identification was closely linked to its consequences, particularly in the redistribution of resources that resulted for districts with Title I Focus Schools.

[14] These safeguards were put into place in 2013-14, two years after the initial cohort was identified. Chapter 2 does not accommodate these new criteria for identification and for exiting Focus status; therefore, the discussion of implementation focuses on the initial regulations for identification and exit.

The funding influenced principals to implement the supports mandated by Focus status. However, as demonstrated in a later section, even non-Title I schools reportedly complied with requirements of Focus status despite the lack of funding; consequently, I conclude that while funding was important, it was not an essential component of buy-in.

Equity versus Equality in Funding

One of the primary consequences of Focus School identification for Title I schools was that districts with Title I Focus Schools were required to set aside 10% of Title I money per building. This was a departure from previous funding mechanisms, in which all Title I eligible schools in a district received equal funding regardless of the needs of its students. In this way, participants saw Focus Identification as a move towards equitable funding in lieu of equal funding for all Title I schools within a district. As one ISD-level interviewee explained:

> It's definitely an equity over equality issue. That's what we did have happen … we had a lot of districts had flat funding formulas, so based on the student population you got a percentage of the general fund, and it was totally blind to any of the issues of neighborhood or the income make-up. A good example would be [X District], you can have one elementary where people have homes that are $250,000 and then you can have another elementary, which mostly is fed by trailer parks and apartments, getting the same exact funding even though the kids that are in those schools need completely different levels of support. So this—Focus [status]—helped change that, because even with [the implementation of the multi-tiered systems of support in 2010], there was still resistance to really changing their funding structure around, but once they were identified Focus, they finally admitted defeat at the central office and said "You know, we need to be a little bit more dynamic in how we fund our schools" (Interview, ISD-level personnel).

These views were echoed at the state level, which indicated a divide between sympathetic state- and ISD-level personnel for Title I monetary set-asides and district-level individuals who (at least initially) saw this policy as diverting funds from lower-performing schools. Here are two opposing viewpoints from the state and district-levels.

> At the district level, we're talking about, just because you get your funding in a pot doesn't mean you have to parse it out equally. And the idea in the ESEA is that you do it

> by need, and district weren't doing that. They were doing "Everyone gets equal." But everyone does not have equal needs. Why would you do that? I think … Focus status has led to that change…. It's a forced equitable piece, that set-aside. Because you can't distribute it equally when you have set-asides. (Interview, state-level personnel).

> And then we understood that we had to move some of our Title funding to those schools. And that made us really angry because we're taking money away from schools that have the highest amount of need. They're not Focus Schools because all of their students, or a high percentage of their students, are on free and reduced lunch. So, why are we taking money away from those to schools that are doing well, that have average or above-average student achievement just because there's that gap? But we understood, too, that just because a student goes to a school like, for instance, [the high performing school in district], doesn't mean that they're going to be high achievers. So we did understand that we need to definitely take a look at what we're doing for our students that are lower-achieving, that lowest 30% at those schools, to make sure that they're not getting left behind because everyone else is moving forward compared to a school that has a high percentage where the tier one instruction is geared towards students that may be a little bit behind. So that was that initial thought. It was kind of angry, but it's like, "Okay, this gives us an opportunity to take a look at something that we should be monitoring (Interview, district-level personnel).

The shift toward acceptance of Focus status and the belief that it could be used to draw attention to struggling students within more affluent or higher performing schools, on average, was one that was echoed in meetings across governance levels. However, it was felt more acutely at the district level than any other level of governance, as this is where the redistribution of funds occurred:

> I think at the school level, [support] was more mixed or more to the positive, where at the district level, [initially], it was difficult, again, to take money from schools with the greatest need and give it to schools that we didn't feel like needed any more funding or needed it more than those other schools (Interview, district-level personnel).

Thus, at the district level, the Title I set-asides were an important boon to be able to redistribute funds, though this redistribution was controversial. However, at the school level, the identification alone was enough for schools to comply with state-level requirements. More than the monetary supports, it was the technical assistance and capacity building efforts that were perceived to be more efficacious in reducing within-school inequality.

Interactions: Schools, Policy Design, and Environment

Accountability versus Accountability Plus Monetary Support

Another way Title I set-asides affected the policy implemented was that non-Title I schools were required to comply with all of the same regulations as Title I schools, but received no extra monetary support to do so. While both Priority and Focus Schools were eligible for School Improvement Grants (SIGs), these primarily went to Priority Schools. In addition, several Focus School principals reported not wanting to apply for them because they came with additional reporting and requirements that were viewed as unwelcome and burdensome. Consequently, the only monetary incentives that Focus Schools received was the 10% of building fund set-asides that Title I schools received. Non-Title I schools did not receive any extra monetary support but were still required to participate in data dialogues and the superintendent's dropout challenge. However, "we had such surprising lack of blowback and resistance" from non-Title I schools in complying with Focus Status, said one state-level interviewee who put that percentage of non-Title I schools that complied with Focus School requirements as "above 90%." This may have been a result of the accountability part of Focus identification. "It forced [Focus Schools] to do that because now they were being watched" (Interview, ISD-level personnel). Such comments provide further support for the idea that a tenet of accountability design – identification – was a productive one, at least in some cases.

The data shows that while there was initial mistrust in the methodology used to identify Focus Schools, it was accompanied by a desire to get off the list. In part, this was due to the fact that schools that were not traditionally identified under the state accountability system were often identified as Focus Schools. As one participant reported:

> So the [district] administrators have to go report to the board their progress, so there's a little accountability that way. And so I think a lot of districts jumped on that, especially if they had points of pride already and this all of a sudden was a blemish on their district. They worked really hard to try to rectify that. (Interview, ISD-level staff)

However, as we see below, the accountability pressure alone was not sufficient to reduce within-school inequality. Districts that coordinated capacity and developed systems of support were more likely to see a larger number of Focus Schools exit Focus status. This exiting of Focus status was seen as due to a real reduction in within-school inequality in high implementing districts and not just a fluke of measurement. This finding is supported in the literature as research suggests that in complex policy environments, reforms that require changing attitudes, beliefs, or daily routines require a combination of pressure and support for successful implementation (see, e.g., McLaughlin, 1987).

Interactions: Schools, Policy Design, Environment, and Organizations

Coordinated Capacity is a Key Mediator of Implementation

The supports provided for Focus Schools have also changed over time as governmental K-12 actors and non-governmental actors have built capacity to understand the policy, its metrics, and what may be needed to support the bottom 30% of students even in schools that may not be the lowest performing schools on average. For example, at the beginning, one Focus principal expressed the notion that Focus status was not as much of a priority to the central office as Priority Schools, which were the lowest performing schools. This perception was pervasive throughout my study of Focus Schools among participants from all levels of governance. Indeed, the documentation backs this up as ISD support staff were only allowed to dedicate eight hours per month to support Focus districts and schools, whereas they spent at least 40 hours per month with Priority Schools. In part this was because Priority Schools were seen as more in need of

support than Focus Schools, for whom "there was already good stuff happening over there, so it was more a matter of trying to figure out how to do it over here, too" (Interview, district level). As one school-level principal recounted, any support provided to Focus Schools, at least initially, was largely a function of school-level leadership:

> We have a monthly administrator meeting [after which] they pulled aside the principals of Focus Schools after the meeting. And so we had a brief discussion there about what [Focus status] entailed, what we might need to do, and setting up some of the guidelines. But the support is going much more to the Priority Schools than the Focus Schools. So whenever I had questions, there was always somebody there to answer the question, but otherwise I would go through and make sure the documentation and fill in the staff in what we needed to do and get supports in place and collect the data for what we needed to find out. (Interview, school-level personnel)

Over time, however, systemic supports, such as curriculum coaches and leadership training, were put into place according to documentation and interview respondents and were reported to be useful in improving student outcomes. The ability to effect change at a systems level was repeated across governance levels as a key mediator of successful implementation of Focus supports. The educational system, in this case, was the statewide system of supports managed by MIExcel. The program used policy instruments such as professional development and training to disseminate a vision of systemic district support that included strong leadership and data-driven decision making. At the state, ISD, and district levels, this systemic approach was seen as the key to "successful implementation" or getting off the Focus list.

> Let me give you an example of [Y District]. Almost all—above 80%—of their schools were identified as Focus Schools and [at first] it was total "It's not us it's the metric." They wouldn't even let anyone in to their district that first year. It was really hard. So second year, they're still having trouble, MDE is on them, "No you're Title I, you have funding reserved, this is what you have to do." [So], they hired—we hired—a retired curriculum … person just to look at gaps, so she did. She said "I've got no skin in this game." She's very skilled in teaching and learning. She found that they were not … looking at the bottom 30%, that at a local level they weren't looking at achievement ... they were in total denial. And they kept getting re-identified. So she took them to task. She said, "This is a problem. We're not reaching all kids. What are we going to do about

> it?" And she just went right at it … and now all but two of their schools have been released [from Focus status] out of … 30 or 40. It's a lot … [And it came from] curriculum changes, teaching and learning, looking at data, holding teachers accountable, talking about it when they weren't improving." (Interview, state-level personnel)
>
> In [our county] we had 26 Focus Schools identified in 2012. In 2016, we had just six …. Schools that are still Focus Schools are schools that lack really strong systems …. For example in [one district], they started out with seven Focus Schools, but now they only have one. The building that's still a Focus School has gone through three administrators in the past three years … and has [sic] never been able to build a strong system of supports for students. (Interview, ISD-level personnel)
>
> Well, we haven't really put into place anything that's totally unique, or out there. It's, "Here's the basics of best practices. This is what needs to occur. Let's help you put that in place." But yeah, we really don't go outside the bounds. It's really about getting people proficient in the activities that we recommend, or are supported. Like putting the systems in place, and then having those systems work for data use, or progress monitoring, or whatever the case may be. (Interview, district-level personnel)

All respondents reported on the need to coordinate capacity of organizations at various levels of governance to align with a unified vision, preferably set at the district level. As one district-level staff member put it:

> We can't have people coming in and saying, "Do this, do this, do this instead" because people get confused and then you lose the vision. Quite honestly, when I first came here, there were so many different folks coming in. Not only from the—well, mostly from state groups. You had the SIG monitor, you had the SRO person, you had the MIExcel [statewide system of support] people. I had to actually make a chart and say, "Who are all these people?" So the goal has been to say, "Yup, come and help us, but you need to be within our vision. (Interview, district-level personnel)

The idea that supports must be aligned with a district-driven vision is different from prior implementation stories in which a federal initiative was thought to direct local reform efforts. In this case, the federal policy—specifically, the ESEA waivers—defined the parameters of reform, but the local district was charged with creating an education system and vision that could support its policy aim, i.e., the reduction of within-school supports. In high-implementation districts, it would be the district improvement team, comprised of the Title I coordinator, school

improvement director, curricular coaches, and often the district superintendent, that would set that vision. District size was reported to be a mediating factor, as smaller districts often did not have the capacity to develop a system of support and therefore leaned on larger districts or relied more heavily on ISD support to supplement—or even at times, supplant—this function. Another factor was the number of Focus Schools in a given district. Districts with only one Focus School reported offering less structured support than districts with multiple Focus Schools, where building a system was often a more efficient way to approach the problem as it took advantage of economies of scale.

Another change reported by interview respondents and supported by observations and documents is that Focus identification led to a change in the way that ISD supports were deployed to high achievement-gap schools. In the past, ISD support involved deploying disciplinary coaches to classrooms to change teacher practice. However, after the identification of Focus Schools, the MI Excel Program, which is the statewide system of supports for MDE, began to a build a systems-based approach to school improvement, in particular for the reduction of within-school inequality.

> ISDs were already responsible for the classroom-level support, but they were doing it, I guess, removed from what was happening in the rest of the building—one teacher at a time. By moving [Focus supports] to the ISDs now they have both. They have the ability to do the [building] leadership piece if that's what they choose to do, but they also have the wraparound supports with the literacy coaches and math coaches. (Interview, ISD-level personnel)

This is another example of how local vision collaboratively developed by the school and district could dictate how and when ISD supports were deployed to supplement but not supersede reform efforts. This effort represents further evidence that coordinated capacity, not just capacity, was important to achieving the policy aims set forth in ESEA waiver legislation. For example, ISDs

that still operated as "plug-in" curriculum coaches without a coordinated vision were perceived by interview respondents as less successful than those who were deployed in strategic ways. The same size considerations for districts were true of ISDs as well. For example, smaller ISDs were encouraged to partner with larger ISDs who could develop systems of support for schools in order to use their curricula, staffing, expertise, or other developed infrastructure.

Much of the rhetoric surrounding current educational reforms elicits calls for the implementation of a "systemic approach" to these reforms. However, what constitutes or facilitates a systemic approach to reform is still being understood. What I refer to as coordinated capacity in this context—a systemic effort to harness individual and organizational capacity—breaks down the notion of a systemic approach into several characteristics. First, local entities (such as schools and districts) would collaboratively set the vision for local priorities so that state resources, including Title I set-asides, ISD curricular coaches, and professional development opportunities, can be harnessed in service of that vision. Second, school leaders are an important conduit in marshaling those resources to the appropriate staff and students. Third, as we see in the next section, access to data and data use play an important role in democratizing the vision and empowering teachers to enact pedagogical strategies informed by data review.

Data Use Can Be an Instrument to Help Build Coordinated Capacity

A common theme throughout is the importance of data and feedback to improvement. One district staff member said:

> The key piece in every district that needs to move kids is getting that data in the hands of the teachers so that they can use it to target the interventions. So that's been a huge piece. (Interview, district-personnel).

At the school level, there is some evidence that this data-driven culture is becoming pervasive, in part because of the metric of identifying the bottom 30% of students. One principal reported using this metric to create lists for teachers and structure discussion around those students:

> At the point that we became a Focus School, I started taking our data. Our teachers had the classroom data sets that they could pull from what we were using data-wise at the time. But I can pull the grade levels sets and I would—I just started ranking students from top to bottom, or bottom to top. By grade. And we had grade-level meetings monthly so we focused on those students and the interventions. And that became part of the grade-level meetings that became part of our staff meetings. We looked at data. I start most staff meetings or professional development with some data sets that we have, that we've gathered in the last four weeks, that teachers might not have seen. (Interview, school-level personnel)

This is demonstrative of the importance of coordinated capacity. For example, the ESEA waiver set the aim that the bottom 30% of students was to be the locus of improvement for Focus Schools. School leadership was important in coalescing staff attention and efforts on the data to achieve these goals. But a competent and visionary principal could not have individually enacted the teaching and learning strategies needed to address any deficits found in the data. And individual or even collective grade-level teacher cohort capacity would not have been enough to affect changes for the bottom 30% independently of other schoolwide efforts. Indeed, as a consequence of data review, district or ISD resources or expertise was often marshaled to address issues found in the data. In this way, it was not any one of these components but the coordinated interactions between the school, policy design, environment, and organizations that yielded successful efforts to reduce within-school inequality.

Another principal reported training teachers on data usage as teachers may not be used to reviewing data in such a systematic way. This need for training on data management and review was echoed throughout each governance level. Respondents touted a strong faith in data and evidence, though some admitted that in addition to identifying students, it was important to know

what to do after patterns are found in the data—that is, the knowledge of what types of interventions each tier of students should get was not always clear. The importance of collaborative efforts between schools, organizations, and the environment to develop solutions and innovate in these cases became even more pertinent when the solution set was uncertain than when it was more certain. Still, most of the conversation focused on data use and capacity building for data use to achieve policy goals.

> [Teachers are] getting a lot more with the ILCs (Instructional Learning Cycles). They're getting a lot more with heavy data usage. A lot of teachers know kind of where the kids are at, but they don't like doing all the data. They know what the data is that they're putting in, but to track it and then see the progress and see why certain groups are making it and certain groups aren't, I don't think they're used to that. And so that's what we're trying to get—a culture where they can look at the data and be accountable for it for those kids that aren't making [it]. (Interview, school-level personnel)

The focus on data shows that data usage can be an instrument by which the policy goals of reducing the achievement gap may be achieved. Again, however, data usage is dependent on the ability of teachers, principals, and central office administrators to understand and review data together. This is a necessary but insufficient step, as the interventions that are deployed after data use will determine the extent to which struggling learners are able to catch up to their peers.

Another consideration is that even if data is readily accessible and used and the appropriate interventions are deployed, high performing students may take up the intervention provided at a higher or faster rate than their lower performing peers. In this scenario, lower performing students may improve but higher performing students improve at an equal or faster rate, thus maintaining or exacerbating the gap between students. The requirement, then, is to review data and implement interventions in such a way that low performing students show real gains in achievement at a faster rate than their higher performing peers. This was a dilemma that

practitioners at all governance levels struggled with and admitted was a challenge in implementing supports for Focus Schools.

Consequently, data-driven decision making and data-led efforts can be a powerful instrument to achieve policy aims. However, in the absence of coordinated capacity, data access and even data use will not have the magnified effects it could have at scale given the ability to amass and align both individual and organizational capacities to achieve a policy aim.

Study Significance and Implications

Though there are limitations to this study, there are also some interesting insights that arise from examining implementation of supports to Focus Schools. The first is that buy-in for the policy changed as the policy design was amended based on schools that were identified. The interaction between these elements eventually led to changes in the policy design that excluded schools from being identified that were higher achieving than the state average or that were showing improvement for the bottom 30% of students that was equivalent to or above the state average. This change indicates that there was significant learning about how high achievement-gap schools should be identified as a result of the ESEA waivers.

A second insight is that schools were encouraged to develop a multi-tiered system of supports to help improve achievement for the bottom 30% of students. However, many practitioners reported that they did not differentiate their instruction and that most people received only the first tier of support. Consequently, the emphasis on MTSS indicates a need for innovation of interventions for the bottom 30% of students even if the other potential weaknesses of implementation outlined earlier are accounted for.

Third, if coordinated capacity is indeed a mediating factor for successful implementation of supports to improve achievement for the bottom 30% of students, then the question becomes

how this kind of capacity can be developed within the current infrastructure of schooling. Chapter 4 begins to tackle this question by exploring the extent to which a Networked Improvement Community can be used to develop coordinated capacity within and between individuals and organizations that share a common aim.

This study has limitations that are important to keep in mind when interpreting results. First, the sample of respondents is small and may not be representative of Focus Schools or districts in Michigan. Instead, the sample represents a vertical cross-section of practitioners across two mid-sized urban districts in Michigan. Some schools and districts that were initially selected for representativeness by size, urbanicity, and Title I status opted out of participation. Consequently, the remaining sample of respondents may present a biased viewpoint of Focus School implementation. For example, if only those individuals who bought into Focus School status agreed to participate, then this study does not capture dissenting voices except for the dissent publicly voiced during statewide meetings and captured in field notes. While the use of observations and documentation to triangulate interview responses minimizes this bias, it does not eliminate it completely.

A second limitation is that the interviews and observations were conducted during the 2015-16 school year, but interview respondents were asked to reflect on the initial years of Focus identification. However, as the supports provided to Focus Schools and the criteria for exiting Focus status evolved over the four years of Focus identification, respondents may have recounted or remembered perceptions from these later periods that may or may not have been true in the first year of identification and the first two years of implementation of supports. Again, to minimize this bias, I corroborated respondent interviews using documentation provided by the Michigan Department of Education and MI Excel, the program manager for the Michigan

statewide system of supports. Despite these limitations, this study offers a unifying vision of important factors for the implementation of a statewide system of supports to reduce within-school inequality, including coordinated capacity among and between governance level and the use of data as a policy instrument to achieve policy aims.

References

Boyle, A., Le Floch, K., & Bowles Therriault, S. (2008). *State systems of support under NCLB: Design components and quality considerations.* Washington, DC: American Institutes for Research.

Cohen, D.K., Peurach, D.J., Glazer, J.L., Gates, K., & Goldin, S. (2014). *Improvement by design: The promise of better schools*. Chicago, IL: The University of Chicago Press.

Cohen, D. K., & Moffitt, S. L. (2009). *The ordeal of equality: Did federal regulation fix the schools?* Cambridge, MA: Harvard University Press.

Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic fieldnotes.* Chicago, IL: University of Chicago Press.

Honig, M.I. (2003). Building policy from practice: District central office administrators' roles and capacity for implementing collaborative education policy. *Educational Administration Quarterly, 39*(3), 292–338.

Joachim, A., & Murphy, P. J. (2013, December). The capacity challenge: What it takes for state education agencies to support school improvement. Seattle, WA: Center on Reinventing Public Education.

Kober, N., & Rentner, D. S. (2012). *Year two of implementing the Common Core State Standards: States' progress and challenges.* Washington, DC: Center on Education Policy.

Marsh, J., Kerr, K., Ikemoto, G., Darilek, H., Suttorp, M.J., Zimmer, R., & Barney, H. (2005). *The role of districts in fostering instructional improvement: Lessons from three urban districts partnered with the Institute for Learning*, MG-361-WFHF, Santa Monica, Calif.: RAND Corporation. Retrieved from http://www.rand.org/pubs/monographs/MG361/

McLaughlin, M. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis, 9*(2), 171–178.

Mintrop, H., & Trujillo, T. (2007). *School improvement under test driven accountability: A comparison of high-and-low-performing middle schools in California.* CSE Report 717. National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Murphy, P. J., & Rainey, L. (2012). *Modernizing the state education agencies: Different paths toward performance management.* Seattle, WA: Center on Reinventing Public Education.

Oliff, P., & Leachman, M. (2011). New school year brings steep cuts in state funding for schools. Washington, D.C.: Center on Budget and Policy Priorities. Retrieved from http://www.cbpp.org/research/new-school-year-brings-steep-cuts-in-state-funding-for-schools

Pressman, J.L., & Wildavsky, A. (1973). *Implementation.* Berkeley, CA: University of California Press.

Rhim, L. M., & Redding, S. (Eds.). (2014). *The state role in school turnaround: Emerging best practices.* San Francisco, CA: WestEd. Retrieved from http://centeronschoolturnaround.org/wp-content/uploads/2014/01/The_State_Role_in_School_Turnaround.pdf

Weiss, J., & McGuinn, P. (2016). States as change agents under ESSA. *Phi Delta Kappan 97*(8): 28–33.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Zavadsky, H. (2012). *School turnarounds: The essential role of districts.* Cambridge, MA: Harvard Education Press.

Chapter 4

The Promise of Continuous Improvement: Developing a Networked Improvement Community to Support Focus Schools at the District-Level

"We live stuck between two polar views. On the one hand, a robust infrastructure has emerged for examining narrow, focused propositions through large, randomized field trials. On the other hand, there is a long tradition in education of local learning from the actions of individual practitioners. We … argue for a third way." —Bryk, Gomez, & Grunow (2011)

"Rather than asking whether an 'intervention works,' a network improvement community asks, 'what works, when, for whom and under what sets of circumstances?'" —Bryk, Gomez, & Grunow (2011)

Abstract

Although not new, over the past decade collaborative research partnerships between researchers and schools, school districts, and community-based organizations increasingly have been seen as productive approaches to addressing persistent problems of practice in education (Coburn, Penuel, & Geil, 2013). One such type of research-practice partnership is a Networked Improvement Community (NIC), defined as "a distinct network form that arranges human and technical resources so that the community is capable of getting better at getting better" (Englebart, 2003 in Bryk, Gomez & Grunow, 2011). This networked approach could potentially serve as a design feature to develop the coordinated capacity discussed in Chapter 3 that is central to implementing complex reforms in institutions with historically weak capacity; however, there are scant cases that describe the implementation of an NIC approach in a real-world context. To examine the conditions that facilitate and constrain the development of such a

network, I use the case of the Michigan Focus NIC, initiated and sustained over the course of the 2015-16 school year to reduce achievement gaps in Michigan Focus Schools. As a result of my role as leader of the network hub, I am able to provide a unique perspective on the elements of social organization needed to initiate and sustain a Networked Improvement Community including 1) identifying organizations and individuals for membership; 2) building legitimacy; 3) establishing norms; and 4) building trust, buy-in, and ownership. I also discuss implications on the use of such a methodology for alleviating within-school achievement gaps at scale.

Introduction

Although not new, over the past decade collaborative research partnerships between researchers and schools, school districts, and community-based organizations increasingly have been seen as productive approaches to addressing persistent problems of practice in education (Coburn, Penuel, & Geil, 2013). In part, the increased focused on research-practice partnerships is a response to 1) policy design of educational reforms in the past 15 years and 2) the impact evaluations of these policies. Many of the educational reforms enacted in the U.S. since the 1990s have focused on building an exoskeleton of standards and assessments to monitor and (dis)incentivize schools to act in politically desirable ways. In this design, the daily routines and strategies to accomplish those goals were often left to practitioners with historically weak capacity (Cohen & Moffitt, 2009). Over a decade of rigorous impact evaluations of these efforts have indicated null or small effects for such educational interventions (Jacob, 2015; Dynarski, 2016). These converging trends belie the need for the development of what Peurach (2016) calls "improvement infrastructure" analogous to the impact infrastructure developed to fund and conduct rigorous evaluations of educational interventions.

This study describes one such effort to develop "improvement infrastructure" through a collaborative effort between the Regional Educational Laboratory (REL) Midwest and the Michigan Department of Education funded by the Institutes for Education Sciences. In this effort, these partners initiated and sustained a Networked Improvement Community (NIC), a type of research-practice partnership, in Michigan during the 2015-16 school years to support Focus Schools. This NIC, referred to as the Michigan Focus NIC, is comprised of Focus School principals, district administrators, intermediate school district (ISD) personnel, state department of education personnel, and REL Midwest researchers. Together, the Michigan Focus NIC worked together to identify a problem of practice, develop an intervention, and study its effects. This chapter describes the efforts of the network hub, REL Midwest, to attend to the elements of social organization needed to initiate and sustain the NIC and provides a narrative of NIC activities to illustrate how these elements played out throughout each phase of NIC work.

The Promise of Networked Improvement Communities

Networked Improvement Communities, or NICs, are comprised of individuals or organizations from diverse contexts who are focused on a shared problem of practice and use systematic, scientific inquiry to develop and test innovations at scale. NICs are focused on improving practice in a particular industry through defining a data problem, developing a theory of practice improvement, and conducting inquiry cycles to test an intervention through Plan-Do-Study-Act (PDSA) cycles. In a framing popularized by the Carnegie Foundation for the Advancement of Teaching, improvement science is concerned not only with "What works?" but rather "What works, for whom, and in what context?" This approach seeks to capitalize on contextual variation to learn what works in diverse contexts faster in lieu of seeing contextual variation as an obstacle to program implementation.

NIC participants can include educators working in a range of capacities and at different levels of the education system (for example, the school, district, or state level). One example is the Carnegie Foundation's Building a Teacher Effectiveness Network (Bryk, Gomez, Grunow, & LeMahieu., 2015). This NIC is facilitated by the Carnegie Foundation, the Institute for Healthcare Improvement, and the American Federation of Teachers. Participants include principals and teachers from the Austin Independent School District, Baltimore City Schools, and the charter network New Visions for Public Schools. Another example is the Association of Public and Land-Grant Universities' Mathematics Teacher Education Partnership (Martin & Gobstein, 2015). This NIC is facilitated by the Association of Public and Land-Grant Universities and an unnamed member university. Participants include leaders from over 90 member universities and over 100 K-12 school districts. NICs are characterized by four salient features (Bryk et al., 2015):

1. They are focused on a well-specified, common aim.
2. They are guided by a shared working theory of the system and how it can be improved.
3. They utilize improvement science methods and measures to spur improvement in testable iterations, such as rapid Plan-Do-Study-Act (PDSA) cycles. (See the glossary in box 1 for an explanation of PDSA cycles.)
4. They are organized to share and integrate practices and processes developed within the NIC to other contexts.

The Carnegie Foundation provides guidance for researchers and educators who intend to form NICs using several examples from their work, including forming NICs around community college graduation and improving supports for novice teachers (Bryk et al., 2015). While these

existence proofs are important, they are often efforts to enact NICs in near-ideal circumstances with leading industry experts in research design and measurement, high capacity practitioners as partners, and functional educational environments in which to operate. Outside of these examples, researchers and educators have few additional examples upon which to draw guidance. In particular, the guidance provided by the Carnegie Foundation focuses on conducting continuous improvement research within the NIC once it has been established rather than the critical process of establishing NICs (Russell et al., 2016). Moreover, existing literature is limited in addressing the social aspects of establishing an NIC, including the dynamics of the people and groups participating in the NIC and the way they negotiate their roles and responsibilities (Patton, 2011). This study adds to this literature by providing a case study of what initiating and sustaining an NIC in a real-world context—that is, with little extra monetary or personnel investment, no extra time or resources set aside for its work, and voluntary efforts on the part of practitioners—may look like.

Conceptual Framework

Forming an NIC requires 1) fostering a community with shared culture, norms, and identity organized around networked aims and 2) the work of acting as a "hub" that leads, organizes, and operates the network. Once an NIC community is defined and developed, an NIC's work is defined by three primary tasks: develop a theory of practice improvement; use continuous improvement research methods (such as PDSA cycles) to implement, test, and redesign an intervention in an iterative manner; and build a measurement and analytics infrastructure (Russell et al., 2016). Figure 8 from Russell et al. (2016) provides a framework for understanding these five components of initiating and sustaining an NIC and presents a theoretical framework for understanding how the Michigan Focus NIC was initiated and

operated during the 2015-16 school year. To discuss the initiation of the Michigan Focus NIC, I collapse the domains of "leading, organizing, and operating the network" and "fostering the emergence of culture, norms, and identity" to one domain, "attending to the social organization of forming an NIC community." Within this domain, I consider four subcategories of activities: membership; building legitimacy; establishing norms; and building buy-in, ownership, and trust.

Figure 8. Framework for Initiating Networked Improvement Communities
from Russel, et al., 2016

Although educators are the primary participants in this process, a "network hub"—often composed of researchers—facilitates the process (Bryk et al., 2015). The network hub provides expertise on continuous improvement research, guiding the participants through each step of the process. The network hub also serves an observational role, monitoring the process of continuous

improvement research across sites and providing feedback to participants about what is working well and where there may be challenges. Finally, the network hub plays the practical role of convening participants. As Bryk et al. (2105) conclude, "tending to the needs of the community is foundational for everything else" (p. 159). Here, the network hub is referred to as the network initiation team and its activities are embedded into the work of attending to the social organization of forming an NIC community and attending to the core technology of NIC work. The next section describes considerations for establishing an NIC and then delineates the three primary tasks that an NIC undertakes.

Attending to the Social Organization of Forming an NIC Community

The importance of attending to the social arrangements of partnerships is well documented in the literature (Coburn, Penuel, & Geil, 2013). One way research partnerships can support efficacious partnerships is in the identification of key organizations and individuals. Partners or stakeholders may be identified for a variety of reasons, including power over or proximity to the locus of change; contextual, content or methodological knowledge needed to support the partnership; or ability to build and sustain clout for the partnership, among others. Another component of the social organization of collaborative research partnerships is the ability of partnerships to build legitimacy, both between partners and between the partnership and the external context. One difficulty in building legitimacy in collaborative research partnerships may stem from the differing perspectives and goals of research and practitioners (Firestone & Fisler, 2002; Coburn, Penuel & Geil, 2013). However, doing so is a prerequisite for building a collaborative partnership in which partners work effectively together.

Once key partners are identified who respect one another's realm of expertise, collaborative partnerships must establish norms for working together including the frequency,

duration, content, tone, and mode of communication; the method by which the group will come to consensus; and expectations for ownership and dissemination of products that emerge as a result of the partnership, among others. Finally, perhaps the most important component of social organizations is trust between members and between the partnership and the external environment. Building buy-in, ownership, and trust have been shown to be key characteristics of both successfully implemented educational reforms as well as productive collaborative research partnerships (Bryk & Schneider, 2002; Vangen & Huxham, 2003). These components of the social organization of partnerships—membership, legitimacy, norms, buy-in, ownership, and trust—evolve over time and are key considerations when forming an NIC (Bhatt & Proger, forthcoming; Russel, et al., 2016).

Identifying organizations and individuals who are involved in that problem space within the context. First, a hub leader or network must identify partners and participants who are involved in the problem space. The universe of potential partners and participants may include those with a specific geographic location, organization type, expertise, or position. Initiating an NIC may mean partners often choose individuals or organizations who are geographically proximal, such as the Houston Education Research Consortium (HERC), a collaborative effort between researchers at Rice University and the Houston Independent School District (HISD). Another model is to cast a wide net to recruit partners and then work with those individuals and organizations who express interest. For example, in recruiting participants for a Networked Improvement Community in Michigan, researchers used a snowball method to recruit participants at the state, intermediate school district, district, and school levels respectively. Oftentimes, the identification of partners in the planning phase is ad hoc and based on prior relationships. This method of identification provides an initial sample of potential partners but

does not guarantee that partnerships present the ideal configuration of expertise and knowledge. As the work of the NIC matures, partnerships become more fluid. There is a core set of partners that is committed to the project or research agenda, but initiators of the collaborative relationship are still exploring additional partnerships that can be used to build legitimacy within a specific context or content area.

Building legitimacy. Diverse perspectives can make it difficult for partnerships to build internal legitimacy, or legitimacy between partners. For example, while researchers may prize formal credentials and academic publications as a source of legitimate knowledge, practitioners may instead prioritize experiential and contextual knowledge. Partnerships in which all entities are viewed as legitimate must prize and prioritize each of these types of knowledge. In addition, the partnership should be sanctioned by local entities in such a way that it is viewed as an expert source of information on the content area of interest.

At first, NIC hub leaders should focus on building and maintaining internal legitimacy between partners. In order to build legitimacy between partners in the planning phase, initiating partners often choose individuals who are geographically proximal, as is the case with the Michigan Consortium for Education Research (MCER). In this partnership, researchers from the University of Michigan and Michigan State University have partnered with the Michigan Department of Education (MDE) and its accountability arm, Center for Educational Performance Information (CEPI). The focus of this partnership is evaluation of Michigan policies for K-12 education; consequently, partnering with Michigan institutions of higher education was a vital part of legitimizing the partnership. Doing so also allows meetings to occur in person, which can be beneficial in establishing relationships across institutions. In the case of MCER, using

geographically proximal research partners also aided the partnership in gaining external legitimacy (Conaway, Keesler, & Schwartz, 2015).

Another way to build legitimacy between partners is through prior work. In Denver, the American Institutes for Research (AIR) has been involved in conducting an evaluation of Denver's teacher residency program. As a consequence of that work, the district is exploring initiating a research practice partnership around teacher residency. In this case, prior working relationships between the two organizations proved to be a way that the partners built mutual respect for one another, which is essential to the social organization of a partnership.

Partners must work together beyond respect for expertise and knowledge and begin to see the partnership itself as a legitimate enterprise. Part of this legitimacy stems from follow through on stated activities and consistent, frequent contact. This consistency signals the partnership as a priority to the other partners, thereby legitimizing its work as a worthwhile endeavor. As a partnership matures, it must turn an eye towards building external legitimacy—that is, legitimacy of the partnership in the external context. This may mean creating a dissemination strategy that is tailored for a specific population, receiving endorsements from legitimate stakeholders in the field, or transferring an intervention or strategy to another context. Presenting at conferences and local stakeholder meetings, publishing articles or editorials, and producing multiple sources of data may all be part of this strategy. For example, in design-based research, researchers are primarily concerned with implementation of interventions and contribution to theory. However, after several years of work, researchers attempt to collect and analyze standardized data points such as students' standardized test scores across contexts in order to build legitimacy in the research arena of the design-based research approach. For example, in the Mathematics in the Institutional Setting of Teaching (MIST) project, researchers worked in four separate districts but

reported results on standardized mathematics test scores because it is a measure of interest in the current educational reform environment (Cobb & Jackson, 2012).

Establishing norms. Any emerging partnership in the planning phase must establish norms for how communication between individuals and organizations will occur. Establishing these expectations early in the planning stage reduces the probability of hurt feelings or misinterpreted actions that can undermine efforts to build trust within the partnership. These norms should encompass the frequency, duration, and mode of communication. In addition, the planning phase sets the tone with which communication occurs, which can have lasting implications for later partnerships. For example, researchers at the Houston Education Research Consortium had to negotiate norms for providing and receiving criticism in academic settings versus school districts early in the partnership to avoid missteps (Turley & Stevens, 2015).

Research-practice partnerships must develop norms for decision making, particularly in dynamic contexts where the lack of norms could lead to conflict. For example, early in the organizational history of the CCSR, researchers instituted a "no-surprises" policy that allows senior leadership and pertinent stakeholders time to review reports prior to their release. This policy can help normalize relationships between practitioners and policymakers when tenuous or controversial findings are released (Roderick, Easton, & Sebring, 2009).

As partnerships mature and become sustainable, these informal, shared norms often must be codified due to growth in membership, expansion of the scope of the collaborative partnership, or in some cases, as a result of or in anticipation of turnover. The partnership must also decide how proprietary content and knowledge generated as a result of its activities will be shared or disseminated. Though these conversations may have happened in the emerging or planning phases, there can often be unanticipated social, legal, and business challenges that must

be addressed in later stages of partnerships. For example, a collaborative effort under the Information Infrastructure Systems (IIS) project brought together an interdisciplinary center housed within the University of Chicago's Center for Urban School Improvement, a private company called Teachscape, and individuals associated with a whole-school reform program known as the Literacy Collaborative. As the interorganizational partnership progressed to the third "gamma" or dissemination phase of development, partners had to deal with competing interests between Teachscape, who wished to commercially disseminate the tool developed by the partnership (PDS2) for royalties, and universities, who in principle owned the content. The complexities of developing such an agreement were not insurmountable but threatened the stability of the partnership and therefore would have benefited from prior established norms and expectations (Rosen, 2010).

Building trust, buy-in, and ownership. The importance of building trust in educational research partnerships is well-documented in the literature (Bryk & Schneider, 2002; Coburn & Stein, 2010; Roderick, Easton, & Sebring, 2010; Coburn, Penual, & Geil, 2013; Turley & Stevens, 2015). Webb (1991) documents that some level of trust is required to initiate collaboration. Paradoxically, however, Creed and Miles (1996) document that "trust begets trust." Clifford and Millar (2010) document that trust in P-20 partnerships is an input required for partnership, a process by which partnership activities occur, and an outcome of the partnership. This quality of trust as both an input and outcome of collaborative relationships suggests that building trust happens through a cyclical, iterative process (Vangen & Huxham, 2003). At the beginning of initiating an NIC, trust can be built through "small-wins," or mutually beneficial experiences achieved through the implementation of low-risk initiatives (Bryson, 1988). When trusting attitudes are reinforced through these initially modest outcomes, there is a greater basis

for increasingly more ambitious collaboration in mature and sustainable phases of collaborative partnerships (Vangen & Huxham, 2003). For example, in the MIST project, researchers articulate that their repeated insistence and demonstration that the relevance of research to district priorities was of utmost importance helped the team to build trust over time (Rosenquist, Henrick, & Smith, 2015).

Trust is a component of initiating and sustaining an NIC that is paramount to building relationships. However, partnerships require initial buy-in to the premise of the partnership and the ideas it puts forth. This can be established through devotion of time to the recruitment and onboarding process (Coburn, Penuel, & Geil, 2013; Bhatt et al., 2015, forthcoming). In addition, partnerships can institute quick turnaround projects, as in the case of the Baltimore Education Research Consortium (BERC), which reported using Rapid Response projects to build buy-in and bide time when waiting for results from long-term projects that might have led to publications for the researchers who were part of the partnership (Connolly, Plank, & Rone, 2012). It is important for partners to be committed to mutualism through the life stages of the partnership—that is, the idea that the partnership should be beneficial to both practitioners and researchers (Coburn, Penuel, Geil, 2013). However, mutualism should evolve from buy-in from internal partners during the initial stages to ownership of the work of the partnership in mature and sustainable phases of the work. This is largely because building and maintaining buy-in is a process that requires time and effort. Once an NIC matures, individuals within the network have collegial working relationships, high levels of knowledge of each other's contexts and roles, and a shared identity as part of the partnership.

Conducting the Work of an NIC: Operationalizing Core Technology

Concurrent to establishing an NIC with attention and care devoted to membership, building legitimacy, culture, norms, and trust, NIC participants must work together to conduct the work of practice improvement (Russell, et al., 2016). To do so, NICs engage in its core technology comprised of three interrelated "domains of activity" (Russell, et al., 2016). The three primary tasks of an NIC are to develop a theory of practice improvement, use improvement research methods to learn and improve, and build a measurement and analytics infrastructure.

An NIC's first task is to develop a theory of practice improvement, which involves specifying a problem to focus on, determining the factors that drive that problem by engaging in a root-cause analysis, formulating an aim statement, and generating hypotheses about the changes in practice that will drive improvement in order to meet the aim. Bryk et al. (2015) recommend using a fishbone diagram—a tool adapted from the business sector—to guide participants in visually representing the problem, its root causes, and the factors that contribute to each root cause (Berwick, 2008; Bryk et al., 2015). For example, in the case of the Community College Pathways NIC described by Bryk et al. (2015), participants identified low success rates in developmental math as the problem. They then identified several root causes, such as that instructors lacked the skills and beliefs that students can succeed and that institutional structures did not support student success (Bryk et al., 2015). After the root causes are identified, the NIC selects one or two root causes to focus on and develops an aim statement, or goal, that aligns with the root causes selected. For example, participants in the Community College Pathways NIC focused on improving teaching and learning in community college classrooms and agreed upon the following aim: to increase the proportion of students who complete college math credits within one year of continuous enrollment from five to 50 percent (Bryk et al., 2015). Once the

aim is established, the NIC then hypothesizes what changes in practice would drive the improvement needed to achieve their aim.

The second task of the NIC is to use improvement science research methods, such as Plan-Do-Study-Act (PDSA) cycles, to implement a change in practice based on the theory of practice improvement (Bryk et al., 2015). All NIC participants may implement the same change or may choose to implement variations of the change identified in their respective contexts. Next, the participants implement the change and observe the process, collecting formative evidence of its success. In the analysis stage, participants examine the evidence and determine what tweaks need to be made for the change in practice to be more successful. Finally, they implement the modified practice and test it again, often at additional sites in order to learn from variation across these sites. The scale of the change is small initially—for example, changing the way teachers respond to student questions in a particular subject—but then grows in scale as the NIC continues its work. Cycles may be as short as two weeks or as long as three months.

The third and final task of the NIC is to build a measurement and analytics infrastructure to determine whether the change in practice led to improvement. At first, measurement may be informal, based on the observations of participants and focused on "process-oriented measures" (that is, measures related to how well the change has been implemented). But as the change in practice is scaled to additional sites, the need arises for a common measurement system that involves the collection of data across sites, with the goal of being able to examine whether the change in practice led to a change in the primary driver and the ultimate aim (Bryk et al., 2015). Bryk et al. (2015) emphasize the importance of "practical measurement"—that is, measures that are directly related to the problem identified by the NIC (rather than more traditional accountability measures), and measurement should occur often and be embedded within regular

organizational routines. As participants engaged in this process, they developed an understanding of the NIC process and how it could be used to scale other initiatives in their respective work sites.

The Case of Initiating and Sustaining the Michigan Focus NIC

The ideas set forth in this essay provide guidance on how to establish an NIC to scale improvement in educational practice. However, as this approach is relatively new to the field of education and requires practitioners and researchers to work in ways that differ from historical models (Russel, et al., 2016), case studies of how these ideas are implemented in practice are useful to demonstrate the potential power and limitations of this approach. This case study examines the initiation of one such example, the Michigan Focus NIC, from the perspective of the network initiation team leader. Using the "Framework for Initiating Networked Improvement Communities" (Russell, et al., 2016), I describe the formation of the Michigan Focus NIC, leadership activities, and how the three interrelated domains of NIC core technology played out. This case demonstrates how an NIC may be used to achieve the policy goals set forth in an accountability framework, such as closing achievement gaps, and the considerations of network initiators and participants in doing so.

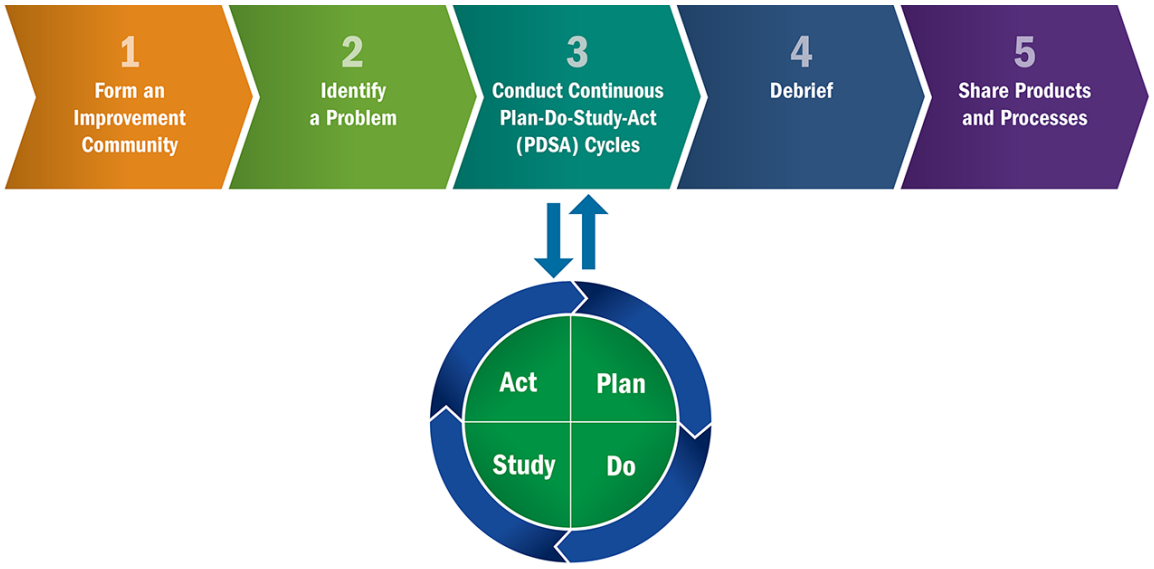
The Social Organization of the Michigan Focus NIC

Focus schools in Michigan were identified in summer 2012 using 2011-12 data, along with Priority Schools and Reward Schools under the Michigan Department of Education's (MDE) ESEA waiver (see the introductory chapter for a more in-depth explanation of ESEA waivers). In spring 2015, MDE approached the Regional Education Laboratory (REL) Midwest and discussed the desire to scale best practices from schools that were "beating the odds" to schools that were in need of improvement under the accountability framework. In these

discussions, MDE mentioned that while the needs of Priority Schools (the lowest performing schools in the state) were known—for example, Priority Schools needed stronger leadership, an improved school climate, and more support for teacher learning—the needs for Focus Schools in terms of how to reduce achievement gaps within their schools was less clear (see Chapter 3 for a more in-depth discussion on this point). One of the consequences of not knowing exactly what was needed to support the bottom 30% of students, which comprised a variety of students including English learners, racial and ethnic minorities, economically disadvantaged students, and special education students, was that the core technology of the intervention discussed in Chapter 2 needed to be invented. REL Midwest researchers suggested that this project may be a good fit for using an NIC approach with improvement science methods, and MDE agreed. Consequently, a research team with REL Midwest developed and instituted an NIC that used improvement science methods to refine supports provided to Focus Schools in the state. This NIC was supported in partnership with the Michigan Department of Education and the Regional Education Laboratory (REL) Midwest at the American Institutes for Research. Between April - September 2015, the network initiation team, consisting of REL Midwest researchers and MDE personnel, planned for the composition and purpose of the NIC and worked together to identify, recruit, and onboard participants. I led these efforts and helped design the structure of the Focus NIC and also implemented it as a hub leader. Given this unique position as designer and implementer, this account cannot be construed as a formal evaluation of efforts to establish and initiate the Michigan Focus NIC. However, as a researcher in the process, I can offer my perspective on the conditions that were helpful in both initiating the NIC and those that would be needed to sustain its efforts. The following sections describe how the network initiation team, which I led, conducted each of the steps outlined above to initiate an NIC, leading to the

establishment of the Michigan Focus NIC. Figure 9 summarizes the steps the network initiation team took to form a network and conduct the core work of the NIC.

Figure 9. Improvement Communities in Action



Michigan Focus Networked Improvement Community

1 **Form an Improvement Community**—In Michigan, Regional Educational Laboratory (REL) Midwest researchers were strategic about recruiting networked improvement community (NIC) participants and devoted time and resources to the onboarding process. Researchers worked with partners at the Michigan Department of Education to identify potential participants at the intermediate school district (ISD) level, ISD partners identified participants at the district level, and so on. Researchers had in-depth conversations with all potential participants to explain the goals and structure of the NIC to create commonalities in language and expectation.

2 **Identify a Problem**—Members of the Michigan Focus NIC determined that a recurring issue was: *Focus Schools suffer from a lack of data use to implement, monitor, and evaluate continuous improvement on a daily basis because of time, priority, skills, resources, and commitment.* This initial problem statement was refined through a root-cause analysis that included using continuous improvement tools such as fishbone diagrams. NIC participants identified mathematics, specifically students' lack of mathematics fluency skills, as a primary driver of inequality in Focus Schools. Participants identified students' lack of opportunity to practice mathematics fluency skills as a problem that the group would work on together.

3 **Conduct Continuous Plan-Do-Study-Act (PDSA) Cycles**—The Michigan NIC then refined the problem statement using the findings uncovered from the root-cause analysis and problem identification step (Step 2) to the following: *Students lack the opportunity to practice mathematics fluency skills on a daily basis, which results in gaps in mathematics fluency skills and exacerbates achievement gaps in mathematics on standardized assessments.* With this new problem statement, the NIC began working through PDSA cycles, with the defined goal that *all students will master grade-level fluency benchmarks by demonstrating appropriate strategies and recalling facts.*

4 **Debrief**—The improvement community meets on a monthly basis to review implementation and data. After the first PDSA cycle, participants will discuss the strengths and weaknesses of the intervention and outcome measures in PDSA Cycle 1 and amend accordingly for Cycle 2, and so on.

5 **Share Products and Processes**—The next step will be to share the products and processes from the Michigan Focus NIC with colleagues within the state and other states and regions to encourage implementation of a similar process and garner best practices from Michigan's experience.

6209_05/16

Membership. While the final configuration of the Michigan Focus NIC consisted of four schools nested in two districts within two intermediate school districts (ISD) within one state, the

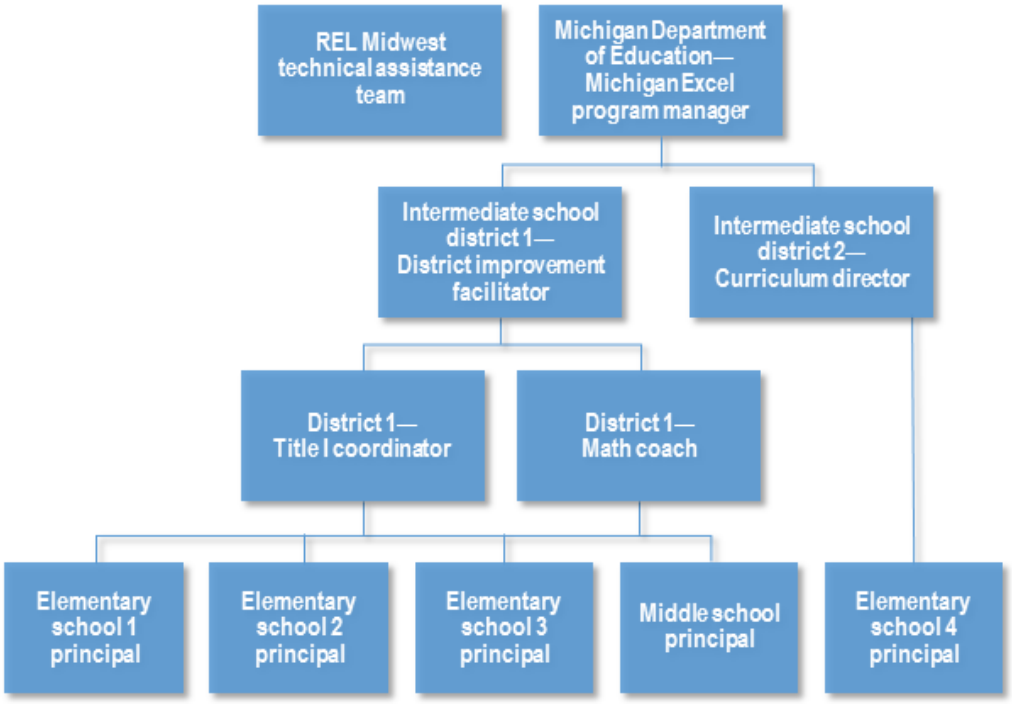
initial configuration planned to include 16 schools across seven districts within three ISDs. The difference in this membership can be attributed to local capacity to implement new ideas, competing interests and initiatives, unfamiliarity with the NIC approach, and distrust of outside or state-level interventions. The network initiation team first identified a series of schools, districts, and ISDs to participate based on the following criteria:

- Title I status
- School size
- Urbanicity
- Number of Focus Schools identified within district
- Number of years school has been identified as Focus School

The initial sample of 16 schools identified using these criteria represented a mix of Title I and non-Title I schools; large and small schools; rural, urban, and suburban schools; districts with varying numbers of Focus Schools; and schools that have been identified in one year only, in two years, and/or in three consecutive years. This variation in contexts and dosage was purposeful per the NIC design in the hope that it would provide useful variation for studying how the intervention worked. The network initiation team then employed a stepwise recruitment strategy in which MDE personnel helped to identify and recruit ISD personnel, who in turn helped to recruit district personnel, who helped to recruit school-level personnel. The network initiation team also made presentations to all ISD-level personnel who supported Focus Schools to recruit any interested participants outside of the initially identified list of districts and ISDs. Through these parallel strategies, the final membership of the Focus NIC consisted of 15 individuals across four governance levels. In addition, teachers in the represented Focus Schools worked to implement the intervention and monitor progress, though they did not attend the monthly NIC meetings due to a shortage of substitute teachers in both districts.

The final sample of Focus NIC participants gave the improvement community several advantages. First, by concentrating on two ISDs that are geographically proximal to one another instead of three disparate ones, the network initiation team was able to conduct in-person meetings. These in-person meetings were vital to creating a sense of community among participants across governance levels and contexts. Second, principals of Focus Schools initially were not counted among NIC participants. Although the network initiation team had felt it important to include school-level voices, it seemed improbable that building principals would be able to commit to a monthly meeting apart from their existing duties. Consequently, these school-level voices would be represented through interviews with principals and teachers. However, after a series of recruitment meetings, ISD and district representatives in the two mid-sized urban contexts felt it was important to include the building principals for several reasons: (1) to represent their school experiences with Focus supports; (2) to identify a change agent responsible for implementing the intervention; and (3) to model how an NIC can be used to drive improvement through tools and activities. The disadvantage of the final list of NIC participants was that the group lost some intentional variation with respect to urbanicity as both areas are relatively midsized urban areas with outlying rural areas. Consequently, a large urban district and small rural district are not represented. Given that an improvement community's primary goal is not generalizability but rather adaptation of interventions to local context, the network initiation team felt that this tradeoff was warranted. The final organization of the Michigan Focus NIC is presented in Figure 10.

Figure 10. The Organization of the Michigan Focus NIC



Establishing Norms. After potential NIC participants were identified, the network initiation team met with them to explain the goals and intended outcomes of the NIC. In Michigan, these early conversations illuminated a need to provide specific information about what an NIC is and how it differed from but built upon other continuous improvement efforts participants already were undertaking, what participation in an NIC involved, and expectations for participating. For example, some participants came into the process with preconceptions about what an NIC is and does based on past participation in communities of practice and professional learning communities. Communities of practice are groups of people that share a common purpose and learn together to improve practice (Wenger & Wenger-Trayner, 2015). While NICs are similar to communities of practice in that members share a common purpose and learn collaboratively, they differ in that they utilize a structured process to identify a problem to

address, decide on and implement an intervention, and measure the outcome of the intervention. Similarly, NICs share some features of professional learning communities—specifically, collaborative learning to develop a shared understanding of a common topic (Hord & Sommers, 2008)—but focus on the impact of a common intervention introduced across sites rather than improving individual practice. Clarifying the differences between an NIC and other collaborative processes provided a common understanding of responsibilities of participants and how the expertise of each member would be utilized to support the goals of the NIC.

To facilitate clarity of what participant expectations were for the NIC, the network initiation team developed a one-page handout of frequently asked questions and a list of participant responsibilities. As participants of the Michigan Focus NIC, members were expected to:

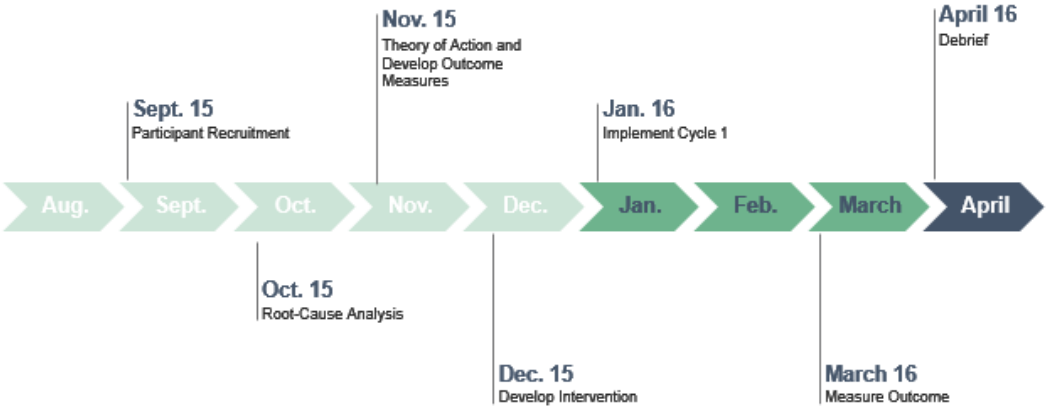
- Participate in regular meetings to identify specific problems.
- Assist in developing a solution to each problem.
- Take the proposed solution back to the local context for implementation and collect data about the process.
- Provide information about how the change has been implemented.
- Review data compiled by the network initiation team and discuss its significance.
- Participate in discussions to refine the change as it is implemented.

In addition, the network initiation team developed a monthly meeting calendar with fixed dates and times to conduct the work of the NIC. This monthly meeting calendar planned NIC activities throughout the 2015-16 school year. One of the norms of the group was that given the complexities of coordinating schedules, these dates would remain fixed and members would commit to meeting in person at a predetermined location midway between the two geographic

locations. During spring testing, when principals needed to be physically present in their buildings during meeting times, the group decided to conduct virtual meetings.

As the group coalesced over time, the need to establish norms not yet addressed arose. These needs were addressed through open discussion with group members at monthly meetings. For example, as external individuals and organizations began to inquire about the work of the NIC, researchers had to ask about the level of anonymity desired by participants and the norms around data sharing both within the group and outside of the group. Figure 11 shows the timeline of Michigan Focus NIC events through April 2016.

Figure 11. Timeline of Michigan Focus NIC Activities Through PDSA Cycle 1



Building legitimacy. The network initiation team primarily used a multi-step recruitment process with known individuals and organizations recruiting the next tier of participants to build legitimacy. In addition, the network initiation team ensured that multiple types of expertise were represented and that the norms of the group did not overvalue one type of expertise over the other. Finally, a third strategy the network initiation team used to build legitimacy was to align NIC work with ongoing efforts throughout the state.

While the REL Midwest technical assistance team served as the facilitator and hub for each NIC, it was critical to do so in partnership with a champion at the state education agency. Champions are usually decision makers in the organization and have the power to commit institutional resources to the project. Champions help recruit participants, contextualize the work for participants, and advocate for the process both within the state education agency and across the stakeholder groups represented in each NIC. It is most effective when other stakeholders view the champion as knowledgeable and valuable in the context. In the case of the Michigan Focus NIC, the participating MDE personnel had long-standing relationships with some of the other NIC participants, which helped to build legitimacy of the endeavor.

In addition to the champion, NICs require distinct types of expertise—including content, context, and research expertise—to ensure accuracy and build legitimacy for its work. Content experts specialize in the content or disciplinary areas that the statewide system of support addresses. Context experts deeply understand the organization of the system, the responsibilities of stakeholders in the system, and the way those stakeholders are both supported and challenged.[1] They also are knowledgeable about the political and personal landscape of the local context and connect the NIC to resources, anticipate and propose solutions to barriers to implementation of the intervention, and provide guidance on how to structure the NIC for sustainability. Finally, research expertise is needed to support NICs' data collection and analysis work and contribute thinking to the development and assessment of outcome measures.

The network initiation team intentionally recruited to fulfill the aforementioned types of expertise. The needs of the group in this respect evolved over time. For example, as the Michigan Focus NIC identified improving students' mastery of benchmarks for mathematics

[1] Here, "system" refers to the complex organization of schooling in each state and how those who work within this organization (that is, educators) interact with one another.

fluency, the network initiation team considered it important to have a mathematics content expert who could speak knowledgeably about key considerations for student proficiency at various grade levels, the availability of assessments, and proposed curricula for the intervention. In addition, the multilayered system of support in Michigan required inclusion of at least one representative from each governance level as context experts to provide insight into the problem and to help address barriers that arose in developing the intervention, such as those related to how participating schools and districts make decisions about mathematics curricula and teaching practices. Finally, data specialists at the district level provided research expertise. This research expertise was supplemented by researchers from REL Midwest, which supported the NIC's data collection and analysis work and contributed thinking to the development and assessment of outcome measures.

Finally, the network initiation team aligned NIC work to ongoing efforts in the state to support Focus Schools. For example, participants wanted to know how this new work was different from the continuous improvement work they were already engaged in, which included annual needs assessments. The network initiation team acknowledged existing work through the statewide system of support and worked with state education agency staff to use terminology and align NIC work to the state's blueprint for its statewide system of supports. In presentations, researchers explicitly demonstrated how improvement science methods could be used to achieve the aims of the statewide system of support. For example, the Michigan Blueprint for School Turnaround references that districts should use a "problem-solving protocol using multiple measures of data" in working with schools; the network initiation team discussed how the NIC model could be used as one such protocol for data review.

Building trust, buy-in, and ownership. Again, the stepwise recruitment process was important for building trust among participants. Engaging in the work of an NIC requires participants to take a close look at their practices and be willing to admit that some practices are not working well and make changes accordingly. For some, this may be a difficult, and at times, personal process. In order to facilitate this process, it was necessary for the network initiation team to build trust between the project team and the participants as well as among the participants themselves. The multi-phase recruitment process helped build trust by providing many opportunities for the project team to interact with the participants prior to beginning the work of the NIC. Although this took several months, it was critical to build the foundation for the success of the NIC.

After the initial on-boarding process, the network initiation team employed a series of in-person monthly meetings so that the NIC participants would be able to come together and build formal and informal relationships to facilitate working together on the NIC work described later in this chapter. Meetings were structured in such a way to maximize participant-led discussion and group work. In addition, the work was rooted in participants' experience to ensure that all participants could contribute to the activities presented to the group. For example, when conducting a root cause analysis, the network initiation team asked participants to write down problems they had encountered in supporting Focus students in their own work in the past day, week, month, and year. Grounding the discussion in daily occurrences helped the network initiation team to build trust, buy-in, and ownership over the work. At other points during the school year, participants were asked to lead group discussions, report out on smaller group activities, research observation protocols or measurement and report back, or solve a problem for the group to move forward (e.g., obtaining permission from the superintendent to align district

professional development activities to Michigan Focus NIC efforts around mathematics fluency). These efforts to build trust were intentional on the part of the network initiation team and were closely linked to efforts to build legitimacy and membership, as well as establish norms.

Using NIC Core Technology to Improve Mathematics Fluency in Focus Schools

According to Bryk, Gomez, Grunow, and LeMahieu (2015): "All activity in improvement science is disciplined by three deceptively simple questions:

1. What specifically are we trying to accomplish?
2. What change might we introduce and why?
3. How will we know that a change is actually an improvement? (p. 114)

The three domains that comprise the core technology of NIC activity—developing a theory of practice improvement, building a measurement and analytics infrastructure, and learning and using improvement methods—all work to answer these questions. In the case of the Michigan Focus NIC, answering these questions led the group to concentrate its efforts on improving mathematics fluency for the bottom 30% of students within participating schools through at least 15 minutes of daily mathematics fluency practice. The group could have picked a variety of root causes on which to focus its efforts, however, and the story of how the group came to consensus on this aim is worth exploring for future improvement science efforts.

Developing a theory of practice improvement. The first component of continuous improvement is to make the work problem specific and user specific. To do this, Bryk et al. (2015) recommend identifying problems of practice through root cause analyses. In this way, participants can "see the system" and come to a consensus of what the diagnosis of the problem is before jumping to a solution. To make the best of data collected to understand the nature of the issues and inform solutions, Bryk and his colleagues (2011, 2015), among others, have adapted

several tools from business to improvement research work. The first tool is the fishbone diagram that facilitates the understanding of problems. The diagram visually presents key factors that might contribute to unsatisfactory outcomes and captures ideas and thoughts on conversations about these factors. The second tool is the system improvement map, an analytic tool illustrating the essential institutional features as the improvement work proceeds. The third tool is the driver diagram, serving as guidance and providing theory of practice for improvement work. The driver diagram highlights a set of hypothetical key levers for improvement, possible changes, and connections among the causal origins of the problem. In developing the driver diagram, it is suggested to use common language to build consensus and prioritize high-leverage ideas (Berg, Hough, & Taylor, 2015).

In the Michigan Focus NIC, the network initiation team used each of these tools to define and refine a problem of practice, identify an intervention, and ultimately develop a theory of problem practice (Appendix D). During the course of four 3-hour meetings in fall 2015, the network initiation team provided tools and activities for participants to identify the root causes of achievement gaps in Focus Schools, identify key improvement hypotheses, develop a theory of action, and define outcome measures. For example, in the October 2015 meeting, participants engaged in activities to bring the group to consensus about the definition of the problem and its root causes. In Activity 1, individuals identified problems of practice they had encountered in supporting Focus Schools in the past day, week, month, and year. Each participant brainstormed an uncapped a number of problems and wrote them down on different sticky notes. Then, participants used an emergent coding technique to group these problems into categories that formed the bases of the root causes of providing supports to Focus Schools. This activity helped individuals create a focused problem statement around supports for Focus Schools:

Focus Schools suffer from a lack of access to, understanding of, and use of data to implement, monitor, and evaluate continuous improvement on a daily basis due to time, priority, skills, and commitment.

This statement was culled from a set of problems that participants said they had faced in the past day, week, month, and year. Participants then grouped these problems into five broad categories:

- Leadership
- Curriculum and instruction
- Resources (for example, data and time)
- Alignment of policies and initiatives
- Family and community engagement

The categories, problem statements, and evidence culled from group members (that is, issues encountered with the provision or receipt of support) were then used to create a fishbone diagram in Activity 2 of root causes that the group would want to focus on to close achievement gaps in Focus Schools (see Appendix D for an example of the group's fishbone diagram). The primary problem statement developed in Activity 1 became the "head" of the fishbone diagram, and participants then worked together to identify the root causes of each problem based on asking a series of "Why?" questions.

After the root causes are identified, the team should pick one root cause on which to focus its efforts (Bryk et al., 2015). This involves creating a theory of action and a theory of change and then determining what change can be implemented to drive improvement on the agreed-upon outcome of interest. This intervention, or change, should be clearly defined, easily implemented, and produce change quickly. These are strong assumptions in any organization, but particularly in a complex environment such as a school, which has many competing interests

and actors. The worry would be that an intervention that fit the description of being clearly defined, easily implemented, and quick to produce change would not affect the kind of large-scale change needed to achieve policy aims of significant reduction in within-school inequality. The benefit, however, is that it breaks down a large-scale problem into a manageable one in order to empower actors at the lower end of the hierarchy to enact and monitor change.

The addition of school-level principals changed the emphasis of the group from the initial conception by the network initiation team. Initially, the network initiation team had intended for the Michigan Focus NIC to emphasize improving supports provided to Focus Schools; through activities to narrow the problem of practice on which the group would concentrate efforts, though, the emphasis gradually moved to the school level and the group decided to concentrate efforts on reducing achievement gaps within Focus Schools. When prompted to develop an intervention that would address the issue of data use for continuous improvement identified in the root cause analysis activity, participants further refined this problem statement to:

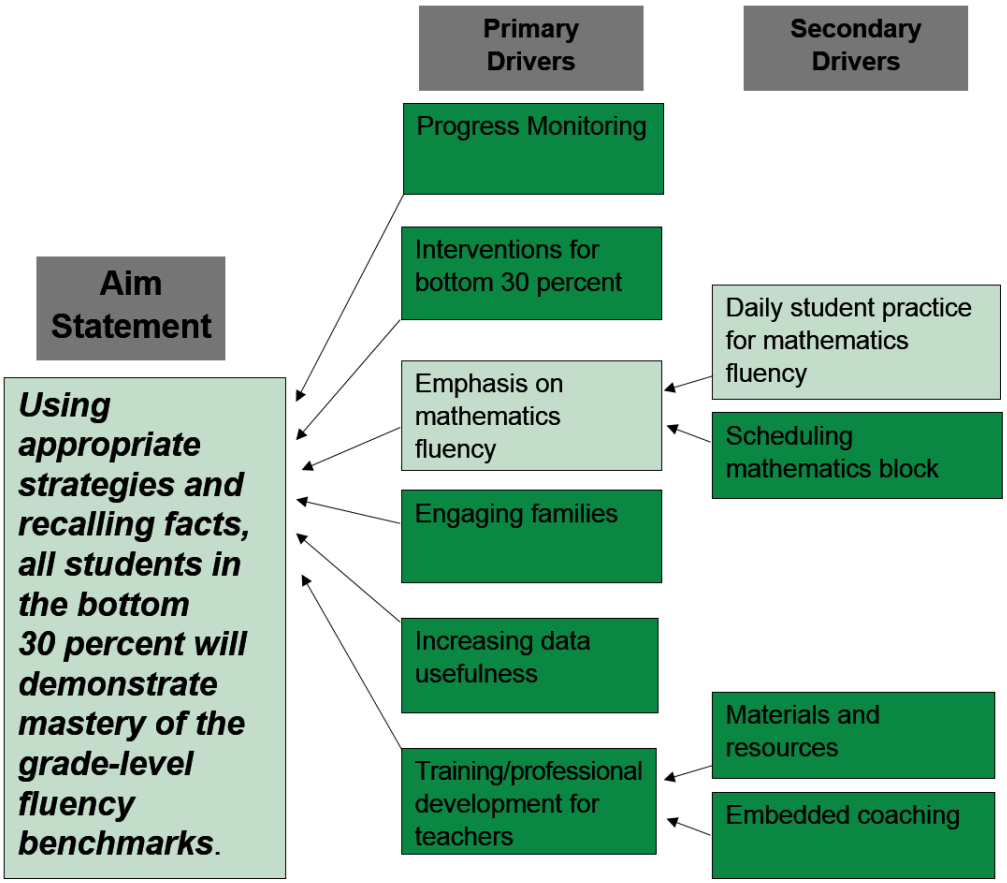
> *Students' lack of daily opportunities to practice mathematics fluency skills exacerbates inequalities in mathematics achievement in Focus Schools.*

It is important to note that participants did not choose to define mathematics fluency further for the purposes of the group. However, over time, principals reported having an increased number of conversations with teachers about what constituted best practices in mathematics fluency.

Developing the driver diagram (Figure 12) solidified the Focus NIC's ideas that gaps in achievement in mathematics between the top and bottom 30% of students were driving Focus identification and that those gaps could be explained by a lack of mathematics fluency skills among the bottom 30% of students. Based on this group work, the Michigan Focus NIC agreed

to target mathematics fluency as a primary driver of achievement gaps in their Focus Schools through the primary driver diagram below.

Figure 12. Michigan Focus NIC Primary Driver Diagram



The progression of the problem statement from data use to daily mathematics practice for struggling students warrants discussion, as this progression was intentionally wrought through a series of discussion in fall 2015 and required revisiting the root cause analysis activity to develop a problem statement that was both high-leverage and highly actionable (Bryk, et. al, 2015). A high level of interest among partners and stakeholders can enable initial and sustained

institutional support, buy-in and ownership of participants, and diffusion of practices and processes that are produced through the work of the NIC. However, problems ought to be actionable if participants are to enact changes in practice through rapid PDSA cycles and learn from those changes. As many schools felt they had been identified because of their large gaps in mathematics achievement, the participants felt that this statement was a more specific lever to address problems of within-school inequality than the previous, more general statement about data-driven decision making. The progression of the group's emphasis from one statement to the other was a direct result of trying to move from the results of one root cause analysis activity to developing an intervention. The group, forced to answer questions such as "Data use in what subject?" and "For whom?" and "Why?" felt that they needed to revise their problem statement to be more specific and actionable.

Consequently, the Michigan Focus NIC agreed to work on mathematics fluency as a primary driver of achievement gaps in their Focus Schools. The measurable aim that they worked together to address was to ensure the bottom 30% of students practice mathematics fluency skills daily for at least 15 minute per day. To do so, school principals in the NIC agreed to identify students in the bottom 30% for mathematics achievement and work with their teachers to ensure they received at least 15 minutes of daily mathematics fluency practice every day. Each school principal worked with the mathematics teachers in their schools to integrate mathematics fluency practice into existing workstations, and principals and the Focus NIC monitored and tracked their progress. In addition, principals worked with the central office and ISD representatives in the Focus NIC to ensure professional development opportunities were geared toward mathematics fluency. This specific focus on mathematics fluency meant including a district-level content expert and bringing in other staff who are knowledgeable in mathematics

fluency; however, it allowed the group to take a large problem—data use for continuous improvement—and implement it on a smaller scale and in a manner in which there is the potential to learn from it.

In November-December 2015, the group took these discussions and developed a theory of problem practice (Figure 13).

Figure 13. Theory of Action for Daily Mathematics Fluency Practice

Theory of Action			
Program Inputs	**Program Activities**	**Program Outputs**	**Outcomes**
Teacher logs to track daily math practice of fluency skills Implementation guide developed by Focus NIC Observation protocol developed by Focus NIC Principal guidance, coaching, and support to mathematics teachers Mathematics curricular kits or workstations District mathematics coach District and ISD-level mathematics fluency professional development and support	Identify bottom 30% of students Teachers track Focus students' ability to practice mathematics fluency skills for at least 15 minutes every day using daily logs Bimonthly walk-throughs using observation protocol Ongoing coaching and data use Daily teacher logs	Increased time for students spent on practicing mathematics fluency skills Increased time spent discussing mathematics fluency between teachers and between teachers and principal Increased mathematics fluency emphasis	Increased percentage of all students mastering mathematics fluency benchmarks by May 2016 Improved mathematics fluency of the bottom 30% of students specifically
Program Targets: mathematics teachers in Ingham ISD and KRESA who teach in Focus schools participating in the NIC; all students in mathematics classrooms in Focus schools participating in the NIC, with an emphasis on the bottom 30 percent of students.			
Program Goal: All students will master fluency benchmarks by demonstrating appropriate strategies and recalling facts.			

Building a measurement and analytics infrastructure. Another key component of improvement science is measurement, which plays a central role in facilitating the understanding of the problem and testing changes that indicate improvement. Measures (or indicators) include

long-term outcomes and short-term changes (Bryk et al., 2011). Three types of measures (i.e., outcome measures, process measures, and balancing measures) can be used to provide a balanced set of measurement for improvement (Bennett & Morales, 2015). Improvement work relies on data about outcome and program process. Bryk and colleagues (2011) suggested embedding data collection into the day-to-day work of program participants. To ensure that the information collected is informative and useful for understanding and improvement, Bryk and colleagues also suggested the data be accessible in a timely manner and provide evidence that is actionable.

The Michigan Focus NIC, with the help of REL Midwest researchers, developed a series of metrics to track leading and lagging indicators of students' daily mathematics fluency practice of at least 15 minutes per day (see Appendix C for the implementation guide). Teachers kept a daily log indicating whether the bottom 30 percent of students had the opportunity to practice mathematics skills on a given day or not. This was meant to be a simple "Yes/No" log that teachers could use to track students' participation in the intervention. However, even such a simple measure spurred discussion amongst the principals. For example, one principal identified students in the bottom 30% and gave each mathematics teacher a list of student names in their course from the bottom 30% —called "Focus Students"—and were asked to track individual students' abilities to participate in daily mathematics fluency practice. In other schools, principals used the checklist to report whether, on average, students in the bottom 30% were *given the opportunity* to receive 15 minutes of mathematics fluency practice daily. Eventually all principals began to identify students individually and asked teachers to track them because it provided for richer data and made it easier to identify which students were actually receiving the intervention. For example, if teachers provided data at the classroom level, those students who

were systematically being pulled out of the classroom for various reading interventions or students who were chronically absent would be counted as having received the treatment when in reality they were not present in the classroom. The important thing to note here, however, was that the group came to consensus on a way to track teachers; the network initiation team developed a measurement rubric, distributed it, the principals tested it, and then came back together with suggestions on improvement. The process of refining this measure was perhaps more meaningful because the principals had experience with it and suggested the more burdensome measure in collaboration with their teachers because they saw the utility of it.

Another issue that arose once principals and teachers began to monitor progress is that in some schools, almost all teachers reported that students had the opportunity to practice their mathematics fluency skills on a daily basis—that is, classrooms reported that 100% of the time, Focus Students in their classrooms were able to practice their skills. This led participants to question whether it was merely the *quantity* of mathematics fluency in practice in their classrooms or the *quality* of the practice. Consequently, the Focus NIC developed a short observation protocol in checklist form that Focus NIC principals used to conduct walk-throughs in their schools every two weeks to ensure students received training in mathematics fluency skills (see Appendix C for an example of this observation protocol form).

In addition to these interim or leading measures, the Focus NIC used AimsWeb and MAP formative assessment scores to track students' progress as well as school-level indicators for mathematics fluency per student. These data were collected and supplied by the school and were policy-relevant measures of students' mathematical fluency abilities, the improvement of which was the ultimate goal of the Michigan Focus NIC. These measures, in addition to the teacher logs and principal observation protocols, were scanned and sent to the research team on a monthly

basis for collation and presentation in a format that was visually appealing and easy to understand. The research team also led data review meetings to ensure that participants interpreted descriptives correctly and made appropriate inferences with respect to causality. An example of a data review packet can be found in Appendix D.

The development and deployment of these measures is demonstrative of how different the buy-in to do this kind of time-consuming monitoring may have been if the district or state had come in and mandated that principals must do these walk-throughs every two weeks. However, principals had been part of defining and refining the problem, developing the theory of action, and developing and testing the measures. They bought into the measurement and analytics infrastructure because they saw the utility of it. However, the research team did the work of taking the Focus NIC's ideas during monthly meetings and then developing the measures for testing, so it required the expertise of everyone in the group to ensure that data would be collected. In spring 2015, when the group reviewed data collected from these measures monthly, the research team led the discussion. It is unlikely that the group would have been able to do so on their own given time and resource constraints.

Learning and using improvement methods. The Michigan Focus NIC spent fall 2015 developing intervention and outcome measures to implement in spring 2016. The group also requested that REL Midwest staff extend the meeting calendar to run through the end of the 2015-16 school year so the group could implement multiple cycles and rounds of data review (Appendix B). The theory of action and corresponding measures allowed the Michigan Focus NIC members to implement two PDSA cycles in spring 2015. The literature suggests that PDSA cycles can be used to refine measures and the intervention and scale work to different and additional contexts (Lozano & Williams, 2015). As Bryk and his colleagues stated (2011), the

ultimate goal would be to develop a self-sustaining improvement system that would require "multiple iterations of refinement and possible larger changes to accommodate network growth and movement." This is the ideal scenario, but given funding cycles and REL resources, the sustainability aspect of this statement proved difficult in the case of the Michigan Focus NIC.

In Michigan, learning to plan for and implement PDSA cycles became part of the learning of the Michigan Focus NIC. The group went through two PDSA cycles, one from January – March 2016, and a second cycle from March – May 2016. The measures were tweaked between the first and second cycles, which separated the two. However, the intervention stayed primarily the same. Participants did not distinguish between cycles in the same way as the network initiation team. Part of the reason for this is because participants reported taking longer than anticipated to get up to "full implementation." For example, in some schools, competing initiatives posed challenges to all teachers using the logs and submitting them to the research team in a timely manner until March or April 2016. Still, throughout the two Plan-Do-Study-Act cycles, participants developed a shared understanding and language for how to implement the ideas they developed in fall 2015. They also reported wanting to take back the activities developed throughout the year to their respective ISDs, districts, and schools. This professional capacity building aspect is a key component of learning and using improvement science methods and was an explicit part of the formation of the Michigan Focus NIC (Russell, et al., 2016). Each part of the PDSA cycle and the ways in which it was implemented in the Michigan Focus NIC is described in greater detail below.

Plan. Over the course of several three-hour, in-person meetings during fall 2015, the Michigan Focus NIC came to a consensus on addressing students' mathematics fluency skills,

which group members saw as a primary driver of mathematics achievement inequality in Focus Schools. Consequently, the group coalesced around the following aim:

One hundred percent of Focus Students[2] will achieve mastery on grade-level math fluency benchmarks.

The group developed the following theory of action to delineate how the intervention—15-minute daily mathematics fluency practice for Focus Students— ultimately would affect this aim. These steps are described in greater detail above. The first planning phase took approximately four three-hour meetings to implement, which is a departure from what the literature suggests (8-12 weeks for an entire PDSA cycle). However, some of that planning time was also spent initiating the Michigan Focus NIC.

Do. With the planning from fall 2015 and the implementation guide, Michigan Focus NIC principals were ready to implement the intervention—daily mathematics practice for students—in their Focus Schools. Principals gave the daily log sheets to teachers and worked with them to identify the bottom 30% of students in mathematics achievement who would be considered Focus students for the purposes of this work.

Principals collected teacher logs weekly, scanned them, and sent them to the research team for data analysis. Principals also conducted observations of each teacher implementing the intervention (i.e., all mathematics teachers) at least once every two weeks and sent the scanned observation form to the research team as well. Teachers also received a copy of their principal observation sheet and retained a copy of their logs. Principals reported that teachers were eager to implement the intervention in their classrooms, viewing it as a potentially high-leverage yet

[2] Focus Students are those students who are in the bottom 30% of mathematics achievement. This designation stems from Michigan's identification of Focus Schools, which are defined as the 10 % of schools in the state that have the highest gaps between the top 30% and the bottom 30% of students within a school.

manageable change in practice. Principals also reported that they were having many conversations with teachers concerning mathematics fluency and that teachers were discussing mathematics fluency among themselves as a result of these efforts.

Study. The research team collected and entered the data sent by Focus NIC principals and created visually appealing data displays to share the information with the NIC in a way that facilitated data review (Appendix D). These data packets were de-identified by removing teacher, principals, school, and district names and identifiers after monthly data review meetings in February through May so principals could share them with their teachers and district participants could share them with district leaders without fear of identifying any particular teacher or school out of context.

At the monthly meetings, the Michigan Focus NIC participants—which included representatives from MDE; intermediate school district support staff for Focus Schools; central office administrators such as assistant superintendents, Title I coordinators, and mathematics coaches; and Focus School principals—reviewed the data packets together to determine what was working and what needed to be improved. The network initiation team facilitated these conversations to ensure that all participants felt comfortable sharing their experiences with implementation of the intervention and reviewing school-level data in a respectful and constructive manner. Agendas and data review prompts were disseminated prior to each monthly meeting so participants were well prepared (Appendix E).

Act. As a result of the data inquiry cycle described above (Cycle 1, January–March 2016), Focus NIC participants revised the teacher log to specify individual students (one school had already done so in Cycle 1 and participants decided this approach was most useful). In addition, the Focus NIC realized there was variability in teacher feedback on students'

mathematics fluency practice (which arose after reviewing principal observation data across participating Focus Schools). The mathematics coach of one district worked with the Title I coordinator and the curriculum specialist from a different intermediate school district to develop a rubric that teachers could reference to model teacher feedback with respect to mathematics fluency. These materials—the revised teacher log and rubric—were disseminated to participating teachers in Cycle 2, which took place from April to June 2016. A similar implementation and data inquiry cycle resulted from Cycle 2.

Participant Perceptions of the Michigan Focus NIC

The following feedback was solicited anonymously from NIC participants over the course of the project by asking three questions at the end of two sessions in the middle and toward the end of the project:

- What about the NIC has worked well?
- What is challenging?
- What about the NIC can be improved?

Respondents were overwhelmingly positive in their feedback. For example, in response to the first question, respondents reported answering:

- "The NIC has allowed me to influence initiatives at the building level."
- "The NIC is forcing me to slow down and learn before I begin working, which is difficult for me, but beneficial."
- "We have brought [multiple] levels together to create a solid plan."

- "The communication has been very good. I receive regular updates and summaries. Questions are always thoughtfully and fully answered. Our goals are clear and organized."
- "The part of this process that has been most helpful is placing a magnifying glass on our Focus Student … implementing an intervention after examining data and factors … professional discussions outside of my district."

With respect to the aspects of the Michigan Focus NIC that were challenging, respondents reported time out of their buildings or away from their jobs as a universal barrier to participation. During an end of the school year debrief in June 2016, participants reported wishing that the work of the Michigan Focus NIC were embedded in their daily work and not something "extra" that they were participating in, but rather counted towards district and state compliance efforts—that is, while schools and districts were held accountable for reducing inequality in Focus Schools under their purview, participation in the Michigan Focus NIC was done on their own time and did not replace other efforts, such as filling out work logs and compliance forms for the state. However, they acknowledged that mandating participation would change the dynamic of the group as well. Other aspects of the Michigan Focus NIC that could have been improved included having a central repository for agendas, data packets, and accompanying materials such that participants could access them and share with colleagues not associated with the NIC.

Lessons Learned: Promises and Pitfalls of the NIC Approach

As described in this chapter, initiating an NIC requires attention to the social and technical aspects of learning, both as individuals and as organizations. In Michigan, the work of the Focus NIC will help build capacity on how to implement improvement science to address problems of practice across governance levels and regional contexts. In addition, in Focus

Schools, an emphasis on improving math fluency skills among the bottom 30% of students may lead to improved outcomes for students on math fluency benchmarks. This particular study was focused on how an NIC might be used to achieve those goals. Due to timing and lack of funds to sustain efforts beyond the REL contract, this study cannot report on real effects on student achievement, but it demonstrates the way that such an improvement community has the potential to produce effects on achievement if monitored for a longer period of time.

More broadly, however, there is growing interest in the ability of improvement science to spur innovation and to address complex problems of practice through iterative PDSA cycles. In education, this methodology is often implemented through collaborative partnerships in which researchers and practitioners work together to systematically test and refine theories of change in real-world settings. NICs are an example of one such partnership. While the central work of NICs is primarily conducted at a local level, the lessons learned from these local contexts can be brought back to the network and potentially impact education practices more globally. The work of the Michigan Focus NIC is an illustrative example of how to establish a Networked Improvement Community in the context of high-stakes accountability to spur school improvement.

Despite the significance of these efforts, however, there are limitations on the ability of an NIC to break through the social organization of schooling. For example, without an external facilitator or hub of the network, such as REL Midwest, it is unlikely that the individuals in the Michigan Focus NIC would have come together to be able to work on a common problem of practice. There are fewer examples of a network hub that operates within the educational system. One question is whether an organization embedded within the system can maintain the same function as an external organization that serves as the hub of a network.

Another limitation is the inability to link Michigan Focus NIC efforts with student outcomes in a causal manner. For example, if the schools where daily mathematics fluency practice had been selected randomly by the district, then even a short PDSA cycle would provide participants with the ability to conduct rigorous experiments in their contexts and learn from them. As part of the 2015-16 initiation of the Michigan Focus NIC efforts centered on developing professional capacity for improvement efforts, this was not feasible. Consequently, any improvements in students' mathematics fluency practice cannot be attributed to the efforts of the Michigan Focus NIC. I would argue, however, that a better outcome measure would be the development of an improvement mindset and the establishment of relationships and trust between governance levels within the Michigan educational system. Future research efforts in this arena should seek to measure these types of outcomes and systematically track the way they change over time.

Still, the Michigan Focus NIC is evidence that individuals from schools, districts, ISDs, and states—which may have reasons to interact with one another in antagonistic, bureaucratic ways—can come together to focus on problems of practice related to teaching and learning to help struggling learners. This promise relies, as all educational endeavors, on the capacity of individuals to commit to and fully participate in an improvement community. However, it yields promise for scalability in ways unlike siloed programmatic efforts of the past.

References

Bennett, B., & Morales, L. (2015, March). *How do you know that a change is an improvement? Using data for improvement*. Presentation given at the second annual Carnegie Foundation Summit on Improvement in Education, San Francisco, CA.

Berg, J., Hough, H., & Taylor, J. (2015, March). *From aim to action: Developing a theory of practice improvement.* Presentation given at the second annual Carnegie Foundation Summit on Improvement in Education, San Francisco, CA.

Berwick, D.M. (2008). The science of improvement. *Journal of the American Medical Association*, *299*(10),1182–1184.

Bhatt, M., & Proger, A. (2016). Life stages of research-practice partnerships. Unpublished manuscript.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building Networked Improvement Communities in education. In M. Hallinan (Ed.), *Frontiers in sociology of education* (pp. 127–162). New York, NY: Springer.

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*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement.* New York, NY: Russell Sage Foundation.

Bryson, J. M. (1988). A strategic planning process for public and non-profit organizations. *Long Range Planning, 21*(1), 73–81

Clifford, M., & Millar, S. B. (2007). K–20 partnership: A definition and proof of concept (WCER Working Paper No. 2007-9). Madison, WI: Wisconsin Center for Education Research. Retrieved from http://www.wcer.wisc.edu/publications/workingPapers/working_paper_no_2007_09.pdf

Cobb, P., Jackson, K., & Dunlap, C. (2015). Design research: An analysis and critique. In L. English & D. Kirshner (Eds.), *Handbook of international research in mathematics education* (3rd ed.). Routledge.

Cobb, P., & Jackson, K. (2012). Analyzing educational policies: A learning design perspective, *Journal of the Learning Sciences, 21*(4), 487–521.

Coburn, C. E., Penuel, W. R., & Geil, K. (2013). *Research-practice partnerships at the district level: A new strategy for leveraging research for educational improvement*. New York, NY: William T. Grant Foundation.

Coburn, C. E., Penuel, W. R., & Geil, K. E. (2013, January). *A strategy for leveraging research for educational improvement in school districts.* New York, NY: William T. Grant Foundation.

Coburn, C. E., & Stein, M. K. (Eds.). (2010). *Research and practice in education: Building alliances, bridging the divide*. Lanham, MD: Rowman & Littlefield.

Cohen, D. K., & Moffitt, S. L. (2009). *The ordeal of equality: Did federal regulation fix the schools?* Cambridge, MA: Harvard University Press.

Conaway, C., Keesler, V., & Schwartz, N. (2015). What research do state education agencies really need? The promise and limitations of state longitudinal data systems. *Educational Evaluation and Policy Analysis*, *37*(1), 16S–28S.

Connolly, F., Plank, S., & Rhone, T. (2012). Baltimore Education Research Consortium: A consideration of past, present, and future. Retrieved from http://baltimore-berc.org/wp-content/uploads/2013/05/BERCNurturingConsortiaPaper.pdf

Creed, W.E.D., & Miles, R.E. (1996). Trust in organizations: a conceptual framework. In R. M. Kramer & T.R. Tyler (Eds.), *Trust in organizations: Frontiers of theory and research.* London, UK: Sage, 16–39.

Deming, W. E. (1993). *The new economics for industry, government & education.* Cambridge, MA: Massachusetts Institute of Technology Press.

Englebart, D. C. (2003, September). *Improving our ability to improve: A call for investment in a new future.* Paper presented at the IBM Symposium on the Coevolution of Technology-Business Innovations, Boulder, CO.

Firestone, W. A., & Fisler, J. L. (2002). Politics, community, and leadership in a school–university partnership. *Educational Administration Quarterly*, *38*(4), 449–493.

Hord, S., & Sommers, W. (2008). Leading professional learning communities: Voices from research and practice. Thousand Oaks, CA: Corwin Press.

Jacob, B. (2015). Harnessing the value of "failure." Brookings Institution. Retrieved from https://www.brookings.edu/research/harnessing-the-value-of-failure/

Langley, G., Moen, R., Nolan, K., Nolan, T., Norman, C., & Provost, L. (2009). *The improvement guide* (2nd ed.). San Francisco, CA: Jossey-Bass.

Lewis, C. (2015). What is improvement science? Do we need it in education? *Educational Researcher, 44*(1), 54–61.

Lozano, M. and Williams, D. M. (2015, March). *A powerful engine for change: Applying the method for improvement.* Work presented at the Carnegie Foundation Summit on Improvement in Education in San Francisco, CA.

Martin, W.G., & Gobstein, H. (2015). Generating a Networked Improvement Community to improve secondary mathematics teacher preparation: Network leadership, organization, and operation. *Journal of Teacher Education, 66,* 482–493.

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York, NY: The Guilford Press.

Postlethwaite, T.N. (2005). *Educational research: Some basic concepts and terminology.* Paris, France: International Institute for Education Planning/UNESCO. Retrieved from http://www.unesco.org/iiep/PDF/TR_Mods/Qu_Mod1.pdf.

Peurach, D. (2016). Innovating at the nexus of impact and improvement: Leading Educational Improvement Networks. *Educational Researcher,45*(7), 421-429.

Roderick, M., Easton, J., & Sebring, P. (2009). The Consortium on Chicago School Research: A new model for the role of research in supporting urban school reform. Chicago, IL: Consortium on Chicago School Research. Retrieved from https://ccsr.uchicago.edu/sites/default/files/publications/CCSR%20Model%20Report-final.pdf

Rosenquist, B.A., Henrick, E.C., & Smith, T.M. (2015). Research–practice partnerships to support the development of high quality mathematics instruction for all students. *Journal of Education for Students Placed at Risk 20*(1-2), 42–57.

Russell, J. L., Bryk, A. S., Dolle, J., Gomez, L. M., LeMahieu, P., & Grunow, A. (2016). A framework for the initiation of Networked Improvement Communities. *Teachers College Record*. Retrieved from http://researchandpractice.org/wp-content/uploads/2016/02/NIC-Initiation-Framework_Russell_TCR_FINAL-SUBMISSION.pdf.

Turley, R.N.L., & Stevens, C. (2015). Lessons from a school district- university research partnership: The Houston Education Research Consortium (HERC). *Educational Evaluation and Policy Analysis, 37*(1S), 6S-15S.

Vangen, S., & Huxham, C. (2003). Nurturing collaborative relations: Building trust in interorganizational collaboration. *Journal of Applied Behavioural Science 39*, (1), 5–31.

Webb, A. (1991). Coordination: A problem in public sector management. *Policy and Politics*, *19*(4), 229–241.

Wenger, E., & Wenger-Treyner, B. (2015). Communities of Practice: A Brief Introduction. Retrieved from http://wenger-trayner.com/wp-content/uploads/2015/04/07-Brief-introduction-to-communities-of-practice.pdf

Chapter 5

Conclusion

Educational reformers in the U.S. have been successfully advocating for standards-based reform policies since the early 1990s (Smith & O'Day, 1991). The resulting exoskeleton of standards, curriculum, and assessment allowed practitioners the freedom to reach set achievement targets however they best saw fit in their specific contexts. However, after 25 years of employing such a strategy, policymakers are recognizing the problematic nature of leaving the development of best pedagogical practices to chance and happenstance (Cohen & Moffitt, 2009). Consequently, recent educational reforms have made efforts to connect individuals and organizations particularly to support schools in need of improvement (Peurach, 2016).

Since No Child Left Behind (NCLB), the federal government has suggested that schools that consistently fail to meet benchmarked goals are in need of extra support. ESEA waivers offered flexibility from many provisions in NCLB but still relied on a basic system of identifying schools and supporting them in their improvement efforts. The resulting statewide systems of support (SSOS) built on prior efforts to support low-performing schools while attempting to develop new capabilities and economies of scale. With respect to Focus Schools, however, states had no prior playbook to rely on when it came to supporting high achievement-gap schools in reducing within-school inequality. Consequently, the development of a system to identify and support Focus Schools presents an interesting case to examine whether and how states can play a role in reducing inequality within schools under an accountability framework.

In this dissertation, I ask and answer four separate questions related to identifying and supporting Focus Schools in Michigan. First, what was the impetus for the policy with respect to its goals, measures, and design? Second, what are the merits of this design in achieving its policy goals, namely the reduction of within-school achievement gaps? Third, given such a policy, what are the organizational elements needed for an optimal response? Fourth, can improvement science provide a methodology to develop and sustain those organizational elements? These questions and the findings that correspond to them represent a shift in education reform from a hands-off approach to accountability to developing coordinated capacity across geographic locations and governance levels to support struggling students.

The first set of questions provides contextual and historical explanations for the federally designed ESEA waivers. In the midst of mounting criticism towards No Child Left Behind coupled with the lack of Congressional action, the U.S. Department of Education issued ESEA waivers to provide states with a reprieve from NCLB regulations. These waivers continued to require identification of low-performing and high-performing schools, but added a third category of schools to identify and support: Focus Schools, or high achievement-gap schools.

Inequalities in student outcomes have long been a topic of concern in U.S. public schools. Most recently, researchers have highlighted growing gaps in educational achievement and attainment between high-income and low-income students (Reardon, 2011; Duncan & Muranne, 2014). However, the school's role in these discussions around inequality primarily centered on inputs such as resources and funding levels. Even under NCLB, achievement gaps were identified at the subgroup level, and schools were held accountable if specific subgroups of students did not meet benchmark standards. The ESEA waivers changed this focus by allowing states to identify schools with high achievement gaps between high- and low-performing

students. In this sense, the policy goals of reducing within-school inequality are a departure from previous attempts to improve students' access to educational opportunities. For example, one could imagine reducing variation in student achievement within a school by lowering student achievement for high-performing students or by creating homogenous student bodies. These potential unintended consequences must be guarded against in any continuation of ESEA waiver stipulations under the newly authorized Every Student Succeeds Act of 2015, which codified many of the regulations set forth in the waivers.

In Michigan, this Focus set was defined as schools with the highest gaps between the top third and bottom third of students. This measure of within-school inequality created "super subgroups" of students who comprised the bottom 30 percent. For example, students in the bottom 30% might be economically disadvantaged, receive special educational services, or speak English as a second language. Under Michigan's accountability system, these students would be treated as a monolithic block of struggling students. On the one hand, this measure has identified schools with struggling students that escaped identification under previous accountability systems. By providing a measure agnostic of students' demographic characteristic, it also allows teachers and principals to focus on students who are struggling academically instead of typecasting students based on characteristics such as race or ethnicity. On the other hand, this undifferentiated treatment may not be a useful or meaningful metric for teachers in supporting struggling students. In addition, it uses student achievement as the only metric of identification for struggling students in lieu of looking at more meaningful outcomes, such as educational attainment or GPA or even socioemotional outcomes.

Finally, the ESEA waivers provided flexibility to states to identify and support schools. In some states, the resulting accountability system was to develop a strong statewide system of

supports. In Michigan, the resulting reforms were more aligned with a "results without rancor" model in which districts played a key role in supporting Focus Schools (Jochim & Murphy, 2013). This devolved responsibility has the advantage of allowing for contextual variation. However, it presupposes that districts understand the steps needed to support the reduction of within-school inequality and also have the capacity to implement these interventions. It also ignores the reality that local conditions likely contributed to within-school inequalities in the first place.

Despite these types of post hoc questions raised by the policy goals, measures, and design, the reality for many districts and schools in 2012 called for them to act within the system developed by the state department of education. Results from a regression discontinuity design suggest that the effects of Focus identification and support were null, though the study is underpowered and may not be able to detect smaller effects. If these results are to be believed, they likely mask an underlying distribution of effects that denote heterogeneity based on school or district characteristics. Results from the qualitative case study suggest that districts that relied on a strategy of coordinated capacity, or a systemic effort to harness individual and organizational capacity, were more successful in supporting Focus Schools and reducing within-school achievement gaps.

Chapters 2, 3, and 4 answer questions regarding policy design, impact, and implementation, respectively. Oftentimes, these questions are asked and answered separately: What is the design of a policy? What effect did it have, if any? Given this design, how do individuals and organizations implement it? Chapter 4 suggests a methodology, improvement science, in which these questions are asked and answered in a collaborative fashion. The promise of such an approach, represented here by a Networked Improvement Community to support

Focus Schools, has been well documented in the literature. The ability of states and districts to implement such an approach with limited resources, however, has not been proven at scale. In Chapter 4, we see that efforts to utilize this type of improvement science approach in a real-world context requires a strong hub leader, varying levels and types of expertise, and elements of "improvement infrastructure" to enable such work. This type of educational reform is still in its "fragile" and nascent stages; consequently, whether it can be enacted at scale is an open question. However, it is likely that efforts under ESSA will demand more, not less, of this type of coordinated approach. Consequently, it is worth investigating how micro and macro efforts can be intertwined to produce more equitable outcomes for students across schools, districts, and states.

Appendix A

Interview Protocols

Central Office Representative Interview Protocol

District / ISD Context

1. I would like to start by asking you to tell me about your district / ISD. How would you briefly describe your district / ISD to someone who has never been here before, including who it serves, how it is organized, its strengths, challenges and priorities?

   Listen for comments regarding:

   - *The students*
   - *The schools and differences among schools*
   - *The faculty (e.g., quality, shortage)*
   - *The fiscal context*
   - *Relationship with the school board*
   - *Administrator turnover*
   - *Other school reform efforts*

2. How many Focus Schools do you have in your district / ISD? How many of these are Title-I schools?

3. It is my understanding that Focus Schools are, at least in part, aimed at reducing within-school inequality in student achievement. Is that an accurate characterization?

4. Have there been efforts to reduce within-school inequality in the past? In your opinion, have these efforts been successful? If so, what has been helpful? What has hindered improvement efforts in this area? Probe as necessary:

   - *How do elementary schools differ from middle and/or high schools in what contributes to within-school inequality and what they need to improve?*
   - *Are there any federal, state, ISD or district policies and practices that may have (inadvertently) contributed to these achievement gaps?*

   Things to listen for:

   - *Characteristics of the students, parents, community*
   - *Leadership*
   - *Teacher capacity*
   - *Resources*
   - *School climate (e.g., safety, discipline, policies)*
   - *Implementation of a Multi-Tiered System of Supports (MTSS) structure*

5. Generally speaking, what do you think the Focus Schools in your district /ISD need in order to improve performance?

   Probe as necessary:

   - *Funding*
   - *Expertise*
   - *Staff capacity*
   - *Others?*

Identification of Focus Schools

6. Which schools in your district / ISD would you describe as having high levels of inequality among students? How many schools are identified as Focus Schools?

7. Why do you believe these schools in your district / ISD are identified as Focus Schools?

8. Do you believe that the list of Focus Schools in your district / ISD targeted the schools that were most in need of improvement from an equity standpoint? Why or why not?

9. Were there schools left off the list that you thought needed to receive targeted support for reducing inequality? If so, what criteria or characteristics would you use to identify these schools?

   Interviewer note: If the central office representative says that all schools with achievement gaps are supported whether or not they are labeled as Focus Schools, tweak wording in following questions to continue to probe on whether distinction between Focus School supports and non-Focus School supports within this district / ISD exists at all and in what ways.

Supports for Focus Schools

10. I'd like to ask you questions about supports provided to Focus Schools. If I were a principal in a Focus School in your district / ISD, what type of support would I receive, whether from the state, ISD, district, or another support provider?

    Probe as necessary:

    - *What supports are emphasized at the Focus Schools in your district / ISD?*
    - *To what extent is this support different from that of non-Focus Schools, if at all?*
    - *Has the district / ISD hired additional staff for the purpose of supporting Focus Schools?*

    Things to listen for:

    - *Professional development for principals / leaders (e.g., form, content, intensity)*
    - *Professional development for teachers (e.g., form, content, intensity)*
    - *Intensity of support services (how often)*
    - *Fit between support provider and school needs*

11. Is this support received by Focus Schools different than the support that other schools receive? If so, how?

12. Is there an district / ISD-wide approach to instructional improvement or other areas planned or being implemented for Focus Schools? To what extent is the approach different from that of non-Focus schools, if at all? What specific strategies are involved in this approach? Please explain.

    Things to listen for:

    - *Curriculum changes*
    - *Pedagogical strategies*
    - *Change in use of time / time for instruction*
    - *Parent / community involvement*
    - *Support for at-risk students*

13. On a scale from 1 to 10, where 1 is "not at all prescriptive" and 10 is "completely prescriptive," how prescriptive is the district / ISD in planning or implementing these strategies? If schools have some flexibility, please describe this. In what areas or to what extent do schools have flexibility?

14. How are these instructional improvement strategies different from changes you or others in your district / ISD have tried to make in the past?

    Note: Central office representative may not have been in district / ISD long enough to answer this question. In this case, move on to next question.

15. Are there any schools within your district / ISD that have had **unusual** or **noteworthy** experiences with implementation of supports for Focus Schools?

    Probe as necessary:

    - *What makes these cases noteworthy (e.g., school size, resources, leadership, community engagement)?*
    - *How would you characterize the average experience of schools within your district / ISD with implementation of supports for Focus Schools?*

16. In your view, have Focus Schools been successful in reducing achievement gaps within schools? What has been most helpful in supporting Focus Schools to reduce their within-school achievement gap? What has been least helpful?

17. How would you and/or your colleagues define success for the implementation of supports for Focus Schools?

State Role

18. What support and guidance is your state department of education providing to your district / ISD to administer supports to Focus Schools?

Probe as necessary

- *What has been most helpful? What has been least helpful?*
- *Prescriptive or flexible?*
- *Implementation of supports for Focus Schools?*
- *Use of funds?*
- *Reporting requirements?*

19. How does information about Focus School supports get filtered down from the state to the school-level?

20. Is there anything I have not yet asked that you think is important to know about identification, measurement of, or supports for Focus Schools?

School Administrator Interview Protocol

1. Can you to tell me a bit about your background as an educator? When and how did you come to be a principal at this school?

   Probe if not mentioned:

   - *How many years of experience do you have working as a principal?*
   - *Do you have experience working in schools that are similar to this one?*
   - *Have you worked in a school identified as having significant inequalities between the highest performing and lowest performing groups of students?*
   - *Do you have experience trying to reduce within-school inequality?*
   - *Did you work in the school before becoming principal? If yes, for how long and in what capacity?*
   - *When did you become the principal of this school? [month/year]*
   - *Did you work in the district (at central office or in another school within the district) before becoming principal of this school? If yes, for how long and in what capacity?*

2. I know that every school is unique, with unique circumstances that influence daily life. How would you describe your school to someone who has never been here before, including who it serves, how it is organized, its strengths, challenges and priorities?

   Things to listen for:

   - *Faculty: capacity (e.g., teacher knowledge and skills), commitment, turnover*
   - *Structures (e.g., departments/clusters, specialized classes, Multi-Tiered System of Supports [MTSS])*
   - *Students (diversity of student population, ethnicity, socioeconomic status, English language learners, special education students)*
   - *Community*
   - *Funding level*
   - *For high schools only: How is your high school organized (e.g., departments or clusters)?*
   - *Family / community ties*

3. Why do you believe this school has been identified as a Focus School?

   Note: If the respondent's answer includes a discussion of inequality, then move forward. If not, say "I also understand that Focus Schools aim to address problems of within-school inequality. Would you agree with that characterization?"

4. Generally speaking, what do you think your school needs in order to address the problem of inequality you just described?

   Things to listen for:

   - *Funds*
   - *Expertise (e.g., about effective practices, struggling students, etc.)*

- *Data needs*
- *Staff capacity*
- *Teacher quality*
- *Parent involvement*

5. How would you and/or your colleagues define success for the implementation of supports for Focus Schools?

6. How are you implementing MTSS? How are you monitoring the achievement of the bottom 30 percent and the gap between the top and bottom 30 percent?

7. In your view, have Focus Schools been successful in reducing achievement gaps within schools? What has been most helpful in supporting Focus Schools to reduce their within-school achievement gap in your school? What has been least helpful in your school?

   Probe as necessary:

   - *School size?*
   - *School organization (e.g., departments or clusters, especially in high schools)?*
   - *Professional capacity of staff?*
   - *Professional community?*
   - *Clear vision?*
   - *Data inquiry cycle?*
   - *External support?*

8. Could you describe the specific improvement strategies your school has implemented this school year? How have you prioritized these strategies? *[Interviewer note: The focus of this question is on the current improvement efforts in the school, regardless of the funding source (e.g., even if they are not funded by Focus-School-specific Title I money).]*

   Probe as necessary:

   - *Why were these strategies or combinations of strategies chosen (e.g., after an assessment of needs)? [Interviewer note: The focus of this probe is on leverage points, theories of action.]*

   Things to listen for:

   - *District facilitator*
   - *Data dialogues*
   - *Superintendent's Dropout Challenge*
   - *Overall strategies*
   - *Staffing changes*
   - *Improving instructional programs*
   - *Professional development (for both staff and principal)*
   - *Use of data*
   - *Supplies / materials / technology*
   - *Extended learning time*

- *School climate issues*
- *Community engagement*
- *Facilities improvements*
- *Other support*

9. I would like to focus on the principal's role in implementing supports for Focus Schools and what leadership, governance and decision-making roles look like in your school. To start, what is your main priority as principal?

   Things to listen for:

   - *Fostering a coherent instructional vision and strategies*
   - *Supporting the faculty and staff members in the change process*
   - *Creating opportunities for professional learning (e.g., professional development, coaching, fostering collaboration)*
   - *Monitoring instruction and monitoring student learning (e.g., observing classrooms, meeting with teachers to discuss students' progress, tracking progress of individual students, etc.)*
   - *Managing the internal environment (e.g., building a leadership team, motivating students and teachers, managing facilities, etc.)*
   - *Managing the external environment (e.g., meeting with parents, communicating with the district, building support from external organizations, etc.)*
   - *Ensuring that the MTSS structure is implemented effectively*
   - *Ensuring that Tier One instruction is effective for 85 percent of students*
   - *Ensuring that teachers are not putting kids into Tier Two or Tier Three without reason*
   - *Ensuring that teachers who are not successful at Tier One have support*

10. [If not answered in Q5] What are your activities in a typical week?

    Probe as necessary:

    - *What percentage of that typical week is spent on instruction and learning (what's going on in the classrooms), and what percentage is taken up by the "external environment"?*
    - *How much time do you spend in the classrooms during a typical week?*

11. I would like to turn to the supports you receive external to your school environment. From the district, what support do you receive (or have access to) to implement changes in your school?

    - Other district staff?
    - ISD?
    - State?

12. What resources or supports do you have within the school that help you implement desired strategies and changes? *[PAUSE]*

Follow up questions, if needed: From the School Improvement team? Other school staff? Do you think you have the support you need to achieve your goals? Why or why not?

13. Are the resources and support provided by your ISD / district equally shared across Focus Schools?

14. Does the ISD / district differentiate support for Focus Schools?

15. I'd like to ask you a few questions about your staff. How would you describe the staff at this school?

    Probe if not mentioned:

    - *Strengths/ weaknesses of the staff?*
    - *Do the staff have the capacity to implement the new reforms being asked of them? If no, why not?*
    - *Do the staff have the will to implement the new reforms being asked of them? If no, why not?*
    - *Do the staff have the time to implement the new reforms being asked of them? If no, why not?*
    - *Were there any changes in personnel this year?*
    - *Is there high staff turnover?*
    - *How would you characterize the feelings of the school staff toward the Focus School efforts? Do they support them?*

16. How many of your staff members are able to effectively teach 85 percent of their students, so that they are proficient with Tier One instruction and do not require intervention?

17. Are the staff in your school participating in professional development or learning activities designed to help them implement the Focus School supports? If so, please describe them for me.

    Things to listen for:

    - *What is the form of professional development (e.g., in-person, off-site, distance learning)?*
    - *Who participates?*
    - *Who provides professional development?*
    - *What is the content?*
    - *How frequent, over what time frame, how many hours?*
    - *Is it embedded in the school day?*

18. Is there anything I have not yet asked that you think is important to knowing about identification, measurement of, or supports for Focus Schools?

Teacher Interview Protocol

19. I'd like to start by learning a bit about your background, including how long you've been a teacher. Can you please describe your career to me?

    Probe as necessary:

    - *How many years have you worked at this school? In what capacities (e.g., department chair, coach, group leader, technology mentor)?*
    - *What do you currently teach at this school (grade / subject area)?*
    - *How many years have you worked in this district?*
    - *Where were you teaching prior to this school?*

Classroom Organization

20. I'm interested in learning more about your students. How would you describe the students in your classes?

    Probe as necessary:

    - *What is the range of student backgrounds and abilities?*
    - *How do you ensure that the students in the bottom 30 percent are achieving and that there is no achievement gap between the top and bottom 30 percent?*
    - *How do you ensure that your Tier One instruction is effective for 85 percent of all students?*
    - *How do you respond when students do not learn in Tier One?*
    - *Are there systems in place to screen for appropriate interventions?*
    - *Is the Multi-Tiered System of Supports (MTSS) structure implemented in the school ensuring that 100 percent of students are successful?*

21. What are your goals for your students this year? What do you want them to accomplish, and how will you know if they have succeeded?

    Probe as necessary:

    - *Could you give me some examples of how you know you are successful with the students in your classroom?*
    - *Have your goals changed at all over the past year or two? Why?*
    - *In a typical class period, how likely is it that you are able to meet your goals for that class?*
    - *What proportion of your students typically are able to meet the class goals?*

22. How do you decide what you teach, especially given the range of students you've described? When I say "decide what you teach," I'm thinking both about content and pacing.

23. What instructional and curricular approaches do you rely on most? This year, are you (or your team) intentionally making any changes in your approaches? If it makes it easier to

answer, feel free to discuss instructional techniques or pedagogy first, and then curriculum or content second. .

Probe as necessary:

- *What pedagogical approach (e.g., direct instruction)?*
- *What curriculum (e.g., content area, scope, and sequence)*
- *What strategies do you use for meeting the needs of specific groups of students (e.g., English language learners, special education students, the lowest performing students)?*

24. What types of challenges do you face with regard to helping your students improve academic outcomes? If there are too many to name, what are the biggest 3-4 challenges that come to mind?

25. *Note: This will probably result in a long and varied list. Encourage the participant to be forthcoming by listening and nodding.*

26. Please describe the students who are struggling in school. How do you address their needs?

Probe as necessary:

- *What kinds of problems do students have? Academic problems (i.e., keeping up with material)? Behavioral problems?*

School Organization

27. Now let's talk about the whole school. Can you tell me a bit about your school? How would you describe this school to someone who has never been here before? Who does it serve? How is it organized? What are its key strengths, challenges and priorities?

28. Is there a clear set of goals for the school as a whole? If so, what are your school's goals? (*If the teacher did not already refer to reducing achievement gaps, ask the following*): Are there any school goals related to reducing achievement gaps?

29. How would you describe the teaching staff at this school? What are their strengths and weaknesses as a staff?

Probe as necessary:

- *Are there staff members with whom you plan or teach lessons or assess your students? What does working together look like?*

30. How would you characterize the leadership of this school? Who are the key leaders, and what do they do to move the school forward and support you as a teacher? To what extent do you think they are effective in leading particular aspects of the school?

Things to listen for:

- *Providing instructional leadership*

- *Spending sufficient time in classrooms*
- *Developing vision or goals for the school*
- *Creating opportunities for professional learning*
- *Motivating teachers and students*
- *Evaluating teachers*
- *Building relationships with parents and others outside the school*
- *Making sure that things run smoothly on campus*

Focus School Supports

31. Now I'd like to talk about your school's Focus School designation specifically. Why do you believe this school has been identified as a Focus School?

32. *Note: If the respondent's answer includes a discussion of inequality, then move forward. If not, say "I also understand that Focus Schools aim to address problems of within-school inequality. Would you agree with that characterization?*

33. Generally speaking, what do you think your school needs in order to address the problem of equality you just described (e.g., funding, expertise, staff capacity, etc.)?

    Things to listen for:

    - *Funds*
    - *Expertise (e.g., about effective practices, struggling students, etc.)*
    - *Data needs*
    - *Staff capacity*
    - *Teacher quality*
    - *Parent involvement*
    - *Implementation of MTSS*

34. How would you and/or your colleagues define success for the implementation of supports for Focus Schools?

35. Can you describe some of the specific improvement strategies that you know have been adopted at your school as part of its Focus School designation?

    Probe as necessary:

    - *Implementation of MTSS?*
    - *Are these strategies similar to or different from prior practice?*
    - *Do you know why your school is embarking on these strategies?*
    - *How are the improvement strategies prioritized?*
    - *Does the School Improvement Plan provide the guidance for improvement efforts? Does it provide guidance for closing the achievement gap between the top 30 percent and bottom 30 percent of students?*

36. What has been your role in selecting, developing, or implementing any of these strategies?

37. What kinds of supports or professional development have you received to prepare you for implementing these strategies? Do you feel prepared to implement these strategies?

38. In your view, has your school been successful in reducing achievement gaps within schools? What has been most helpful in reducing your school's within-school achievement gap? What has been least helpful?

    Probe as necessary:

    - *MTSS?*
    - *School size?*
    - *School organization (e.g., departments or clusters, especially in high schools)?*
    - *Professional capacity of staff?*
    - *Professional community?*
    - *Clear vision?*
    - *Data inquiry cycle?*
    - *External support?*
    - *Family / community support?*

39. Is there anything I have not yet asked that you think is important to knowing about identification, measurement of, or supports for Focus Schools?

Appendix B. Focus School Networked Improvement Community: 2015-16 School Year Focus NIC Monthly Meeting Calendar

Table B-1 Monthly Meeting Calendar

Date and Time	Activity	Participants	Location
September 2015	Hold introductory meetings	Regional Educational Laboratory Midwest (REL Midwest) Michigan Department of Education (MDE) MI Excel District Improvement Facilitator (DIF)/Intermediate School District (ISD) Consultant Central Office Representative	• Virtual conference or video call
October 20, 2015 **9:00 a.m.–12:00 p.m.**	Perform root-cause analysis	Focus NIC REL Midwest MDE MI Excel DIF/ISD Consultant Central Office Representative	• Ingham ISD, Room AB 2630 W. Howell Rd. Mason, MI
November 17, 2015 **1:00–4:00 p.m.**	Develop theory of action Define intervention	Focus NIC	• Ingham ISD, Room AB 2630 W. Howell Rd. Mason, MI • West Main Professional Development Center 1627 West Main Street, Kalamazoo, MI
December 3, 2015 **9:00 a.m.–12:00 p.m.**	Define outcome measures	Focus NIC	• Ingham ISD, Room AB 2630 W. Howell Rd. Mason, MI
January 20, 2016 **1:00–3:00 p.m.**	Review baseline outcome data Implement the intervention	Focus NIC	• Calhoun ISD
• **February 19, 2016**	Monitor intervention	Focus NIC	• Calhoun ISD

Date and Time	Activity	Participants	Location
• **1:00–3:00 p.m.**			
• **March 9, 2016** • **1:00–4:00 p.m.**	Measure outcome (posttest)	Focus NIC	• Calhoun ISD
• **April 13, 2016** • **1:00–4:00 p.m.**	Define Plan-Do-Study-Act Cycle 2 activities and measures	Focus NIC	• Virtual meeting
• **May 11, 2016** • **1:00–4:00 p.m.**	Monitor intervention	Focus NIC	• Calhoun ISD
• **June 8, 2016** • **1:00–4:00 p.m.**	Measure outcome Debrief of Continuous Improvement model Determine next steps	Focus NIC	• Calhoun ISD

Appendix C. Michigan Focus Schools Networked Improvement Community Quick Implementation Guide

The Michigan Focus Schools Networked Improvement Community (NIC) designed the following intervention to address the differential attainment of mathematical fluency benchmarks by students in focus schools by achieving the following measurable improvement aim:

All students in the bottom 30 percent in Focus NIC schools will have the opportunity to practice mathematical fluency skills daily.

Intervention Start Date: January 11, 2016
PDSA Cycle 1: January 11–March 11 2016

Table C-1 Intervention Roles and Responsibilities

Role	Responsibility	Time Frame
Students	**Daily practice**	**Daily**
• Teachers	• Plan daily instruction including a minimum of 10 minutes of mathematical fluency; assess and progress monitoring of mathematical fluency skills of students; maintain completion log Ingham will implement RocketMath Kalamazoo will implement workstations	• Daily
• Principals	• Conduct fidelity walkthroughs looking specifically for mathematical fluency instructions and practice time; collect teacher logs of daily practice; ensure teachers are filling logs out.	• Walkthroughs every two weeks • Collection of logs weekly
• Central Office	• Assist with collating mathematical fluency data; collaborate with building-level administrators to provide professional development in mathematical fluency when appropriate.	• Ongoing
• NIC	• Ensure building-level staff have tools and resources necessary to implement the intervention; aligning resources with state expectations.	• Monthly

PDSA Cycle 1: Instructions

Principals should work with teachers to identify the bottom 30 percent of students on their mathematical fluency skills per mid-year MAP scores or AIMSweb scores for the 2015–16 school year.

Principals will provide mathematics teachers in Grades 2–5 (or 6–8) with a list of students in the bottom 30 percent.

Teachers and principals will assess baseline mathematical fluency scores on grade-level benchmarks for each student in the bottom 30 percent.

Teachers will work to implement opportunities for these students every day.

In Ingham focus schools, this opportunity will be provided through RocketMath

In KPS, focus schools will utilize workstations to provide opportunities for students to practice mathematical fluency skills daily.

Teachers will track daily whether students received an opportunity to practice mathematical fluency skills for at least 10 minutes per the log provided in this document.

Principals will collect teacher logs weekly. Principals will conduct a walkthrough in mathematics teachers' classrooms at least every two weeks using a checklist developed by Focus NIC participants. The purpose of this walkthrough will be to ensure students are getting opportunities to practice their mathematical skills in the way that the workstations and RocketMath modules were envisioned.

Central office representatives will try to make mathematical fluency benchmarks the focus of any professional development provided to teachers throughout the remainder of the school year.

Focus NIC participants will monitor students' progress and implementation of the interventions and make adjustments as needed throughout the cycle.

Figure C-1 Math Fluency Practice Daily Log—Template

Instructions: First, enter the dates of interest in the Week column. Each day, complete the log by checking or circling "Yes" if students in the bottom 30 percent had the opportunity to practice mathematical fluency skills that day, or by checking or circling "No" if students in the bottom 30 percent did not have the opportunity to practice mathematical fluency skills that day.

• Week	• MONDAY		• TUESDAY		• WEDNESDAY		• THURSDAY		• FRIDAY	
• Ex: • 1/11–1/15	• Yes	• No	• Yes	• No	• Yes	• No	• Yes	• No	• Yes	• No
•	• Yes	• No	• Yes	• No	• Yes	• No	• Yes	• No	• Yes	• No
•	• Yes	• No	• Yes	• No	• Yes	• No	• Yes	• No	• Yes	• No
•	• Yes	• No	• Yes	• No	• Yes	• No	• Yes	• No	• Yes	• No

Figure C-2 Mathematics Fluency Principal Observation Protocol

Date	
Teacher	
Grade level	
School	
Substitute: yes/no	

Core (Tier 1) vs. intervention (Tier 2)	
Length of observation	
Length of mathematics fluency work	
Percentage of Focus students observed	

Activity	**Observation notes**	**Implementation score** (Circle one)
Students are engaged in mathematics fluency skill building.		<50% engagement = 0 50–75% engagement = 1 >75% engagement = 2
Students have the necessary materials.		<50% engagement = 0 50–75% engagement = 1 >75% engagement = 2
Students exhibit routines and procedures regarding work and transitions.		<50% engagement = 0 50–75% engagement = 1 >75% engagement = 2
Students practice mathematics fluency for at least 10 minutes.		<50% engagement = 0 50–75% engagement = 1 >75% engagement = 2
Students can articulate the learning objective.		Not acceptable = 0 Acceptable variation = 1 Fully implementing = 2
Students receive corrections or descriptive feedback.		Not acceptable = 0 Acceptable variation = 1 Fully implementing = 2

Table C-2 Mathematics Fluency Principal Observation Rubric Exemplars

Activity	Fully Implemented	Acceptable	Not Acceptable
Students can articulate the learning objective.	Students can identify which aspect of fluency they are trying to improve and how it will help them solve more complex mathematics problems.	Students can identify which aspect of fluency they are trying to improve.	Students cannot identify which aspect of fluency they are trying to improve.
Students receive corrections or descriptive feedback.	All students receive actionable: • Corrections that identify a more efficient strategy to build fluency. • Feedback that acknowledges the strategy they used to solve the problem. • Feedback that identifies a more challenging strategy.	All students receive corrections or feedback, and a few students receive actionable corrections and feedback.	Students do not receive corrections or feedback.

Appendix D: Michigan Focus Schools Networked Improvement Community: Data Review Through May 1

The Michigan Focus Schools Networked Improvement Community (NIC) has decided to concentrate its efforts on improving students' mathematics fluency skills in spring 2016 through two Plan-Do-Study-Act (PDSA) cycles. As part of the effort to monitor implementation, principals have collected data on program activities and program outputs in the theory of action developed in fall 2015.

Table D-1 Michigan Focus NIC Theory of Action

Theory of Action			
Program Inputs	**Program Activities**	**Program Outputs**	**Outcomes**
Teacher logs to track daily math practice of fluency skills Implementation guide developed by Focus NIC Observation protocol developed by Focus NIC Principal guidance, coaching, and support to mathematics teachers RocketMath kits (Ingham) or workstations (Kalamazoo) District mathematics coach District and ISD-level mathematics fluency professional development and support	Identify bottom 30% of students Teachers track Focus students' ability to practice mathematics fluency skills for at least 15 minutes every day using daily logs Bimonthly walk-throughs using observation protocol Ongoing coaching and data use Daily teacher logs	Increased time for students spent on practicing mathematics fluency skills Increased time spent discussing mathematics fluency between teachers and between teachers and principal Increased mathematics fluency emphasis	Increased percentage of all students mastering mathematics fluency benchmarks by May 2016 Improved mathematics fluency of the bottom 30% of students specifically
• **Program Targets:** mathematics teachers in Ingham ISD and KRESA who teach in Focus schools participating in the NIC; all students in mathematics classrooms in Focus schools participating in the NIC, with an emphasis on the bottom 30 percent of students.			
• **Program Goal:** All students will master fluency benchmarks by demonstrating appropriate strategies and recalling facts.			

Figure D-1: Michigan Focus NIC: Primary Driver Diagram

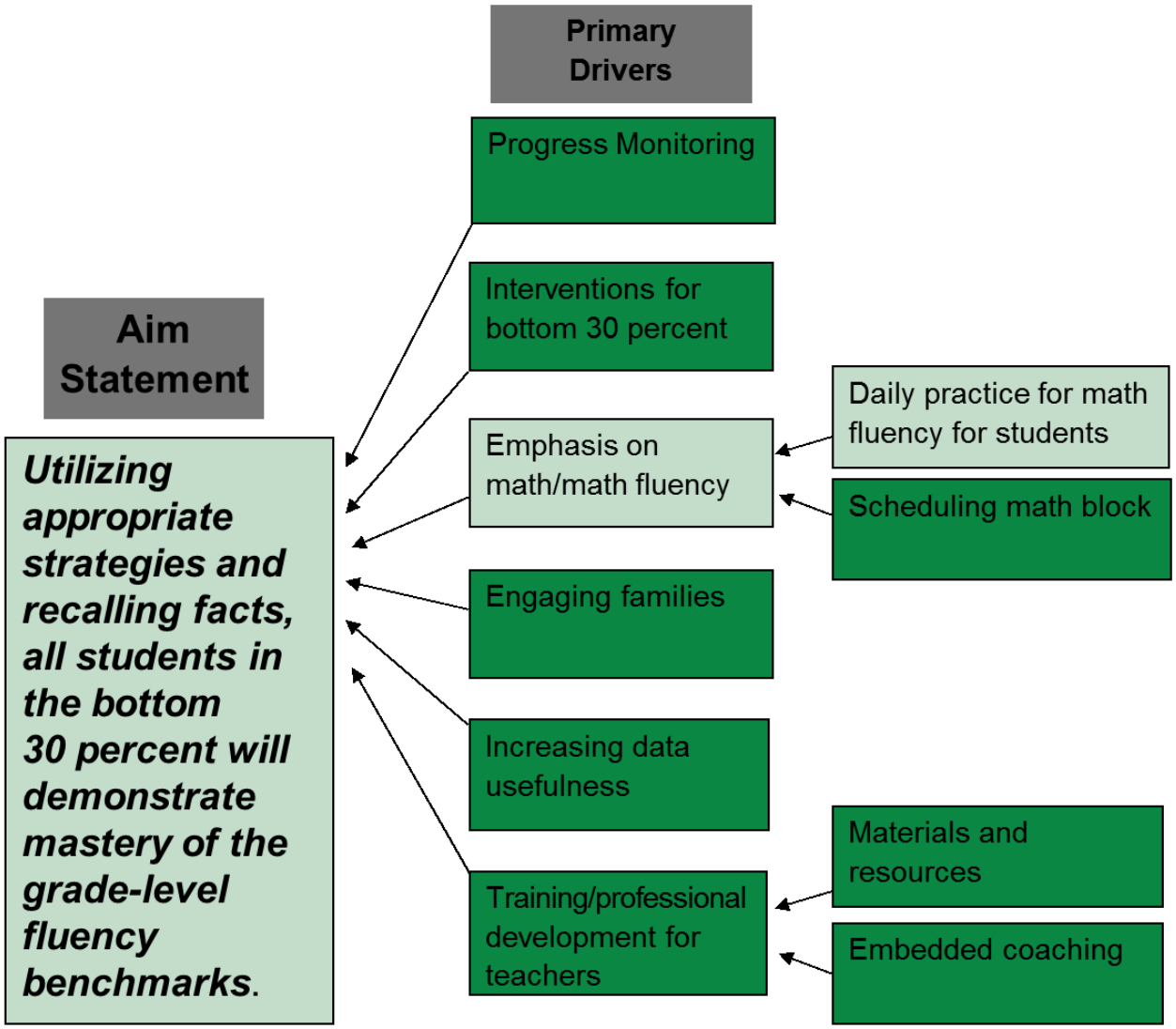


Figure D-2: School 1 - Student Performance on MAP

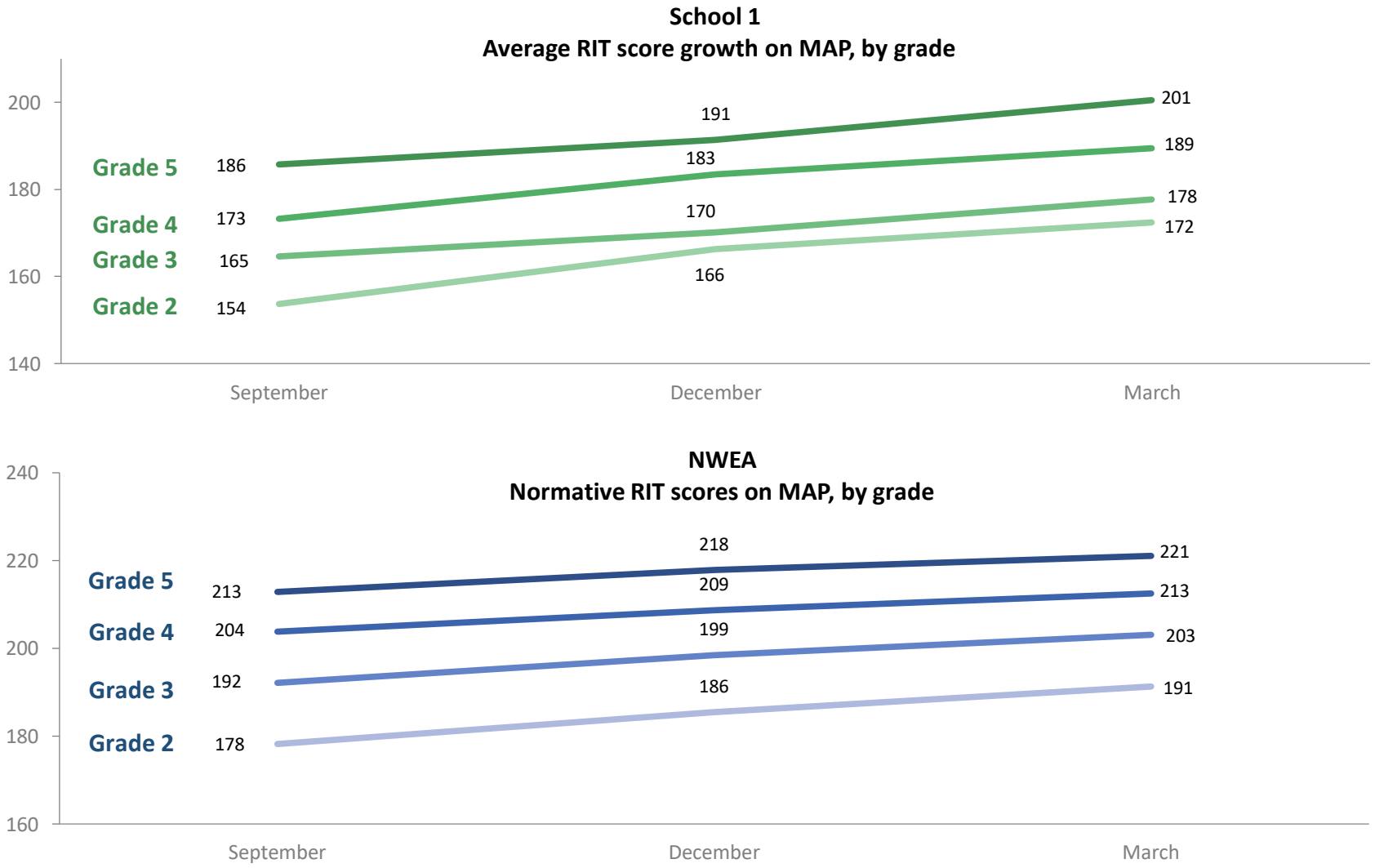
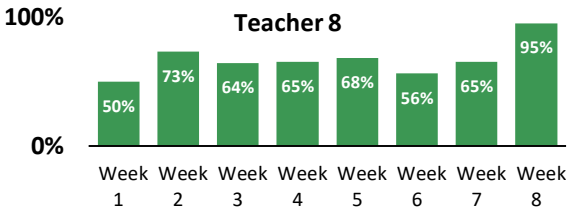


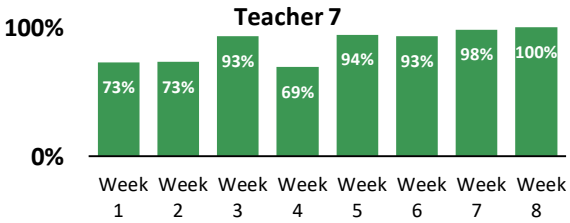
Figure D-3: School 1 - Teacher Daily Logs

Teacher logs were submitted for 11 different classrooms at School 1 Elementary from January 11, 2016 (Week 1), to February 29, 2016 (Week 8). The following figures present the average weekly percentage of Focus students who received mathematics fluency instruction during the data collection period, broken out by classroom.

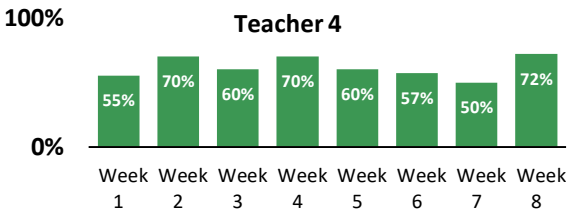
Grade 5

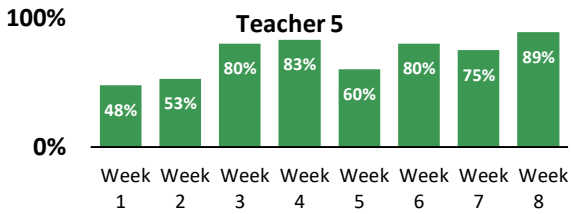


Grade 4



Grade 3





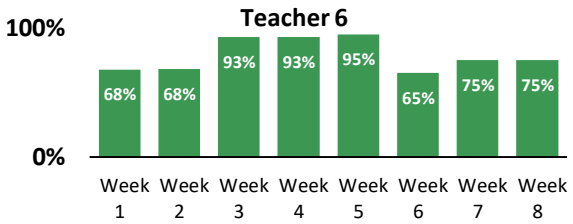
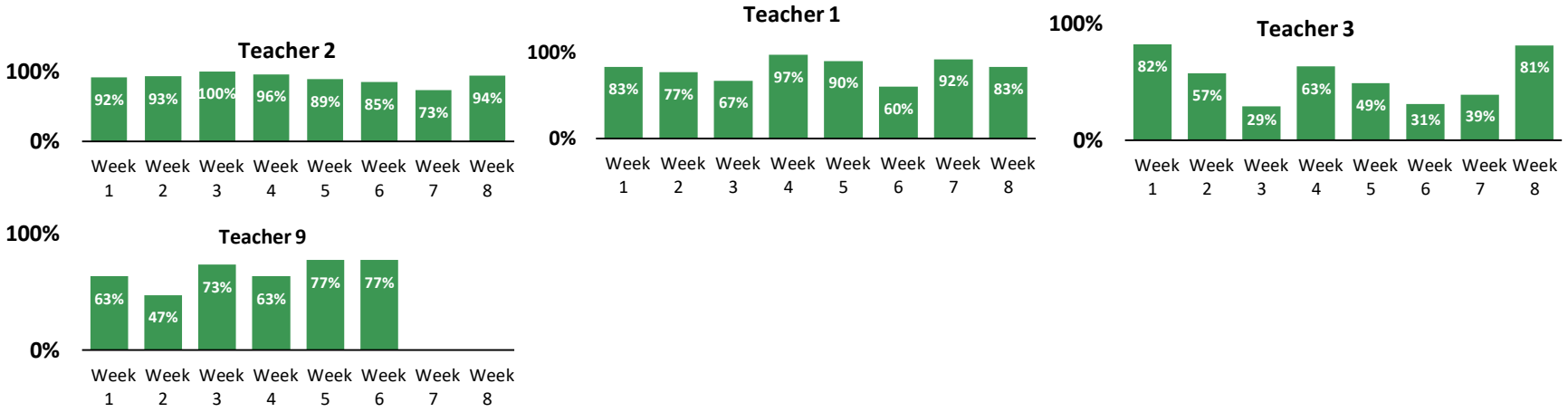


Figure D-3: School 1 - Teacher Daily Logs (continued)

Grade 2



School 1: Principal Observation Summary

Eleven classroom observation protocols were submitted from February 2, 2016, to February 28, 2016. Based on their observations, principals provided ratings of 0–2 for several domains of teachers' instruction, with 0 representing unacceptable practice or low student engagement, 1 representing acceptable practice or moderate student engagement, and 2 representing full implementation or high student engagement. The following table presents a summary of this data, as well as the average length of practice and the percent of Focus students observed.

Table D-2: Summary of mathematics fluency principal observation protocols (*N* = 11), School 1

Protocol information	Average
Length of observation (in minutes)	12.50 (n = 11)
Length of mathematics fluency work (in minutes)	10.5 (n = 10)
Percentage of focus students observed	80.77 (n = 11)
Students engaged in mathematics fluency	1.75 (n = 11)
Students have the necessary materials	2.00 (n = 11)
Students exhibit routines and procedures regarding work and transitions	1.70 (n = 11)
Students practice mathematics fluency for at least 10 minutes	1.65 (n = 11)
Students can articulate learning objective	1.75 (n = 10)
Students receive corrections or descriptive feedback	1.20 (n = 11)

Figure D-4: School 2 - Student Performance on AIMSWEB

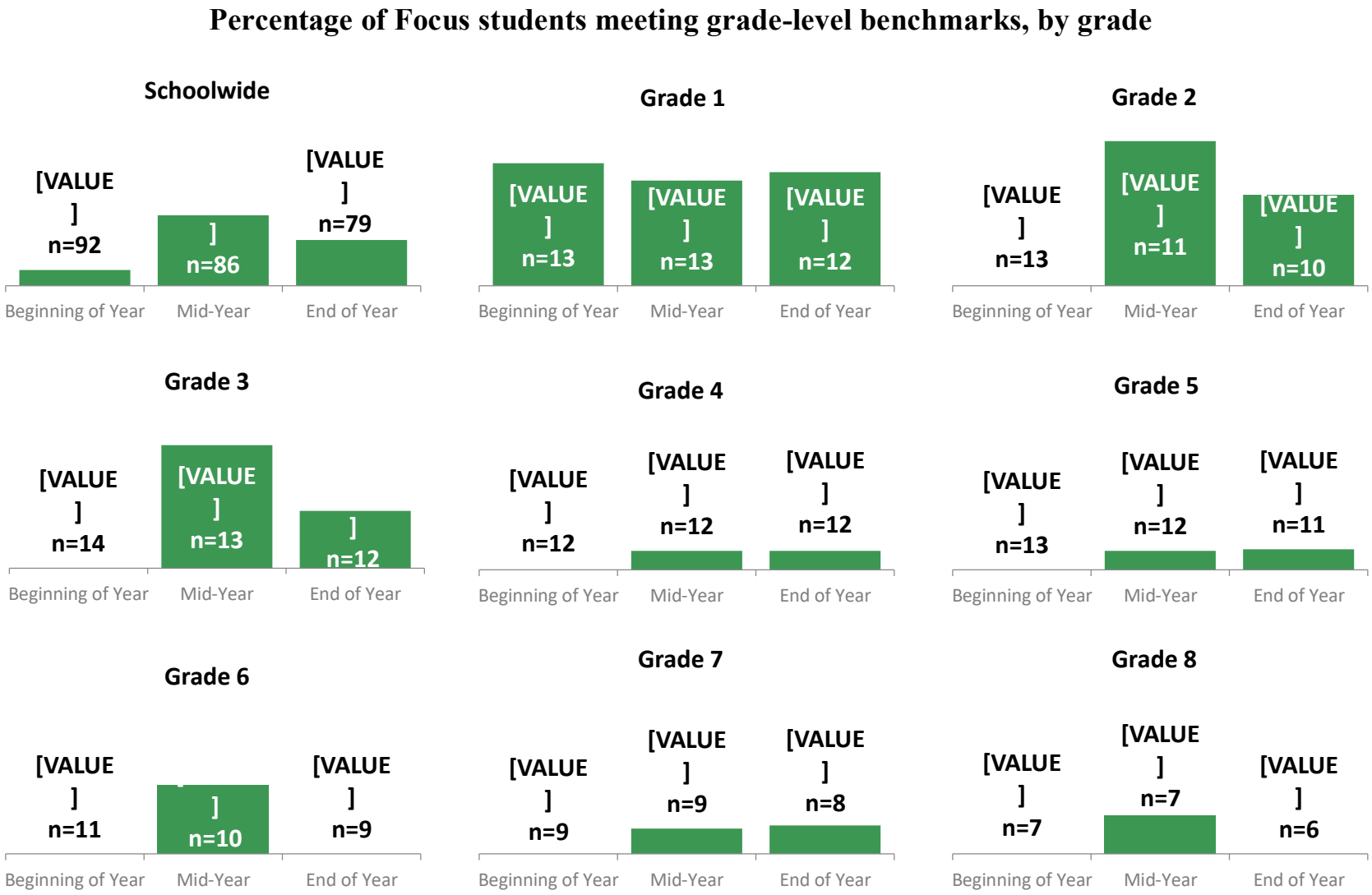


Figure D-5: School 2 -Teacher Daily Logs

Teacher logs were submitted for 13 different teachers at School 2 from February 8, 2016 (Week 5), to March 28, 2016 (Week 12). The following figures present the percentage of days per week Focus students in the class received mathematics fluency instruction, broken out by classroom.

Grades 7, 8

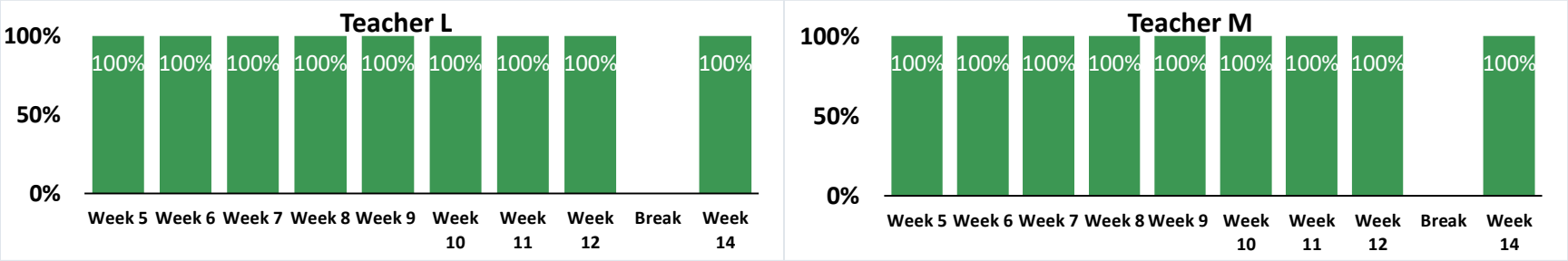


Figure D-5: School 2 -Teacher Daily Logs (continued)

Grades 4, 5, 6

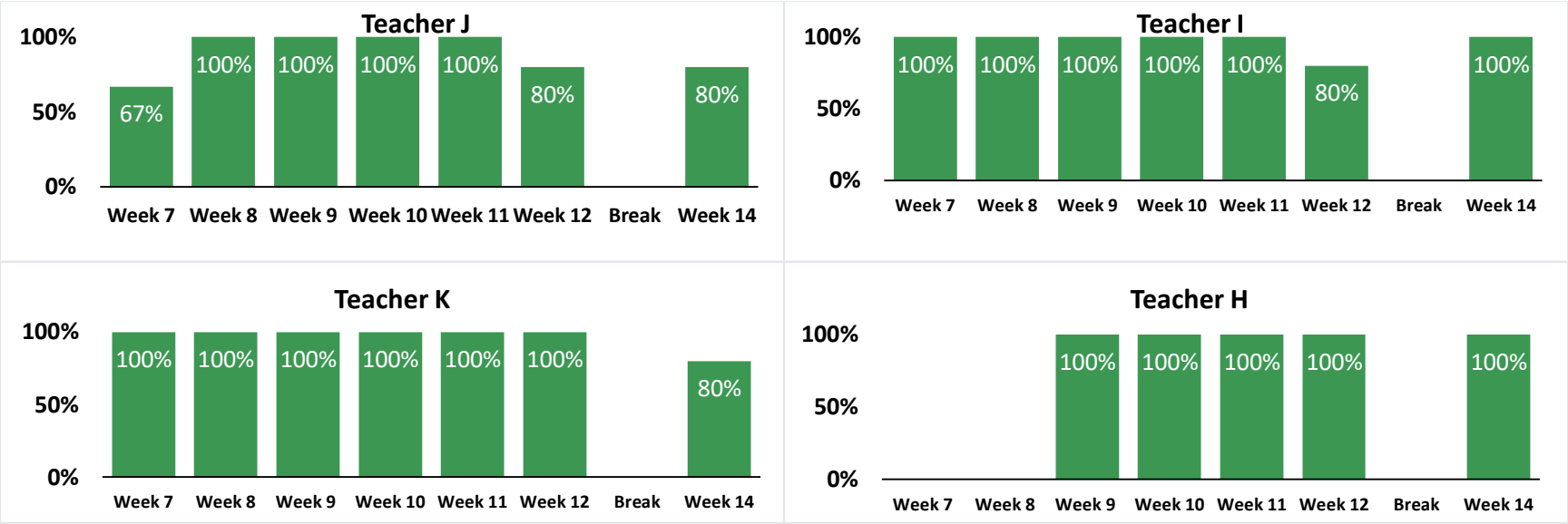
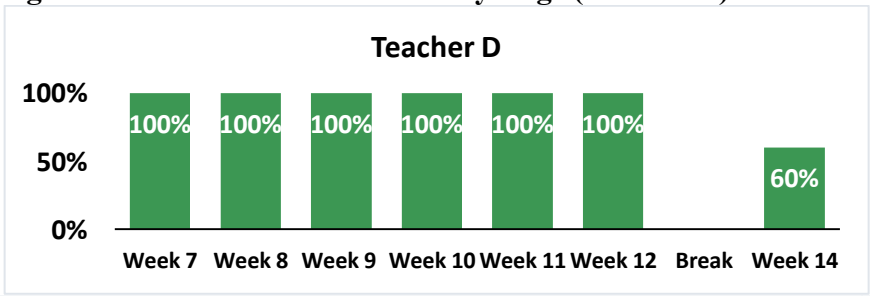
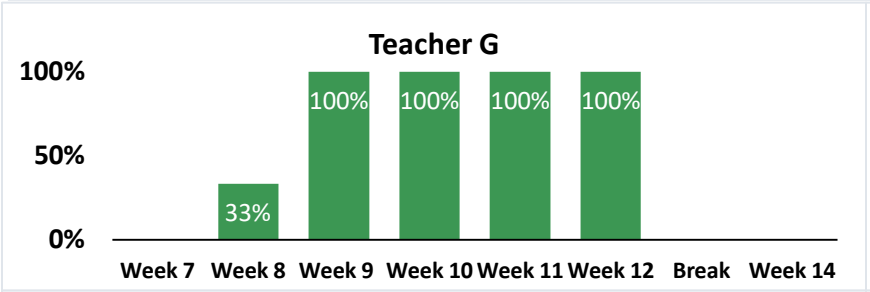
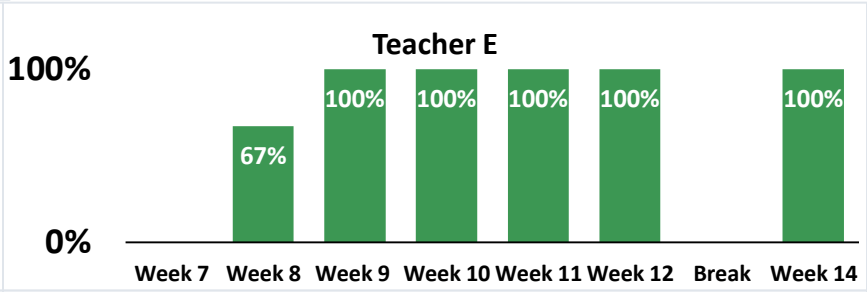
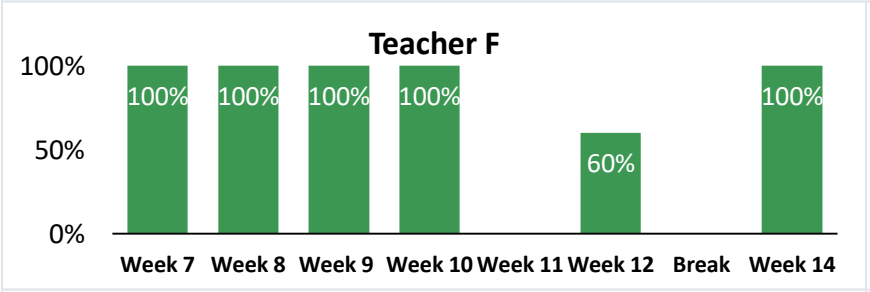


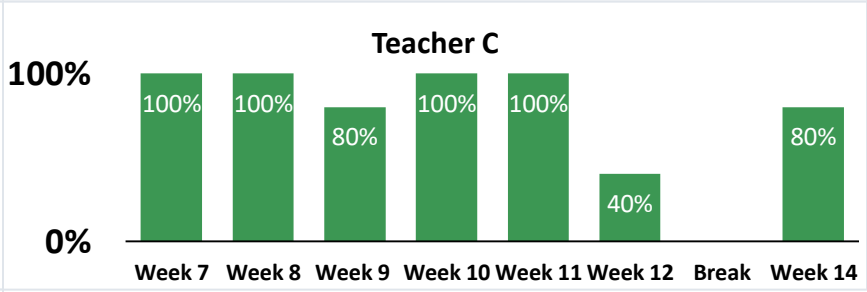
Figure D-5: School 2 -Teacher Daily Logs (continued) Grades 1, 2, 3











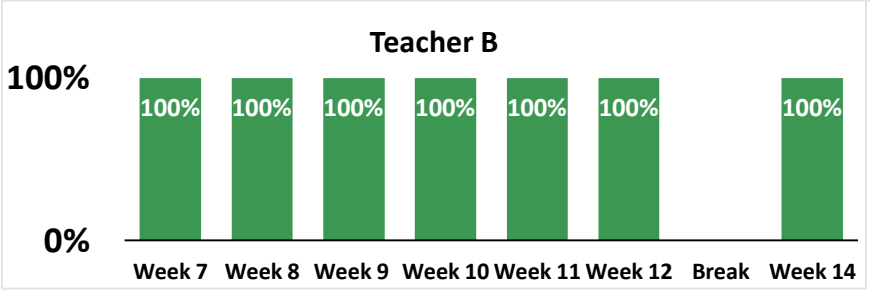
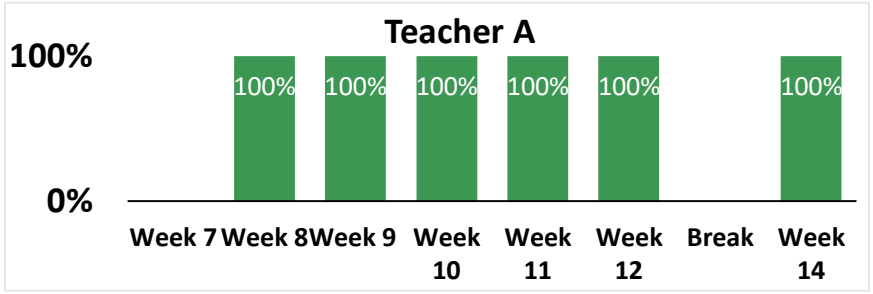


Figure D-5: School 2 -Teacher Daily Logs (continued)

Kindergarten



School 2: Principal Observation Summary

Thirty-one classroom observation protocols were submitted from February 29, 2016, to May 5, 2016. Based on their observations, principals provided ratings of 0–2 for several domains of teachers' instruction, with 0 representing unacceptable practice or low student engagement, 1 representing acceptable practice or moderate student engagement, and 2 representing full implementation or high student engagement. The following table presents a summary of this data, as well as the average length of practice and the percentage of Focus students observed.

Table D-3: Summary of mathematics fluency principal observation protocols (*N* = 31), School 2

Protocol information	Average
• Length of observation (in minutes)	15.77 (n=31)
• Length of mathematics fluency work (in minutes)	15.38 (n=26)
• Percentage of focus students observed	100 (n=21)
• Students engaged in mathematics fluency	1.87 (n=31)
• Students have the necessary materials	1.94 (n=31)
• Students exhibit routines and procedures regarding work and transitions	1.90 (n=29)
• Students practice mathematics fluency for at least 10 minutes	1.96 (n=30)
• Students can articulate learning objective	1.60 (n=30)
• Students receive corrections or descriptive feedback	1.73 (n=30)

Observation Protocol Side-by-Side Comparison

Table D-4: Summary of mathematics fluency principal observation protocols, School 1 and School 2

Protocol information	School 1 average (n = 11)	School 2 average (n = 31)
• Length of observation (in minutes)	• 12.50 (n = 11)	• 15.77 (n = 31)
• Length of mathematics fluency work (in minutes)	• 10.5 (n = 10)	• 15.38 (n = 26)
• Percentage of focus students observed	• 80.77 (n = 11)	• 100 (n = 21)
• Students engaged in mathematics fluency	• 1.75 (n = 11)	• 1.87 (n = 31)
• Students have the necessary materials	• 2.00 (n = 11)	• 1.94 (n = 31)
• Students exhibit routines and procedures regarding work and transitions	• 1.70 (n = 11)	• 1.90 (n = 29)
• Students practice mathematics fluency for at least 10 minutes	• 1.65 (n = 11)	• 1.96 (n = 30)
• Students can articulate learning objective	• 1.75 (n = 10)	• 1.60 (n = 30)
• Students receive corrections or descriptive feedback	• 1.20 (n = 11)	• 1.73 (n = 30)

•

January Agenda: Michigan Focus School Networked Improvement Community: Reviewing Baseline Data

The Focus Networked Improvement Community (NIC) has spent the past few months determining the root causes of inequality in our schools, picking one factor on which to focus, and developing a theory of action and defining outcome measures to change that factor. The group has coalesced on addressing students' mathematics fluency skills, which are essential benchmarks to achieve for mathematics learning and assessment.

The group is focused on the following problem statement:

Students in the bottom 30 percent will achieve mastery on mathematics fluency benchmark assessments in Focus NIC schools by May 2016.

To do so, principals will track teachers' self-reports of opportunities for students in the bottom 30 percent of mathematics assessments to practice their mathematics fluency skills on a daily basis. Teachers will collect a daily log of whether students had an opportunity to practice mathematics fluency skills. Principals will collect teacher logs on a weekly basis and conduct a walk-through with an observational checklist every two weeks. District and intermediate school district supports will be used to emphasize mathematics fluency whenever possible.

The goal for this meeting is to review baseline data, address questions and concerns about implementation of the intervention, and refine and practice with the observation protocol developed by Focus NIC members. The group will also review the long-term timeline for the Focus NIC, including the goal of completing a second Plan-Do-Study-Act cycle in April–May 2016.

Agenda

1:00–1:20 p.m.	**Road Map: Where Are We Going?**
1:20–2:00 p.m.	**Principal Presentations of Baseline Data**
2:00–2:15 p.m.	**Strengths and Challenges of the Intervention**
2:15–2:30 p.m.	**Break**
2:30–3:15 p.m.	**Principal Observation Protocol**
3:15–3:45 p.m.	**What Is Next for Focus Schools?**
3:45–4:00 p.m.	**Next Steps—Next Meeting: February 17, 2016**

Principal Presentations of Baseline Data

Goal: Review each school's baseline data. Each principal will have five to seven minutes to present on their three slides:

Who is participating?

How many students comprised the bottom 30 percent? How were they identified? In which grades?

Which teachers are participating in the intervention? How many? In which grades?

What are students' baseline scores?

On mathematics fluency assessments? On MAP? On AIMSweb?

What are the baseline scores for all students or the bottom 30 percent of students?

How is the implementation going?

What are teachers' perceptions of the intervention? Are they completing their daily logs?

Are principals able to conduct walk-throughs and collect daily logs?

What are other barriers facilitators face with respect to implementation?

Principal Observation Protocol

As part of the implementation of the intervention, principals in the Focus NIC have committed to conducting a walk-through with a five- or six-question observation protocol focused on mathematics fluency instruction and workstations. Matt, Sean, and Dodie will present a draft of the observation protocol with scenarios to practice using it. The community will also have an opportunity to provide feedback and refine the observation tool.

What Is Next for Focus Schools?

Given the passage of the Every Student Succeeds Act (ESSA), the Focus School designation may transition to something else. However, the focus on achievement gaps will remain. Karen Ruple from the Michigan Department of Education (MDE) will lead the group in a discussion of how MDE plans to continue to support Focus Schools and what might be next for schools with achievement gaps.

Next Steps

The next meeting of the NIC will take place on February 17, 2016, at Calhoun Intermediate School District. We will use this time to monitor implementation of the intervention, discuss challenges and facilitators, and share best practices.

February Agenda:
Michigan Focus School Networked Improvement Community: Monitoring Intervention

February 19, 2016

The Focus Networked Improvement Community (NIC) has spent the past four months determining the root causes of inequality in our schools, picking one factor on which to focus, and developing a theory of action and defining outcome measures relevant to that factor. The group has settled its focus on addressing students' math fluency skills, and has developed a math fluency intervention and measures.

Beginning in January 2016, principals and teachers began tracking opportunities for students in the bottom 30 percent of math assessments to practice their math fluency skills on a daily basis.

- Teachers began collecting a daily log of whether students had an opportunity to practice math fluency skills.
- Principals began collecting teacher logs on a weekly basis.
- Principals began conducting walk-throughs with an observational checklist every two weeks.

This virtual meeting serves as a "check-in" at the halfway point of the implementation of the mathematics fluency intervention and data collection activities.

Agenda

1:00–1:10 p.m.	**Recap of January Meeting and MI Excel Quarterly Focus Meeting**
1:10–1:45 p.m.	**Principal Updates on Implementation**
1:45–2:10 p.m.	**Strengths and Challenges of the Intervention**
2:10–2:30 p.m.	**Discussion of Math Fluency Resources**

Principal Updates on Implementation, Strengths & Challenges

Each principal will have 10 minutes to present updates from their school:

Who is participating?

- How many students were in the bottom 30 percent? How were they identified? In which grades?
- Which teachers are participating in the intervention? How many? In which grades?

What are students' baseline scores?

- On mathematics fluency assessments? On MAP? On AIMSweb?

What are the baseline scores for all students? For the bottom 30 percent of students?

How is the implementation going?

What are teachers' perceptions of the intervention? Are they completing their daily logs?

Are principals able to conduct walk-throughs and collect daily logs?

What are other barriers facilitators face with respect to implementation?

What have been the strengths of intervention implementation between January and now? What have been the challenges?

Discussion of Math Fluency Resources and Tools

NIC members will share resources (e.g., program, practice, tool) that have been useful in supporting the bottom 30 percent of students in their district or ISD.

What is the resource?

How has this resource been used to support students in the bottom 30 percent?

Next Steps

The next meeting of the NIC will take place on March 9 at Calhoun ISD from 1:00 – 4:00 p.m. EST. This meeting will focus on measuring outcomes of the first Plan-Do-Study-Act cycle, as well as on discussing actions to begin the second cycle.

March Agenda:
Michigan Focus School Networked Improvement Community: Measure Outcome

March 9, 2016

The Focus Networked Improvement Community (NIC) has spent the past few months determining the root causes of inequality in our schools, picking one factor on which to focus, and developing a theory of action and defining outcome measures to change that factor. The group has coalesced on addressing students' mathematics fluency skills with an intervention of daily practice for students in the bottom 30 percent. To measure implementation of this intervention, principals have been collecting teachers' daily logs that indicate whether students had the opportunity for mathematics fluency practice and conducting walk-throughs with an observational checklist every two weeks.

The goal for this meeting is investigate the data collected during cycle 1 of the Plan-Do-Study-Act process to draw conclusions about implementation so far and to determine next steps for the NIC.

Agenda

1:00–2:00 p.m.	**Activity 1: Collaborative Interpretation of Data**
2:00–2:10 p.m.	**Break**
2:10–2:40 p.m.	**Updates**
2:40–3:20 p.m.	**Activity 2: Focused Conversation**
3:20–3:50 p.m.	**Activity 3: What Is Next for the Michigan Focus NIC?**
3:50–4:00 p.m.	**Next Steps—Next Meeting April 13, 2016**

Activity 1: Collaborative Interpretation of Data

Goal: Through collaborative review, make meaning of the data presented in visual graphs and charts to tell a story about the implementation of mathematics fluency interventions in Focus schools.

Step 1

Break into pairs or small groups. Each group will receive a graph or chart.

Take a few minutes to review the graph to become familiar with format and content. Begin mentally noting any observations about the data.

Work in a small group to construct thematic statements drawn from your observations of the data and record them on sticky notes.

Step 2

Work as a large group to combine the thematic statements and summarize similar statements to construct key findings. Record the key findings on another sticky note.

Discuss the key findings that emerge and how they relate to the individual visuals studied by each group.

Activity 2: Focused Conversation

Goal: Process the information from Activity 1 and derive conclusions by discussing and answering a series of questions based on four levels of thinking: objective, reflective, interpretive, and decisional.

As a large group, take turns to discuss the following questions in order:

Objective—data and sensory observation

Looking at what was presented in Activity 1 and the materials in hand, what caught your attention?

Reflective— personal reactions and associations

In what areas are you really clear? In what areas are you confused or frustrated?

Interpretive—meaning, significance, and implications

What new vantage point has this given you about your work?

Decisional—resolution

What will you do differently as a result of this exercise?

Activity 3: What Is Next for the Michigan Focus NIC?

Goal: Determine additional topics to direct the work of upcoming NIC meetings.

As a group, brainstorm and discuss ideas for topics and resources that would be useful for future NIC meetings, such as sharing resources focused on mathematics fluency for specific subgroups.

Record the ideas on a list, and then vote using sticker dots to determine which ideas to prioritize for the next meeting.

Next Steps

The next meeting of the NIC will take place on April 13, 2016, and it will be held virtually.

April Agenda:
Michigan Focus School Networked Improvement Community: Monitor Intervention

April 13, 2016

The Focus Networked Improvement Community (NIC) has spent the past few months determining the root causes of inequality in our schools, picking one factor on which to focus, and developing a theory of action and defining outcome measures to change that factor. The group has coalesced on addressing students' mathematics fluency skills with an intervention of daily practice for students in the bottom 30 percent of mathematics achievement. To measure implementation of this intervention, principals have been collecting teachers' daily logs that indicate whether students had the opportunity for mathematics fluency practice and conducting walk-throughs with an observational checklist every two weeks.

The goal for this meeting is to continue monitoring the math fluency intervention by investigating the student test data and process data collected during the first three months of implementation.

Agenda

1:00–1:20 p.m.	**Updates on Implementation**
1:20–2:10 p.m.	**Data Presentation & Intervention Examination**
2:10–2:25 p.m.	**Data Tool Demonstration**
2:25–2:30 p.m.	**Next Steps—Next Meeting May 11, 2016**

Updates on Implementation

Principals provide updates from implementation, including results from using the new data collection protocols revised during the March 9 NIC meeting.

Data Presentation & Intervention Examination

As a group, examine each school's data:

Outcome data: Changes in student assessment scores pre- and post- intervention

Process data: Data from teacher daily logs and principal observations

Discuss:

Are there any connections we can draw between the process data and the outcome data?

Are the observed changes in student test scores what we might expect given the theory of action created by the group in January?

Are any changes to the intervention necessary based on observations from the data presentation and the theory of action?

Data Tool Demonstration

A member of REL Midwest will demonstrate the functions of a data tool template designed to help principals track and visualize data related to the Focus Schools intervention.

Next Steps

The next meeting of the NIC will take place on May 11, 2016 at Calhoun ISD.

May Agenda
Michigan Focus School Networked Improvement Community: Measure Outcome

May 11, 2016

The Michigan Focus School Networked Improvement Community (NIC) has spent the past few months determining the root causes of inequality in our Focus schools, picking one actionable and high-leverage factor on which to concentrate intervention efforts, developing a theory of action, and defining outcome measures to change that factor. The group has coalesced on addressing students' mathematics fluency skills with an intervention of daily practice for students in the bottom 30 percent of mathematics achievement. To measure implementation of this intervention, principals have been collecting teachers' daily logs that indicate whether students had the opportunity for mathematics fluency practice, as well as conducting walk-throughs with an observational checklist every two weeks.

The goal for this meeting is to continue monitoring the mathematics fluency intervention by investigating the student test and process data collected during the first four months of implementation, revisiting the driver diagram developed in the December NIC meeting, and beginning to reflect on the NIC process with an eye toward sustainability.

Agenda

1:00–1:15 p.m.	**Updates on implementation**
1:15–2:15 p.m.	**Data presentation and intervention examination**
2:15–2:45 p.m.	**Driver diagram redux**
2:45- 3:00 p.m.	**Break**
3:00- 3:30 p.m.	**Reflection activity**
3:30–4:00 p.m.	**Next steps; next meeting June 8, 2016**

Updates on Implementation

Principals provide updates from implementation in their schools.

Data Presentation and Intervention Examination

Examine the following data from each school:

Outcome data: changes in student assessment scores pre- and post-intervention

Process data: data from teacher daily logs and principal observations

Discuss the following:

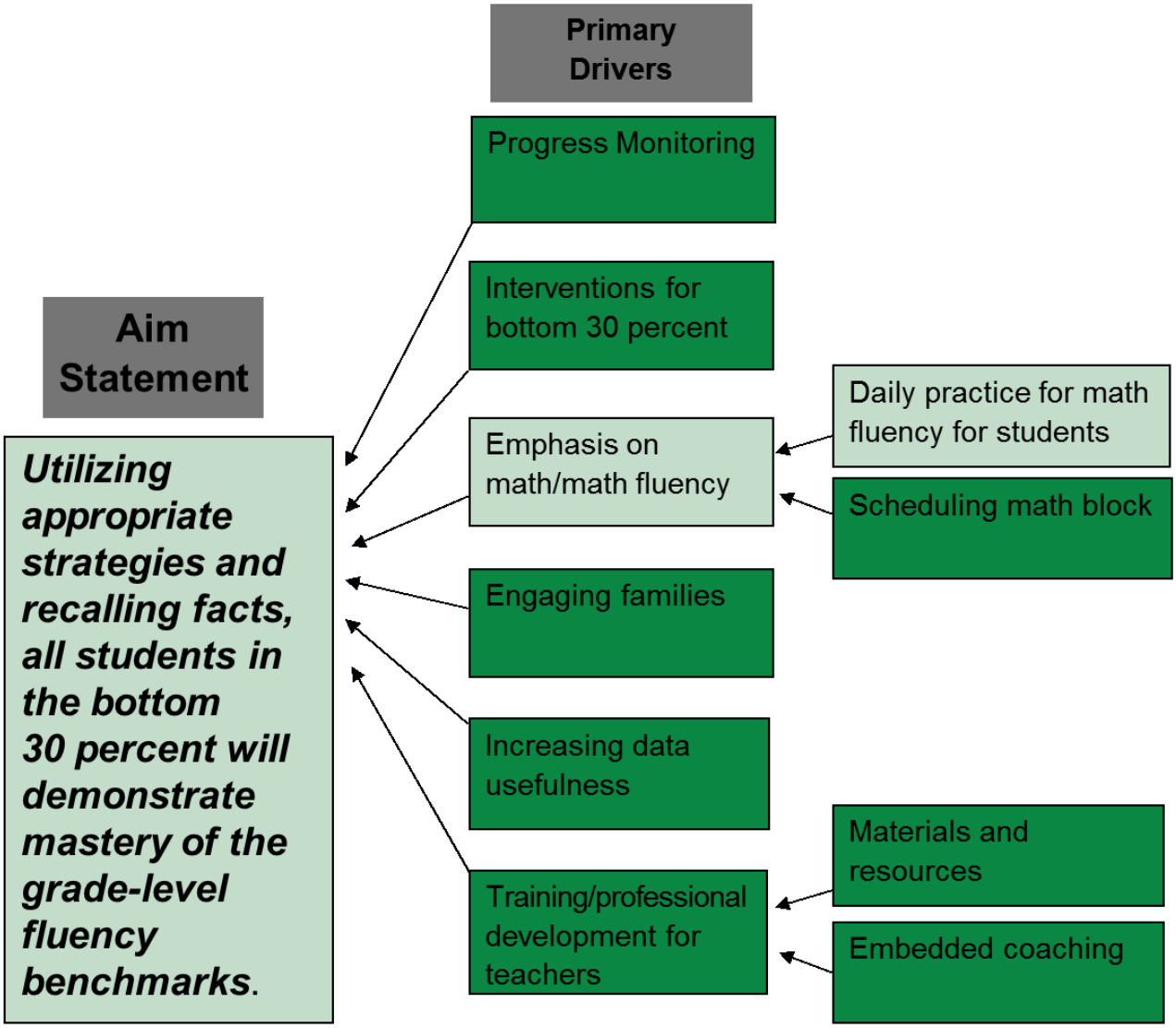
Are there any connections we can draw between the process data and the outcome data?

Are the observed changes in student test scores what we might expect given the theory of action created by the group in January?

Are any changes to the intervention necessary based on observations from the data presentation and the theory of action?

Driver Diagram Redux

Examine the driver diagram created in December. Discuss whether any revisions are necessary given the implementation of the mathematics fluency intervention from January through May 2016 and your review of process and student outcome data.



Reflection Activity

Take a few minutes to reflect on your NIC participation this year before answering the following questions. Then, discuss your responses aloud as a group.

As a result of the intervention, I expect to see…

I would be surprised if…

The part of this process that has been most helpful is…

If we did this again, I would…

If I had to explain the work of an NIC to my counterpart in another school, district, or state, I would highlight…

One thing that would help me continue this work next year is…

Next Steps

The next NIC meeting will take place June 8, 2016, at Calhoun ISD.

June Agenda:
Michigan Focus School Networked Improvement Community: Reflect on NIC

June 8, 2016

The Focus Networked Improvement Community (NIC) has spent the 2015-2016 school year determining the root causes of inequality in our schools, picking one factor on which to focus, and developing a theory of action and defining outcome measures to change that factor. The group has coalesced on addressing students' mathematics fluency skills with an intervention of daily practice for students in the bottom 30 percent of mathematics achievement. To measure implementation of this intervention, principals have been collecting teachers' daily logs that indicate whether students had the opportunity for mathematics fluency practice and conducting walk-throughs with an observational checklist every two weeks.

The goal for this meeting is to reflect on the body of work undertaken by the NIC throughout the 2015–2016 school year, and consider actions and supports necessary to sustain the group's work in the future.

Agenda

1:00–1:15 p.m.	**Updates on Implementation**
1:15–2:00 p.m.	**Case Study #1**
2:00–2:40 p.m.	**Case Study #2**
2:40–3:00 p.m.	**Reflection and Next Steps**

Updates on Implementation

Principals provide updates on implementation from the last month of the 2015-2016 school year.

Case Study #1:

"More than a Network: Building Professional Communities for Educational Improvement"
Jonathan R. Dolle, Louis M. Gomez, Jennifer Lin Russell, and Anthony S. Bryk

1. With a partner, read and review the section of the case study assigned to you. (5-7 minutes)

Discuss the questions below with your partner and jot down notes on your answers (10 minutes).

With your partner, share your summary and observations with the larger group. (15-20 minutes).

As a group, how comfortable do you feel using the language of NICs? What aspects do you think you will retain in your practice? What aspects would you like to incorporate next year?

• **What did the NIC in Case Study #1 do?**	• **How does the case study compare to the MI Focus Schools NIC?**	• **What could be improved?** **In the case study?** **In the Michigan Focus NIC?**

Case Study #2:

"Generating a Networked Improvement Community to Improve Secondary Mathematics Teacher Preparation: Network Leadership, Organization, and Operation"
W. Gary Martin and Howard Goldstein

2. With a partner, read and review the section of the case study assigned to you. (5-7 minutes)

Discuss the questions below with your partner and jot down notes on your answers (10 minutes).

With your partner, share your summary and observations with the larger group. (15-20 minutes).

As a group, how comfortable do you feel using the language of NICs? What aspects do you think you will retain in your practice? What aspects would you like to incorporate next year?

What did the NIC in Case Study #2 do?	**What would this work look like in Michigan? What would be needed?**	**What are some pros about this work? What are some cons?**

Reflection and Next Steps

Reflect on the work of the NIC this year by discussing the following questions:

- What did you learn?
- What will you continue in your practice?
- What would you like to see to continue this work in the future?