Collaborative Improvement in Educational Systems: Exploring Adaptation in Education Reform

by

Maggie Hannan

Bachelor of Arts, Duquesne University, 2008

Master of Arts, West Virginia University, 2010

Submitted to the Graduate Faculty of

the School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Maggie Hannan

It was defended on

November 30, 2018

and approved by

Rip Correnti, Associate Professor, Learning Sciences and Policy

Mary Kay Stein, Professor, Learning Sciences and Policy

Louise Comfort, Professor, Graduate School of Public and International Affairs

Dissertation Chair: Jennifer Lin Russell, Associate Professor, Learning Sciences and Policy

Collaborative Improvement in Educational Systems: Exploring Adaptation in Education Reform

Maggie Hannan, PhD

University of Pittsburgh, 2018

Educational organizations are structurally unique and have a complex relationship to change (e.g., Coburn 2004), and implementation research has long grappled with this fact in efforts to create sustainable, scalable educational improvements (e.g., Spillane et al. 2002). Accordingly, as standards-based reforms become increasingly popular and educational inequities remain entrenched, researchers, funders, and policy makers have become increasingly concerned with the ways in which interventions will scale across highly variable organizational contexts. A variety of improvement strategies have emerged in response to these concerns, and several particularly promising approaches confront the variation and complexity of educational organizations head-on, making those system characteristics key objects for exploration, analysis, and interventions. One such strategy is collaborative problem solving research (Penuel et al. 2017), which tackles many perspectives across different organizational levels, and highlights the structural components of creating and sustaining educational improvement. This dissertation builds on Coburn and colleagues' (2012) "theory of organizational embeddedness" by exploring the theoretical foundations of key change mechanisms in educational systems and investigating the work of two collaborative problem solving research efforts. Together, these three papers examine how change and reform play out in and across educational organizations, a process which is driven by the dynamic interaction between local conditions and the intervention or innovation embedded in

those conditions. In so doing, these studies also explore the relationship between the macro-level structural conditions that shape policy implementation and, simultaneously, the micro-processes of uptake and adaptation. Taken together, these three papers demonstrate the organizational complexity that shapes education implementation research, and suggest several different concepts, perspectives, and strategies that can support researchers and policy-makers as they endeavor to create system-wide educational improvement.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Educational organizations are structurally unique and have a complex relationship to instructional change (e.g. Coburn, 2004), and implementation research has often grappled with this in efforts to understand sustainable, scalable educational improvements (e.g., Spillane, 2002). Accordingly, as standards-based reforms become increasingly popular, and familiar educational inequities remain entrenched despite years of attempts to attenuate them, researchers, funders, and policy makers have become increasingly concerned with the ways in which interventions will scale across highly variable organizational contexts. A variety of research and improvement strategies have emerged in response to these concerns, and several particularly promising approaches confront the variation and complexity of educational organizations head-on, making those system characteristics key objects for exploration, analysis, and interventions.

One such strategy is collaborative problem solving research, a term coined by Penuel and colleagues (2017). These approaches take many forms, which include, but are not limited to research-practice partnerships (RPPs), design-based implementation research (DBIR), community based design research (CBDR), strategic education research partnerships (SERPs), and networked improvement communities (NICs) (Coburn & Penuel, 2016; Penuel et al. 2017; Cohen-Vogel et al. 2015; Tichnor-Wagner et al. 2018; Bryk et al. 2015; Russell et al. 2017; in press; Hannan et al. 2015; Lewis 2015). Penuel and colleagues (2017) explain that this family of approaches share three commitments: to highlighting the perspectives of many stakeholders, to solving problems related to equity, and to using systematic inquiry methods to explore problems and solutions. These approaches acknowledge the complex and multi-faceted nature of implementation and reform

processes in organizations, building on foundational, cooperative infrastructures and routines to enable innovative engagement with entrenched educational problems.

By tackling many perspectives across different organizational levels, collaborative problem solving research also enables a system perspective; in making the voices of many stakeholders observable and intelligible for research, this approach to research can highlight the structural components of creating and sustaining change. Systems theorists argue that taking this kind of perspective is critical to understanding change processes, particularly since change typically takes place in the meso-level of a system, and is largely driven by a few influential actors (Fligstein & McAdam, 2010; Holland, 2014; Miller & Page, 2007). However, understanding those mechanisms and actors requires understanding the broader system structure itself (e.g. Opfer & Pedder, 2011), and taking an expansive view of change processes by thinking across the many levels, interactions, and configurations that make up a system. By explicitly attending to contextual factors, collaborative problem solving research turns our attention to the shape and form of the system in which any educational improvement process is embedded.

This balance between the macro and micro is an important condition for understanding how change and reform play out in and across educational organizations, a process which is driven by the dynamic interaction between local contextual conditions and the intervention or innovation playing out under those conditions. In research on teachers' social networks, Coburn (2012) and colleagues have argued for a "theory of organizational embeddedness" (p. 313) to account for the multidirectional flow of resources and relationships taking place in school environments that effect teacher learning and improvement, a concept that these studies expand to other key actors in educational systems. This idea provides a helpful framing concept for educational implementation

research, in that it presumes that interventions and innovations exist in dialogic interaction with the system structures that surround them.

Furthermore, studying certain kinds of educational innovations and interventions, particularly those that provide the most insight into the organizational boundaries where change takes place, is an important part of building theory about these dynamics. In this dissertation, I examine two state-led reform initiatives that explore the dynamics of novel approaches to supporting implementation and improvement in complex systems at scale. These large-scale, macro-level projects are fertile ground for this kind of analytical work for two reasons; first, because they consider the macro-level of the system by design, in their aim to create scalable change; second, because they attend to the flexibility that is required to support implementation across many different local contexts. Collaborative educational improvement efforts like these respond to important questions about place-based implementation, local conditions that support successful reforms, and strategies for fostering improved teaching and learning reliably across many different contexts (Cohen-Vogel et al. 2015; Honig, 2006; Mclaughlin, 1991; Spillane et al., 2002).

These state level reform initiatives represent two approaches to collaborative problem solving research; the first was a large research practice partnership (RPP) that contained a network of instructional coaches who collaborated with researchers and professional developers from a university, as well as officials from a state department of education, and the second was a networked improvement community (NIC) led by a state department of education with support from the Carnegie Foundation for the Advancement of Teaching. I found that these contexts offered robust empirical opportunities to explore the mechanisms and interactions that can drive change and improvement across variable school districts. In the case of instructional coaches, their

work is entirely located in boundary spaces in educational organizations, and their task is generally to create targeted change at the system level by working with individual teachers. Studying a network of instructional coaches revealed important insights about the many different ways that coaching is implemented and supported across sites, responding to an important gap in the coaching literature about cross-context studies of coaching interventions (Blazar & Kraft, 2015; Kraft et al. 2016).

Additionally, NICs engage the concepts of complex systems, networks, and embeddedness directly, and so provide important learning opportunities for researchers to build a deeper understanding of how to leverage organizational system dynamics in education. NICs provide a social structure that organizes and accelerates learning about how to collaboratively improve in education (Russell et al, 2017). Designed to be scientific learning communities, NICs provide a collaborative infrastructure that connects educators, researchers, and designers around solving a high leverage problem (Bryk, Gomez, & Grunow, 2011; Bryk et al. 2015). NICs are built around four essential characteristics: they focus on a specific, common aim; they draw guidance from deep understandings of a problem, the system that produces it, and a common working theory for how to address it; they use improvement science to discipline their work; and they accelerate the design, testing, refinement, and diffusion of innovations through intentional coordinating mechanisms (Bryk et al. 2015; Russell et al. 2017).

This improvement science approach is particularly significant in that it responds to the need for disciplined and contextualized knowledge building in the education field. Improvement science tools, particularly Plan-Do-Study-Act (PDSA) cycles, use the power of the scientific method to structure small, rapid experiments that allow users to test and refine adaptations to their practice (Barab & Squire, 2004; Cohen-Vogel et al. 2015; Langley et al. 2009; Morris & Hiebert, 2011).

This approach to the uptake of new ideas fosters important learning about contextualized implementation across many local sites, which, when connected by a network structure, contributes to learning about taking an innovation to scale (Bryk et al. 2015; Cohen-Vogel et al. 2015; Russell et al. 2017). The Carnegie Foundation for the Advancement of Teaching has led the development and refinement of NICs in education for nearly a decade, and this particular collaborative approach to educational problem solving is beginning to show promising results. The NIC that I explore here illuminates important insights about using this networked approach in educational organizations to create improvement across diverse organizational contexts.

The conceptual framing of this dissertation expands on this idea of organizational embeddedness in complex systems by drawing together ideas from policy implementation, coaching, systems theory, continuous improvement, and networks research to build theory about the dynamic processes of uptake and improvement in educational organizations. I am particularly concerned with exploring the interactions between the structural conditions that shape policy implementation and uptake and adaptation. In this introduction, I briefly discuss the relevant literature that forms this overarching conceptual framework, and foreshadow how each of my three papers grapples with these concepts.

1.1 Implementation in Variable Complex Systems

These three papers take up a variety of concepts from the complex systems literature from different levels and perspectives. Paper 1 uses a literature review of instructional coaching research to make an argument for conceptualizing coaching as embedded, boundary-spanning work. This paper explores the shifting and contingent environments that influence coaching activities, and

argues that coaching is an inherently dynamic, dialogical interaction that is shaped by the system in which it takes place. This study also synthesizes the findings about coaching so far, revealing that there is significant emerging evidence that instructional coaching *can* contribute to improved teaching and learning under certain circumstances, but that the field has limited information about the specific practices and contextual conditions that support successful coaching initiatives (Allen et al., 2011; Biancarosa, Bryk, & Dexter, 2010; Blazar & Kraft, 2015; Bryk, et al., 2015; Campbell & Malkus, 2011; Mangin & Dunsmore, 2015; Mudzimiri et al., 2014; Obara & Sloan, 2009; Powell et al., 2010; Sailors & Price, 2010).

The character of instructional coaching also varies based on local contextual factors, and coaches have to navigate and negotiate complex organizational dynamics to enable instructional improvement; in fact, qualitative work on the specifics of high-leverage coaching indicates that understanding coaching requires understanding its embeddedness in the environments that shape their roles (e.g., Huguet et al. 2014; Coburn & Woulfin, 2012). This poses a significant methodological challenge for researchers seeking to understand *how* coaching works; as Kraft et al. (2017) have also argued, the constellations of contributing factors that affect coaching impacts suggest that a nested, contextualized theory of coaching is key to understanding its function and potential in school systems. Paper 1 uses systems theory to further elaborate this claim, showing that explicating the many interactions and dynamics that make up a system in which a coaching effort or initiative is embedded constitutes a necessary contribution to the field.

Accordingly, Paper 2 responds to this imperative by making coaching contexts a key object of analysis, extrapolating this systems view and applying it to an empirical project. By aiming to answer the research question, *what contextual conditions facilitate and constrain coaching, and why?* this study activates the theoretical argument I make in Paper 1 and explores the interaction

between coaching practice and coaching contexts. Examining instructional coaches' work across a variety of locations also provides an important opportunity to think about how different system structures interact with coaching practice. This comparative element of Paper 2 allows me to investigate different kinds of schools and districts with various approaches to coaching practice, and build theories about the significance of different structural conditions based on the patterns that emerge.

Paper 3 also takes a complex systems perspective in that it explores the ways in which four districts, all members of a NIC, took up and enacted continuous improvement methods. NICs are explicitly systems-focused: an axiom of improvement work is, "every system is perfectly designed to get the results it gets".[1] This NIC context provided an excellent opportunity to explore an educational improvement effort from a complex systems perspective, since it allowed me to observe educators using inquiry methods to understand and change their literacy teaching and learning systems. By exploring the research question, *How did four case study districts in the network use and adapt continuous improvement methods to navigate complexity in their organization?* Paper 3 explores the ways in which continuous improvement uptake looked different across 4 distinct district literacy systems, and uses these comparisons to better understand how continuous improvement might be tailored adaptively to suit the unique structural characteristics in educational organizations.

[1] The Institute for Healthcare Improvement, a leader in improvement science in the healthcare sector, attributes this quote to Earl Conway and Paul Batalden. http://www.ihi.org/communities/blogs/origin-of-every-system-is-perfectly-designed-quote

1.2 Collaborative Problem Solving in Networks

Returning to Penuel and colleagues' (2017) idea of collaborative problem solving research is also helpful framing here, since both empirical papers take place in this kind of project context. Paper 2 examines coaching contexts across a statewide coaching network conducted within the context of a research-practice partnership, which used design-based implementation research methods to integrate a coaching model across a variety of contextual conditions. This project assumed local context mattered, and that adaptation would be critical to coaches' uptake and enactment of new coaching practices in their home districts. Examining coaching work that takes place within this collaborative project context responds to the key questions about variation in coach role, program structure, and system characteristics that arise from the coaching literature (e.g. Kraft et al. 2017) that I explore in Paper 1. This approach to understanding coaching effects reflects the recent call in educational research to move beyond apprehending what works in randomized control trials to investigating *how* and *why* interventions work in an effort to make solutions to pressing policy problems more adaptable and equitable (e.g. Bryk et al. 2015; Penuel et al. 2017).

The NIC that I examine in Paper 3 also offers important insights into structuring and orchestrating collaborative problem solving efforts in education; this network provides a case study of this emergent, cooperative approach to educational improvement, and builds theory about how to support educators as they engage in potentially transformative continuous improvement efforts in educational contexts. NICs also make variation between local contexts intelligible and visible by design, in that they engage multiple stakeholders across contexts and across layers of a system (Bryk et al. 2015; Russell et al. 2017). This NIC also engaged individuals strategically, who learned continuous improvement practices from the network and then carried those practices

back to their home districts. Investigating this uptake and implementation process allowed me to further develop theories about how educators use continuous improvement strategies, and in doing so, activate social and technical resources, to confront and navigate the inevitable complexity of school and district environments.

1.3 Brokering and Boundary Spanning

Paper 1 explores another key concept that undergirds the improvement and change processes I examine in Papers 2 and 3; the act of boundary spanning, where key brokers in organizational structures play a critical role in the process of interpreting and enacting change processes and reforms (e.g. Daly et al., 2014). Instructional coaches inherently exist in a boundary space, since they are often tasked with supporting and guiding teachers as they take up instructional reforms and initiatives, thereby connecting the macro-level district policy environment with teacher practice (Woulfin, 2014). This involves a process of negotiation and adaptation that interacts with system characteristics (e.g., the broader policy environment as well as the teachers' capacities and needs). Taking this interaction into account, Paper 2 builds on Paper 1's deep discussion of boundary spanning work by examining the ways that the social, technical, and organizational structures in which coaches are embedded affect observable qualities in their coaching practice.

Paper 3 takes up this concept by exploring the boundary-spanning work of improvement team leaders in a NIC; these key players bridge their local school contexts and the NIC context by learning improvement science methods and taking them back to their home districts. Akkerman and Bakker (2011) noted that boundary spaces can be characterized by discontinuity, providing a

point of similarity or overlap between two dissimilar but connected locations. This idea helps to explain the work of district improvement leaders who learn improvement science methods in a NIC and then translate those methods into their local organizations: they essentially bridge and broker a connection between the network, which encourages the adoption of new and often disruptive practices, and their home district, which has its own separate, pre-existing set of organizational and cultural norms. This bridging mirrors the work of instructional coaches as well, since they are the point of translation and transmutation as new policies and practices move from one distinct organizational location (the central office) to another (the classroom). Finally, Paper 1 suggests that these boundary-spanners have the capacity to learn from the multiple system locations in which they are embedded, suggesting they can be resources for system change and improvement.

1.4 Continuous Improvement and Adaptive Integration

District improvement leaders in NICs are distinct from instructional coaches writ large in that they take up a specific continuous improvement methodology (improvement science) to facilitate change and improvement in their schools and districts. However, the adaptive work that enables these continuous improvement efforts has important parallels to instructional coaching, and suffuses any boundary-spanning role. Continuous improvement enables the process of adaptive integration, where a tool or practice developed in one place is changed and refined to fit into a new context, while retaining the integral components of the original design (Cannata et al. 2017; Hannan et al. 2015; Russell et al. under review; Tichnor-Wagner et al. 2018). While coaching does not necessarily involve a systematic method for adapting practices, it does involve

translating ideas, policies, and tools that originate at one system level (the district office and the state department of education) and must be integrated with a degree of alignment into others (the school and the classroom). Both roles activate the potential of the key actors that create change in systems, suggesting that people in boundary-spanning roles are key players for adaptive integration efforts.

These intertwining conceptual threads form the backbone of these three dissertation studies and respond to current trends in the field that encourage collaboration and creative problem-solving in response to the deeply entrenched problems that the education system faces. Paper 1 creates a conceptual foundation for taking a complex systems theory approach in implementation studies, and argues for the importance of structural perspectives in educational interventions. Papers 2 and 3 support this argument through empirical explorations of two interventions that took these considerations about structure and complexity into account, and demonstrate the significant learning and theory-building potential of these kinds of projects.

1.5 References

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, 81(2), 132-169.

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333, 1034-1037.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1-14.

Biancarosa, G., Bryk, A., & Dexter, E. R. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *Elementary School Journal*, *3*(1), 7-34.

Blazar, D., & Kraft, M. A. (2015). Exploring mechanisms of effective teacher coaching: A tale of two cohorts from a randomized experiment. *Educational Evaluation and Policy Analysis, 37*, 542-566.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In *Frontiers in sociology of education*, Maureen Hallinan (Ed.). New York, NY: Springer.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. (2015). *Learning to improve*. Cambridge, MA: Harvard.

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *The Elementary School Journal*, *111*(3), 430-454.

Cannata, M., Cohen-Vogel, L., & Sorum, M. (2017). Partnering for improvement: Improvement communities and their role in scale up. *Peabody Journal of Education*, *92*(5), 569-588.

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

Coburn, C. E., & Penuel, W. R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, *45*(1), 48-54.

Coburn, C. E., & Woulfin, S. L. (2012). Reading coaches and the relationship between policy and practice. *Reading Research Quarterly, 47*(1), 5-30.

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

Fligstein, N., & McAdam, D. (2012). *A theory of fields*. Oxford, UK: Oxford.

Hannan, M., Russell, J. L., Takahashi, S., & Park, S. (2015). Using improvement science to better support beginning teachers: The case of the building a teaching effectiveness network. *Journal of Teacher Education, 66*(5), 494-508.

Holland, J.H. (2014). *Signals and boundaries: Building blocks for complex adaptive systems.* Cambridge, MA: MIT Press.

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

Huguet, A., Marsh, J. A., Farrell, C. C. (2014). Building teachers' data-use capacity: Insights from strong and developing coaches. *Education Policy Analysis Archives*, *22*(52), 1-28.

Kraft, M. A., Blazar, D., & Hogan, D. (2016). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 0034654318759268.

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

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

Mangin, M. M., & Dunsmore, K. (2015). How the framing of instructional coaching as a lever for systemic or individual reform influences the enactment of coaching. *Educational Administration Quarterly*, 0013161X14522814.

McLaughlin, M. W. (1991). The Rand change agent study: Ten years later. In A. Odden (Ed.), *Education policy implementation* (pp. 143-155). Albany, NY: State University of New York Press.

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life*. Princeton, NJ: Princeton University Press.

Morris, A. K., & Hiebert, J. (2011). Creating shared instructional products: An alternative approach to improving teaching. *Educational Researcher*, *40*(1), 5-14.

Mudzimiri, R., Burroughs, E. A., Luebeck, J., Sutton, J., & Yopp, D. (2014). A look inside mathematics coaching: Roles, content, and dynamics. *Education Policy Analysis Archives*, *22*(53), 1-32.

Obara, S., & Sloan, M. (2009). The evolving role of a mathematics coach during the implementation of performance standards. *Professional Educator*, *33*(2), 11-23

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376-407.

Penuel, W.R., Peurach, D. J. et al. (2017). Defining collaborative problem solving research: Common values and distinctive approaches. White Paper. Boulder, CO: University of Colorado Boulder School of Education, Research and Practice Collaboratory.

Powell, D. R., Diamond, K. E., Burchinal, M. R., Koehler, M. J. (2010). Effects of an early literacy professional development intervention on Head Start teachers and children. *Journal of Educational Psychology*, 102, 299-312.

Tichnor-Wagner, A., Allen, D., Socol, A. R., Cohen-Vogel, L., Rutledge, S. A., & Xing, Q. W. (2018). Studying implementation within a continuous continuous-improvement process: What happens when we design with adaptations in mind? *Teachers College Record, 120*(5), 1-52.

Russell, J. L., Bryk, A. S., Dolle, J. R., Gomez, L. M., Lemahieu, P. G., & Grunow, A. (2017). A framework for the initiation of networked improvement communities. *Teachers College Record, 119*(5), 1-51.

Sailors, M., & Price, L. R. (2010). Professional development that supports the teaching of cognitive reading strategy instruction. *The Elementary School Journa*l, 110(3), 301–322.

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.

Woulfin, S. (2014). Charting the research on the policies and politics of coaching. *Education Policy Analysis Archives, 22*, 2-12.

2.0 Dissertation Paper #1:

Conceptualizing the Work of Instructional Coaches: Boundary Spanning in Complex Systems

2.1 Introduction

Educational policy initiatives are increasingly relying on instructional coaches to improve classroom teaching. In fact, coaches have become such a popular policy mechanism for the standards-reform era that rates of coach staffing have doubled since the turn of the century, while the rates of other school district personnel staffing have remained relatively stable (Domina et al., 2015). Although there is no standard definition of instructional coaching, in general, coaches are often understood as versatile professional development tools for education, in that they can fill a variety of instructional support and policy implementation roles (e.g., Coburn & Woulfin, 2012). Though the details of their roles vary, instructional coaches generally aim to develop teachers' expertise in content areas and mediate policy as it is translated into the classroom, simultaneously mentoring individual teachers and enacting collective reforms (Coburn & Woulfin, 2012; Domina et al., 2015; Knight, 2004; Mangin & Dunsmore, 2015). Despite the recent proliferation of the coaching role, the immense variability in coaching positions, coupled with the heterogeneity of state, district, and local school organizations, pose conceptual and methodological challenges for researchers trying to analyze coaching and its effects.

Despite these challenges, instructional coaching is emerging as a critical school improvement approach across the education reform landscape. Domina and colleagues (2015) discussed coaches' role in the shifting organizational landscape of education, explaining how the

need for "professional sense-makers" and policy translators to support standards-based reforms has likely driven this growing investment in coaching positions. As educators integrate standards-based reforms into their practice, coaches guide these interpretative processes through cooperative sense-making (Coburn, 2004; Domina et al., 2015; Spillane et al., 2002). Research, however, has not kept up with the rapid rise in coaching implementation (e.g., Mangin, 2009b). We know that coaches can have positive effects, but how and under what conditions these positive effects emerge remains unclear (Blazar & Kraft, 2015; Bryk Gomez, Grunow, & LeMahieu, 2015; Mangin & Dunsmore, 2015). As the role that coaches play in implementation continues to grow, it is crucial for policy researchers to develop a clear and actionable understanding of how coaching works and the conditions that can support it.

Coaches' work often includes giving situated, targeted instructional feedback, helping teachers with data interpretation, orchestrating professional learning communities, and other group professional development activities (Bertrand & Marsh, 2008; Coburn & Russell, 2008; Horn, Kane, & Wilson, 2015; Mangin & Stoelinga, 2010). Other conventional coaching activities include classroom observations, teaching demonstrations and modeling, and coaching "cycles" that include pre- and post-observation conferences with teachers (Neufeld & Roper, 2002; Gallucci et al., 2010; Mudzimiri et al., 2014). Districts and schools also compel coaches to direct those support activities toward capacity development for collective change and improvement, including the adoption of standards-based policies and other collective efforts to improve student outcomes (Coburn & Woulfin, 2012; Mangin, 2009a; 2009b; Mangin & Dunsmore, 2015). However, coaches can get pulled in to all kinds of other roles that are not part of their job description or covered in formal research studies—for instance, they might be asked to take on administrative duties that the principal is struggling to cover. Coaches occupy a nebulous position in school and

district hierarchies; while they typically do not have formal supervisory authority over teachers, they can sometimes be perceived as evaluators or compliance officers; this creates tension and likely undermines coaching effectiveness, since trust is a critical condition for coaching, and school improvement more broadly (Bryk & Schneider, 2002; Coburn & Russell, 2008; Coburn, Russell, Kaufman, & Stein, 2012; Mangin & Dunsmore, 2015).

Coaches also face challenges as they navigate the many different reforms, initiatives, and stakeholders that are inherent to school improvement. In school environments characterized by multiple, and often conflicting, mandates and goals, the necessity of *crafting coherence* from policy messages in educational organizations has been well established in the implementation literature (e.g., Honig & Hatch, 2004; Penuel, Riel, Krause, & Frank, 2009; Russell & Bray, 2013; Spillane, et al., 2002), and coaches can be central to this process. Crafting coherence involves intensive sensemaking, a sociological concept that describes the active process of constructing meaning through individual and social processes of cognition, negotiation, organization, and interpretation (Coburn, 2001; Weick, 1995). Due to their inherent proximity to this instructional sensemaking, connections to building and district leadership, and professional duties that often include targeted capacity building, coaches are uniquely positioned to impact these policy processes. Expectations that coaches perform *both* instructional improvement and policy enactment, coupled with the accelerating investments in coaching across the country, indicate that this is a critical position to understand in educational systems.

The process of policy implementation in US education—particularly, the challenge of creating meaningful, lasting change in classrooms—is notoriously difficult due to the diffuse, fragmented, loosely-coupled design of the education system (e.g., Coburn, 2004; Fusarelli, 2002; Smith & O'Day, 1991; Weick, 1976). Schools, districts, and states across the country are reckoning

with this reality as they adopt college and career readiness- focused standards, such as the Common Core State Standards (CCSS), which demand that teachers make significant changes to their practice (Porter, Fusarelli, & Fusarelli, 2014). This shift requires both high-quality, embedded professional development and interpretive sense-making, so instructional coaches will play a critical role in the process (Blazar & Kraft, 2015; Coburn & Woulfin, 2012; Domina et al., 2015; Mangin, 2009). Furthermore, the preponderance of instructional coaching positions in schools that serve high proportions of students of color and low-income students suggests that understanding the mechanisms of instructional coaching is a matter of educational equity as well as quality (Domina et al., 2015).

Researchers are increasingly finding that coaches' policy brokerage role is both potentially powerful and poorly understood (Coburn & Woulfin, 2012; Domina et al., 2015; Galey, 2016; Mangin & Dunsmore, 2015; Woulfin & Rigby, 2017). Accordingly, to understand the impact of coaching in schools, researchers also need to attend to understanding coaching work itself, and conceptualize it in a way that provides a strong foundation for research design and evaluation. In this paper, I argue that conceptualizing coaching from a systems perspective, with particular attention to the act of boundary spanning, offers potential to address this challenge, since this frame reflects the inherent contextuality and complexity of the coaching role, and its relationship to both implementation and improvement. In this article, I develop a conceptual frame for instructional coaching that employs these concepts to highlight gaps and opportunities for further exploration in the current coaching research landscape. To do this, I begin with discussions of boundary spanning and complex systems theories, and show how these ideas can be used to develop a deeper understanding of instructional coaching roles and activities. I then turn to the coaching literature

and review what we know about the complexity of coaching work, and conclude by discussing the practical implications of these theoretical additions to the coaching literature.

2.2 Conceptual Framework

2.2.1 Boundary Spanning in Complex Systems

Building on Coburn & Woulfin (2012), who found that instructional coaches play both educative and political roles in their work, this essay uses the extant literature on instructional coaching to theorize coaches' part in policy implementation, specifically positioning coaches as boundary spanners in complex systems. Boundary spanners are key evidence and policy brokers in school organizational structures; they occupy highly influential positions in terms of enacting and interpreting policy initiatives (e.g. Daly et al., 2014). Coaching activities take place in a critical boundary between state and district-level policy and classroom instruction; as schools respond to mounting pressures to adopt policies that improve teaching and learning, coaches are frequently tasked with connecting macro- and meso-level policies with teacher practice (Woulfin, 2014). Consequently, understanding coaching requires understanding policy negotiation, translation, and implementation processes across these system levels. The high degree of variation in coach roles across contexts (Blarney, Meyer, & Walpole, 2009; Galey, 2016; Marsh et al., 2008; Mangin & Dunsmore, 2015; Matsumara & Wang, 2014) also suggests that coaching effectiveness hinges on the contextual factors that shape, enable, and constrain coaches' activities.

Coaching roles are extremely heterogeneous; one district may use coaches primarily to mentor new teachers in grades K-2, while another may use a literacy coach to roll out a new

initiative among all teachers in a small K-12 system, a third may focus coaching efforts on helping teachers' interpret and apply student achievement data in small groups, and yet another may dedicate a mathematics coach to only intermediate grades to build teaching capacity in that content area—there are seemingly endless configurations of coaching roles and responsibilities. Coaches, unsurprisingly, report feeling more supported and confident when their roles are well-defined (Mangin & Dunsmore, 2015), suggesting that developing knowledge about the forms of coaching that are most effective and meaningful could have important practical implications in terms of supporting coaching in various contexts. Therefore, in order to build a better understanding of the complex dynamics that influence coaching, this essay draws on the concept of boundary spanning situated within complex systems theory to conceptualize coaching, its mechanisms, and the system structures that can facilitate it.

Boundary spanning theory calls attention to the ongoing negotiation that coaches do as they work to balance many competing demands. Akkerman and Bakker's (2011) seminal review on boundary spanning in educational research explicates the many applications of boundary theories in education. This piece is particularly helpful in building a conceptual framework for coaching research because it is both theoretically dense and grounded in empirical, educational applications of boundary spanning theory. Using boundary spanning as a conceptual foundation for examining coaching is helpful in two specific ways: 1) dialogical engagement, which is critical to boundary spanning, provides a multi-directional analytical frame for understanding the collaborative sense-making of coaching interactions, and 2) this frame focuses on specific interactions while still emphasizing the organizational structures in which coaches are embedded. While researchers nod to the boundary spanning qualities inherent to coaching work (e.g., Woulfin, 2014), a deep

discussion of the nature and mechanisms of coaches' boundary spanning has yet to appear in the literature.

The many discrete but connected sites, processes, and people in the education system emphasize the importance of boundaries in educational theory and scholarship. Akkerman and Bakker (2011) explain that boundaries suffuse education in a variety of theoretical and practical ways: for instance, the movement from peripheral to central as a learner develops expertise, or the development of identity through distinguishing what is self and what is not-self. Boundaries are spaces of socio-cultural difference that give rise to "discontinuity in action or interaction" (Akkerman & Bakker, p. 133, 2011). They also inherently contain both discontinuity and continuity in that boundaries are a point of similarity between two dissimilar but connected locations. Since coaching work largely takes place at the boundary between the individual classroom and district office, their work is always balancing competing and sometimes incongruent demands. The instructional needs and pedagogical orientations of classroom teachers are not always aligned with district policy priorities, so coaches have to carefully navigate these conflicts while producing improvement, meeting district expectations, and increasing capacity for reform.

2.2.2 Negotiating Boundaries

Boundaries are also characterized by the potential for the production of new meaning; boundary spanning involves "intersections of cultural practices [that] open up third spaces that allow the negotiation of meaning and hybridity—that is, the production of new cultural forms of dialogue" (Akkerman & Bakker, 2011, p. 135). In other words, by virtue of being in between defined locations, these liminal spaces create opportunities for the articulation and negotiation of

what defines the things around the boundary. For instance, coaches have to find ways to translate policy messages so that the teacher can understand and apply them, thereby transmuting policy messages so that they (presumably) retain essential characteristics valued by the district *and* somehow align with a teachers' instructional abilities and mindsets. Boundary crossing involves continual processes that connect the contexts on either side through dialogue (Akkerman & Bakker, 2011). Bakhtin's (1981) theory of dialogicality can be helpful to understanding this aspect of boundaries: dialogic interaction in boundaries enables processes of meaning emergence and generation. That is, in boundaries, embedded, contextualized encounters with multiple perspectives articulate realities that exist beyond the boundary, and build new ways of meaning.

So, for instance, a coach might come to a coaching conversation thinking one thing about policy messages she is responsible for translating, but then, when faced with her teacher's needs, she may adjust the policy messaging to accommodate the teacher, or adjust her instructional support to better integrate the policy into the teacher's development. In both scenarios, the content of the interaction likely shifted, producing a new and different piece of feedback. This adaptable quality requires the coach to process perspectives from both the teacher and the district and devise a way they can coexist, even when they are fundamentally at odds. The importance of the coach responding to context in these examples underscores the idea of boundaries as a *learning resource*—these are not merely transitional spaces, but they offer opportunities to learn about different perspectives and to activate that knowledge in order to connect disparate points of view. Understanding coaches' marginal positioning as boundary spanning between the district office and street-level bureaucrats implies that, as they negotiate the competing demands of policy makers, district personnel, school administrators, and teachers, coaches not only translate policy messages into school contexts, they renegotiate meaning and create altogether new ways of understanding

and implementing those policies. Coaches also grapple with contextual constraints that filter down from the district office and up from the classroom, so coaching interactions can reveal insights into those contexts through continual processes of dialogic negotiation and sensemaking.

Given the dialogical qualities of boundary spanning, it is critical to consider the process of learning that takes place across and within boundaries, and how these processes may contribute to an understanding of coaching interactions. For instance, situative learning theorists also use boundary spanning concepts, as evidenced by Lave and Wenger's (1998) landmark work on communities of practice, to illustrate the learning potential of boundary spaces. Lave and Wenger (1998) explain that communities of practice can lose their vitality without learning at the boundaries; in boundaries, new participants learn how to cycle into central positions, and established participants gain understanding of new ideas encountered through peripheral members. Learning at the boundaries, then, is multi-directional in nature; in contrast to ideas of coaching that frame activities as a simple transfer of knowledge and professional development, understanding coaching as boundary-spanning emphasizes how knowledge emerges through a complex, collaborative interaction. Since, in many cases, coaching activities literally connect policy messages to classroom instruction, these interactions likely contain rich insights into the interplay between policy and practice. Boundaries have the potential to explicate current information and knowledge in the existing system *and* create new knowledge that is then incorporated into the system.

2.2.3 Systems Theory

In order to understand coaches' boundary spanning roles, it is also important to adopt a systems perspective; that is, to consider not only coaches' work and activities within boundary

spaces, but to attend to the structure of the system in which this work is embedded. Given the many different forms of coaching, which can be formal, informal, cognitive, evaluative, supervisory, and/or supportive (Mangin & Dunsmore, 2015), the organizational infrastructures that shape these roles ought to be considered in any analysis of coaching. Taking a complex systems perspective entails conceptualizing performance as a product of interactions of the many people, tools, materials, and processes that compose a complex system (Bryk et al. 2015; Holland, 2014; Miller & Page 2007). As Bryk and colleagues (2015) have explained, educational organizations are complex adaptive systems involving many interactions and interconnections, which may react unpredictably to attempts to change them.

Complex systems cohere around rules and norms that are established through interaction and periodically destabilized during times of change (Fligstein & McAdams, 2012). Social actors are also key forces in shaping and changing these complex systems (Fligstein & McAdams, 2010; Miller & Page, 2007). Systems theorists use structures and system levels to shape their analyses; for instance, Fligstein & McAdam (2012) explained that the meso-level (or middle level, between micro- and macro-) of a system forms a strategic action field, or an organized space of social interaction. In this meso-level, interactions between actors proliferate, and information from these interactions flows into the broad, macro-level of the system to shape large-scale system characteristics. During times of change or instability, actors recursively negotiate meanings, norms, and patterns at the meso-level of a system, and through this process, change coalesces and spreads across the system levels, eventually reaching micro-level, individual actions that make up the components of the system. Instructional coaches, then, can be understood living in these strategic action fields, continually engaging in these potent meso-level interactions where norms of policy enactment and instructional improvement are established, negotiated, and revised

through dialogical learning. Furthermore, this suggests coaches' positioning and role flexibility suggests that they have unique potential for shifting system dynamics and characteristics.

In this meso-level position, coaches serve as a link between macro administrative structures (like district offices and even state departments of education) and micro-level classroom environments. Systems theory tells us that actors in this position critical in their capacity to spark change, and this depends on their ability to empathize across system levels. As socially skilled actors in the meso-level negotiate change, their success often depends on their ability to identify with the "other"—that is, to understand the needs and perspectives of actors across multiple levels of the system, and leverage that understanding to counter, challenge, and convince (Fligstein & McAdam, 2012). Without this empathic capacity, work at the meso-level often fails to produce change; for coaches, this means that if they cannot connect with both macro- and micro-structures in meaningful ways, their work may be less impactful. This illuminates another important element of coaching work which has been underexplored in the literature—the ways in which coaches respond to various demands coming from different levels of the system, and the methods they use to connect with the many different people with whom they work.

These meso-level activities cannot be fully understood (or leveraged) without taking the broader system structure into account. As Opfer and Pedder (2011) argued in their review of teacher learning from a complex systems perspective:

> complex systems thinking assumes that there are various dynamics at work in social behavior and these interact and combine in different ways such that even the simplest decisions can have multiple causal pathways… how these dynamics combine will vary for different people and even for the same person at different times of the day or in different contexts. Furthermore, the ways they combine and the circumstances in which they

> combine are patterned; there may be a large number of reasons for tea drinking, but the reasons are neither limitless nor tend to be random. (378)

That is, understanding complex systems requires thinking across levels, interactions, and configurations of many different processes, people, tools, and orientations. Opfer and Pedder (2011) explained that many conceptualizations of teacher learning (which includes, but is not limited to, the school-based professional development activities of coaches) neglect the interactions between the micro-, meso-, and macro-levels of school systems when investigating how teacher learning works. They therefore propose complex systems theory as a way to understand teacher learning, since it accounts for the many interdependent forces that shape teacher learning activities and outcomes.

Similar to theories of boundary-spanning, complex systems theory emphasizes the interplay of interdependent parts, the disequilibrium inherent to complex structures, and the generative potential of this disequilibrium (Argyris & Schon, 1996; Clark & Collins, 2007; Innes & Booher, 2010; Holland, 2014; Miller & Page, 2007; Opfer & Pedder, 2011; Seashore Louis & Leithwood, 1998). For instance, Opfer and Pedder (2011) discuss how individual's beliefs about learning interact with collective norms of action in a school organization, and then, those school-level learning orientations influence and interact with individual learning theories. Like wicked problems (e.g. Innes & Booher, 2010) complex systems rest on many linked interactions that are situated, influenced, and constrained by structural factors.

The continual interactions of actors and processes enable complex systems to absorb information and feedback, remaining dynamic and responsive to the imbalances that define them (Miller & Page, 2007). This idea of the absorption of feedback and knowledge in complex systems is particularly relevant to coaches' functioning in educational organizations; as actors who move

between classroom level, school level, and district level sites fairly frequently, coaches represent an important point of knowledge absorption and generation in addition to their role of feedback provider. Educational systems struggle with responsiveness due to notoriously loosely-coupled, disconnected structures (e.g. Coburn, 2004), so coaches' work at the meso-level is an invaluable learning resource for the system writ large, not just for teachers receiving instructional feedback and policy messages. Furthermore, coaches are in a unique role to identify and even adapt signals in educational systems. John Holland's (2014) theory of complex adaptive systems posits that adaptive agents exist in boundaries where they process and send signals, accepting some and ignoring others, using mechanisms to create change based on the experiences they accumulate. Coaching work maps on to this theory of adaptive agents; coaches use what they know of state and district policy constraints and classroom-level realities to make practical changes based on what they can reasonably assume is possible. This suggests that coaches' liminal, adaptive roles, which uniquely position them across many system locations (e.g., school and district central offices), provide an opportunity for developing greater insight into signals and connections that make educational systems tick.

Taken together, systems and boundary theories help to elucidate the nature of instructional coaching by attending to the many facets of this complex and poorly understood position in educational organizations. Complex systems theory enhances a conceptualization of coaching activities and roles by embedding coaches in the multivalent organizational structures that they continually negotiate, while boundary spanning theory grounds coaching activities in multi-directional, continual microprocesses that constitute coaches' everyday work-lives. These boundary spaces, what Miller & Page (2007) refer to as the "messy in-between" (p. 229), also

foster adaptation, suggesting that coaching interactions can be a key lever for change in complex school systems.

In order to understand this process, however, a delicate balance between the particular and the general is necessary. Only by acknowledging the nested nature of micro, meso, and macro-level learning activities can a more robust conceptualization of coaching emerge. That said, incorporating boundary spanning theory into a systems-informed conceptualization of coaching emphasizes the microprocesses of coaching interactions *and* embeds these interactions within the other systems and subsystems that influence, and are influenced by, coaching activity. This is not to say that every single possible system interaction must be examined in any analysis of coaching—a comprehensive accounting of this would be impossible. However, as Opfer & Pedder (2011) also argued, trying to understand phenomena within educational organizations by focusing on only a single subsystem cannot provide much insight into either the phenomena or the system itself. Furthermore, taking a systems theory approach requires recasting coaching as a recursive activity that leads to individual *and* system learning, rather than as a one-directional professional development delivery system. While coaches certainly share knowledge and expertise with teachers, they also toggle between classroom-level, school-level, and district-level locations in the system, presumably absorbing information and feedback at each level, which would then shape the way they provide feedback and translate policy messages. Accordingly, in this essay I conceptualize coaching as always nested, multidirectional, and interdependent, and consider how to build on previous research to provide insight into coaching subsystems and their potential for creating systemic change. Figure 2.1 visualizes this theory of coaching dynamics.

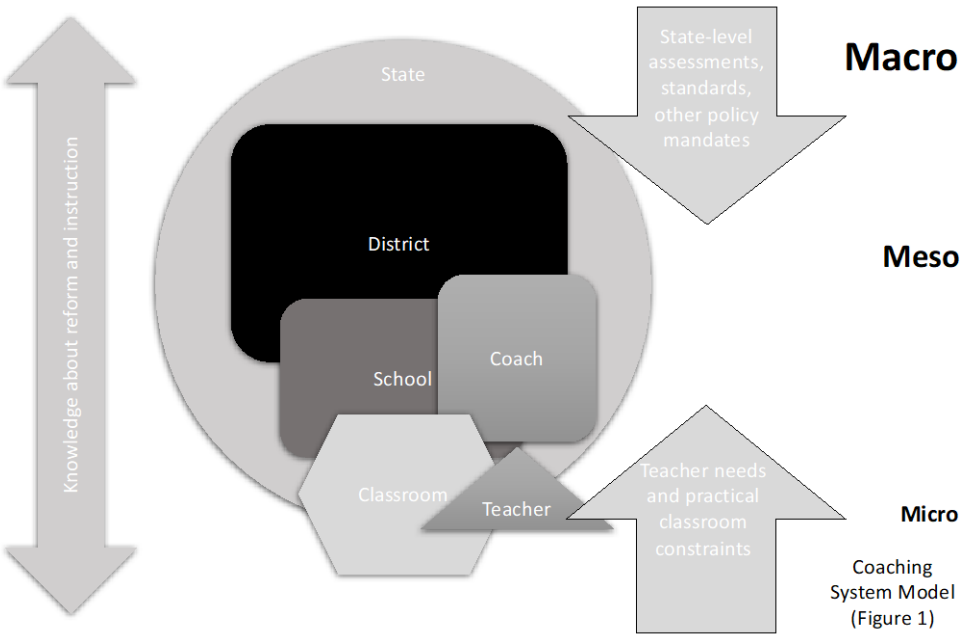

Figure 2.1. Modeling Coaching in Complex Systems

Because this essay is both a reflection on the existing literature and a reconceptualization of coaching, in the next section I begin by sketching the landscape of coaching literature, reflecting on the outcomes-focused, mostly quantitative literature that frames the need for a more nuanced understanding of coaching work. Then, I move to a closer analysis of a sample of primarily qualitative studies that are especially relevant to understanding what is known about coaching in the literature and how researchers have grappled with coaching conceptually. In particular, I focus on studies that have followed the recent call for more complexity in conceptualizing teacher learning (Opfer & Pedder, 2011) by engaging, directly or indirectly, with two themes: 1) coaching microprocesses, (i.e. everyday practices, like feedback conversations with teachers) and 2) system context and structural variability as central features of coaching work. In the spirit of building a rich conceptual contribution to the coaching literature, this focus allows me to focus on pieces that provide the best fodder for my conceptual project while remaining grounded in empirical

knowledge. It is also important to note that, given considerable political pressure to connect coaching initiatives with improved student assessments via teacher professional development effects, many studies of coaching are understandably focused on outcomes rather than contextual conditions. A subset of studies, however, works from conceptualizations that are more explicitly embedded and complex, and unpacking this work is critical to this essay.

To select the literature I review here, I started with a search in Google Scholar for all peer reviewed studies of instructional coaching, cross checking my search results with well-regarded coaching studies to ensure that I was accurately capturing the broad themes in the field. As Galey (2016) noted in her recent coaching review, there are surprisingly few peer-reviewed coaching studies, particularly given the recent increase in interest in the coaching role. Then, I scanned the literature to determine which studies would give insight into the conceptual framework I propose here. These largely fell into two categories: studies that discussed specific coaching practices, or microprocesses—such as conversations with teachers about pedagogical strategies—and studies that discussed specific coaching contexts—such as the structure of coaches' roles and the policy environment of the districts in which they work. Many of the studies I use to illustrate the broad themes in the coaching literature mentioned at least one of these themes superficially but did not engage with them in depth. In the second part of the review, I focus on unpacking the research that deeply and directly engaged with these two themes to develop a sense of what is known about these two critical aspects of the job and to support my conceptualization of coaching.

2.3 Literature Review

2.3.1 Broad Landscape of Coaching Research

The literature shows mounting evidence that instructional coaching *can* have an impact on high-leverage outcomes like instructional quality and student learning (Allen et al., 2011; Biancarosa, Bryk, & Dexter, 2010; Blazar & Kraft, 2015; Bryk, et al., 2015; Campbell & Malkus, 2011; Foster & Noyce, 2004; Matsumura et al., 2010; Matsumura, Garnier, & Spybrook, 2012; 2013; Neumann & Cunningham, 2009; Powell et al., 2010; Sailors & Price, 2010). While there is considerable research that suggests that coaching has the potential to impact teacher and student learning, decisive quantitative proof of positive effects has been elusive, since many coaching studies lack critical details about context, implementation, and specific coaching models (Biancarosa, Bryk, & Dexter, 2010; Blazar & Kraft, 2015). Many researchers have also noted the difficulty in understanding coaching effectiveness given the immense variability in coaching roles and programs across different schools and districts (Campbell & Malkus, 2011; Mangin & Dunsmore, 2015; Mudzimiri et al., 2014; Obara & Sloan, 2009). The distance between coach influence and student outcomes also makes determining coach effectiveness based on student outcomes difficult (Blazar & Kraft, 2015; Mangin, 2009a; 2009b). Despite these challenges, a burgeoning literature on coaching is taking shape.

Quantitative studies of coaching effects show that coaching programs have been associated with significant gains in student and teacher learning *under some conditions*, though there is no consensus on the specific practices, programs, and implementation contexts that support these gains. In one study, Biancasrosa, Bryk, and Dexter (2010) use a quasi-experimental, accelerated longitudinal cohort design and value-added modeling to trace students' growth trajectories in 17

schools implementing comprehensive literacy reforms that relied heavily on coaching interventions. Their results suggest that "well-specified, well-supported coaching initiatives can effect positive change in student learning," (Biancarosa, Bryk, and Dexter, 2010, p. 28), but the authors also called for continuing research on the diversity of coaching models and contexts in order to understand exactly *how* coaching can impact student learning. Significantly, Biancarosa and colleagues (2010) also noted that their results may have emerged from the growth and development of networks of expertise and innovation in the schools studied, gesturing toward a more system-inflected understanding of coaching initiatives. Atteberry and Bryk (2010) explored this networked capacity question in more detail, discussing how network structures in schools can support innovation by promoting trust and providing access to expert actors (like coaches), emphasizing the pivotal role of system qualities in understanding implementation and coaching.

Similarly, Blazar & Kraft (2015) studied teacher development in a blocked, randomized control trial of two coaching cohorts in one school district in New Orleans, focusing on teacher process measures rather than student assessment outcomes in an effort to closely track the impact of the coaching initiative on teacher development. They found significant improvement effects for coach cohort 1, but none for cohort 2. They hypothesized that changes in the coaching model and the teacher's coached in cohort 2 may have caused this effect, and called for more research on specific design features of coaching models and programs in order to understanding the results of coaching interventions (Blazar & Kraft, 2015, p. 564). These results provide evidence of the potential power of coaching, and also suggest that coaching cannot be well understood without detailed contextual information to frame effects and outcomes. Without information about coaching structures, roles, and circumstances, researchers have often been cautious about claims regarding its effectiveness.

Other quantitative studies have found promising coaching results and effects under specific conditions. Matsumura and colleagues (2010) analyzed a longitudinal, randomized field trial of a comprehensive, content-focused coaching program, finding that coaching exposure was associated with higher school-level achievement gains and teacher-level instructional development. It is important to note that Matsumara et al. (2010) addressed system context in that they discussed the importance of involving principals and central office administrators in coaching initiatives; they explained how content-focused coaching is designed specifically to create a professional culture, and how the development of expertise across key school actors likely contributed to coaching gains over time. Campbell and Malkus (2011) also found that mathematics coaches had a positive effect on elementary student achievement in a three-year randomized control trial. Their study aimed to determine whether coach placement in schools affects student achievement; like Matsumara and colleagues (2010), they noted that coaching effects became more pronounced after year 1 of the study, suggesting that the efficacy of coaching interventions is predicated on the development of a collaborative, expert school culture around coaching. Campbell and Malkus (2011) also explained that by virtue of their quantitative design, their study did not address how coaching was implemented or how implementation varied across individuals and schools. These findings provide additional evidence that coaching outcomes cannot be well understood without contextualizing them in the system that produced them.

Additional research has yielded indecisive results about coaching dosage and effects (Garet et al., 2008; Marsh et al., 2008; Murray et al., 2009; Neufeld & Roper, 2003; Van Keer and Verhaeghe, 2005); these studies offer further evidence for both the immense variability of coaching roles and activities as well as the problem with trying to understand coaching without enough contextual information. For instance, Van Keer and Verhaeghe (2005) studied an intensive,

year-long coaching model involving 35 hours of professional development with a more limited coaching treatment of only 15 hours, and found that both had the same effect on student outcomes, possibly confounding the consensus that effective professional development should be intensive and ongoing in nature (e.g., Darling-Hammond et al., 2009). In a small experimental study of a peer mathematics coaching initiative, Murray and colleagues (2009) found that, while teachers reported that coaching was a positive experience, coaching interactions lacked both analytical rigor and depth. These gesture to the importance of understanding the features and practices involved in the many available coaching models; coaching in itself is not a silver bullet. Ultimately, these studies give preliminary insights into the challenge of analyzing and measuring coaching. The field still has limited information about the circumstances that govern coaching effects, and little is known about specific practices that make up successful (or unsuccessful) coaching initiatives.

2.3.2 Specific, Qualitative Explorations of Coaching Practices and System Contexts

Qualitative studies have discussed coaches' roles and activities in more contextualized detail, giving insight into system characteristics and structures that can shape coaching activities. Broadly, coaching has been shown to shape and enhance teacher learning, facilitate capacity building, bolster professional communities, create opportunities for depth of interaction, and affect the quality and integrity of policy implementation (Betrand & Marsh, 2015; Coburn & Russell, 2008; Coburn, Russell, Kauffman, & Stein, 2012; Coburn & Woulfin, 2012; Gallucci et al., 2010; Horn & Little, 2010; Horn, Kane, & Wilson, 2015; Mangin & Dunsmore, 2015). Significantly, coaching has been found to be a particularly effective strategy for helping teachers shift the aspects of their practice that are most difficult to change, such as pedagogical mindsets and interpretive perspectives (Betrand & Marsh, 2015; Coburn, 2004; Coburn & Woulfin, 2012). Variation across

roles and contexts, however, continues to be an obstacle to making clear-cut claims about coaching's effectiveness and its impact on teaching and learning (Matsumara & Wang, 2014; Woulfin, 2014). Studies that engage with the complexity and variability of coaching roles and implementation, however, provide much needed descriptive insights into the conditions that may support or hinder coaching efforts.

Reflecting calls in both the implementation literature and the teacher learning literature to attend to many complex factors that can influence professional development (e.g., Coburn, 2001; 2004; 2005; Coburn & Russell, 2008; Opfer & Pedder, 2011), the studies discussed in this section frame coaching interactions as embedded in implementation systems defined by nebulous and shifting professional roles and highly variable local conditions. This perspective complicates attempts to define coaching effectiveness through unilateral student outcome measures and contributes to researchers' descriptive attempts to better represent coaching phenomena. Qualitative studies are especially well suited to address these concepts of embedded variability; the fields' limited understanding of coach roles and the complex system structures that shape coaching work suggest that qualitative investigations can contribute essential descriptive insights into the character of coaching and its function in organizational systems. Qualitative investigations also provide opportunities to unpack the complexity of coaching work and illustrate the interdependent interactions that shape it.

One strategy for developing insight into situated, specific coaching work is attending to coaching practices, or microprocesses, to unpack variable roles, responsibilities, and functions of coaches in particular contexts. Conversational routines, such as coaching episodes in pairs or in groups, have the potential to reveal insights into organizational dynamics, relationships, and work patterns, and they offer both stability in their routinization as well as flexibility and the potential

for adaptation (Horn & Little, 2010), which suggests that they could provide crucial information into how coaches adapt and negotiate their roles and the other forces that come into play. As they are repeated across teachers and contexts, these practices form the meso-level negotiations of norms and structures in a system, so they are an integral component of educational systems. Attending to the ways in which coaches navigate these interactions, and how they integrate policy messages into their instruction, can illustrate mechanisms and patterns of change in school systems. Though few studies examine coaching practices specifically, those that do confirm the complexity and mutability of the coach role, and suggest that framing coaches as professional development delivery mechanisms is likely a gross oversimplification.

An important insight that has emerged from research on coaching microprocesses is that their work is both instructional and political in nature; coaches play an *educative* role in schools by providing instructional support, and a *political* role in terms of shaping policy as it filters through the meso-level of the system into teachers' classrooms (Coburn & Woulfin, 2012). In their study of Reading First implementation in one Massachusetts school, Coburn & Woulfin (2012) found that teachers who were coached were more likely to "make substantial changes in their classroom practice" (p. 6), and that, in addition to conveying instructional messages, coaches also advised teachers on *how* to take up the Reading First initiative by emphasizing some parts of the policy and deemphasizing others. By "pressuring, persuading, and at times buffering" (Coburn & Woulfin, 2012, p. 13) teachers in coaching interactions, coaches were critical in shaping teachers' variable and often piecemeal uptake of the policy across the school. Coburn & Woulfin (2012) noted that attending to the political roles of coaches was a significant departure from the existing literature; while there is some consensus on the necessity of separating coaching from evaluation and compliance (e.g. Goldstein 2007; Walpole & McKenna, 2004), this study suggests that even

coaches whose roles are framed as purely supportive still influence teacher sensemaking about policy uptake in significant ways.

Other researchers working in this descriptive strand of research have observed themes in the way coaches approach their work with teachers. For instance, Mudzimiri and colleagues (2014) aimed to develop a deeper understanding of mathematics coaching roles and functions by shadowing seven coaches in five school districts, supplementing observational data with coach history questionnaires. They identified categories of coaching strategies that emerged across their observations: "relational," "information exchange," and "facilitating teacher learning" (Mudzimiri et al., 2014, p. 17), and discussed coaching dynamics and content. Their discussion of coaching, however, largely focuses on instructional support, and does not engage the political power dynamics and implementation influence inherent to coaching work. Methodological limitations make it difficult to draw deep conclusions about patterns in coaching practices from this study—in particular, the authors only observed coaches for one to two days total. However, the comparative, multi-school design (studying seven coaches in five districts) is worth noting, and Mudzimiri and colleagues (2014) did conclude that a more flexible observation protocol was necessary to reflect coaches' fluid and shifting roles.

Huguet, Marsh, and Farrell (2014) also considered coaching practices in a comparative case study of coaches' efforts to build teachers' data-use capacity in four different schools. They found that the two "strong" coaches in their study, who had the clearest positive influence on teachers' data use, drew on a broad base of coaching skills, artifacts, and practices to individualize their coaching for each teacher (Huguet et al., 2014). Furthermore, like other researchers, Huguet and colleagues (2014) concluded that the coaching environment is critical to facilitating (or preventing) meaningful coaching interactions; they echoed other's claims (Coburn & Woulfin,

2012; Mangin, 2009a; 2009b; Mangin & Dunmore, 2015; Matsumara et al., 2010) that leadership's framing and facilitation can heavily influence teachers' responses to coaching. These studies suggest that coaching practices are best understood in relation to the context in which they take place, indicating that situating practices within a leveled systems analysis is a viable approach to elucidating the way that coaching works.

It is also important to note that gaining insight into the uptake, diffusion, and interpretation processes involved in coaching is methodologically intensive; for instance, Coburn & Woulfin (2012) conducted interviews with teachers, coaches, school administrators, and district administrators, in addition to classroom, meeting, and professional development session observations, as well as document collection and analysis, over the course of two years. This kind of extensive data collection is necessary to contextualize coaching practices and interactions. Mudzimiri et al. (2014) and Huguet et al. (2014) both use much more limited data sets—Huguet and colleagues (2014) conducted interviews with coaches, principals, and coached teachers, and completed three site visits to each case-study school over the course of a year, while Mudzimiri and colleagues (2014) shadowed coaches for only two days. This is a practical challenge in the literature—while researchers are trying to quickly account for *how* coaches do their work across variable contexts, the few detailed studies of variable coaching activities suggest that more in-depth qualitative descriptions of coaching across school and district contexts are badly needed.

2.3.3 Coaching in Complex Systems

Turning to what we know about coaching contexts and systems, studies of coaching have also highlighted system structure and cross-level interactions in discussions of framing, district capacity, and interdependencies between coaches and school administrators (Gallucci et al., 2010;

Mangin, 2009a; 2009b; 2014; Mangin & Dunsmore, 2015; Mangin & Stoelinga, 2010; Marsh, McCombs, & Martorell, 2010; Matsumara & Wang, 2014; Stoelinga, 2010; Woulfin, 2014). Though these studies do not necessarily use systems theory explicitly, they attend to structural characteristics in conceptualizing coaching activities and research design, situating coaching in the web of organizational forces that influence it. This affords some insight into coaching's role in educational systems, and a starting point for understanding how coaches negotiate their meso-level, in-between positions as they support both policy implementation and instructional improvement.

Complex systems theory emphasizes the nested and reciprocal relationship of actors in an organization, recognizing the interdependence between the individual and collective in order to tease out the many forces that shape outcomes (Opfer & Pedder, 2011). This orientation is particularly relevant to the coaching literature that focuses on district framing of, and influence on, coaching. By considering how districts shape coaching activities and initiatives, these studies recognize the inherently nested nature of coaching by analyzing the interplay between system levels. Mangin (2009a; 2009b) investigated this idea across several studies, considering how districts make decisions about coaching initiatives and the impact of district context. In a study of how districts decide to use coaches, Mangin (2009a) identified three forces that influenced district personnel's orientation toward literacy coaching—the national reform landscape, which encourages districts to view coaches as a way of complying with national and state mandates in addition to raising student achievement; low student performance data, which often becomes a warrant for coaching support; and pre-existing norms in the district around coaching and PLCs. Significantly, these factors span system levels, again drawing attention to the nested nature of

coaching by considering macro- (national reform landscape), meso- (district and school norms, grouped student performance), and micro-levels (individual student performance) of influence.

In another study of how district-level forces shape coaching, Mangin (2009b) discussed specific contextual features and their relationships to coaching. Using qualitative insights drawn from interviews with central office administrators in 20 districts, Mangin (2009b) builds on contextual factors that influence coaching uptake in districts—"national reform context, finances, student performance data, and existing roles and programs" (p. 770-771)—and discusses how district personnel viewed these factors as either barriers or facilitators to coaching. For instance, half of the district participants discussed finances as a constraint to coaching work, and state and national reform mandates were commonly seen as factors that facilitated coaching uptake in districts. Student performance data could either be a barrier or facilitator, depending on performance outcomes—low scores were associated with strong incentives to use coaches, whereas high performance served as a disincentive (Mangin, 2009b). While she noted that the goal of the study was not to trace all contextual interactions, Mangin's (2009b) findings indicated that the existing roles and programs that did not align with coaching initiatives, coupled with financial limitations, were the most commonly cited negative influences on coaching.

These studies also suggest that the way a district defines coaching roles can be a highly influential factor in determining coaching effectiveness. Mangin (2009b) noted that the literacy coaching models districts used to structure coach roles were significant, identifying three different types: classic, where schools developed coaching positions that included release times for teachers to work on instructional improvement; modified, where districts did not provide coaching release time and/or included activities other than instructional improvement in coach role descriptions; and none, where districts did not take up specific coaching roles beyond typical classroom

instruction (Mangin, 2009b, p. 769). Districts using classic coach roles reported perceiving other contextual factors, like national reform climate and student performance data, as facilitating this sort of coaching model. Districts using modified or no coaching roles more frequently reported other contextual factors as barriers (Mangin, 2009b). While Mangin (2009b) noted that the study was not designed to account for *how* these contextual factors interact, these findings provide additional evidence that system context is a critical unit of analysis in research on coaching. These observations also emphasize the idea that organizational support and district norms are key factors in building and integrating coaching initiatives in districts undergoing reform (Coburn & Russell, 2008; Gallucci et al., 2010; Marsh, McCombs, & Martorell, 2010). Furthermore, this study suggests that coaching, particularly careful data analysis in schools using classic and modified models, can help teachers change their perception of student performance. Engaging in rigorous discussion of student data "helped teachers view student learning as a symptom of the problem rather than the source," (Mangin, 2009b, p. 781), confirming other accounts of coaching as being instrumental in accessing deeply held mindsets about teaching and learning (Bertrand & Marsh, 2015; Coburn & Woulfin, 2012).

This finding is also important because it suggests that certain coaching models can help teachers see the complex system that produces learning outcomes. Mangin & Dunsmore (2015) discussed coaching as a systemic change strategy, noting that coaching is often simultaneously framed as a lever to shift individual *and* collective skills and orientations. In this qualitative study, which examined the enactment of literacy coaching roles in one district with established collective literacy aims, Mangin & Dunsmore (2015) found that the framing of a literacy initiative influenced the strategies coaches used to create change. Coaches in this initiative recognized conflicting messages about individual and collective goals in coach training, suggesting that explicit theories

of change shape coaches' work and understanding of their roles. The tension that emerged between individual change strategies (low-depth interactions with individual teachers) and collective reform goals (district-wide literacy improvement targets) emphasizes the importance of creating coherence and shared meaning across system levels, since these results indicate that conflicts between individual and systemic aims made coaches feel confused and uncertain about their work (Mangin & Dunsmore, 2015). This suggests that explicit attention to the liminal qualities of coaching work may also help coaches develop a more nuanced understanding of their positioning and role within the system.

Principals are also critical to the framing process that shapes coaching work; across studies, school level administrators are identified as driving forces in the change process by virtue of their power over coaching responsibilities and their mediation of policy messages (Coburn & Russell, 2008; Gallucci et al., 2010; Mangin, 2009a; 2009b; Matsumara, et al., 2010; Matsumara & Wang, 2014; Stoelinga, 2010). In their study of administrators' sensemaking about instructional practice and improvement, Matsumara & Wang (2014) found that principals' negotiations of misaligned reform visions and coaches' relationships to teaching and learning are likely important factors in coaching uptake. This corroborates additional findings that principals' explicit framing of coaches as sources of expertise likely improves coach efficacy (Coburn & Russell, 2008; Matsumara et al., 2009). Principals, however, understand coaching in a wide variety of ways, contributing to the variation in coach role structure and implementation (e.g. Camburn, Kimball, & Lowenhaupt, 2008). The interdependent dynamic between coaches and principals also plays out in Stoelinga's (2010) discussion of coach role negotiation, which highlights the tension between support and evaluation, and returns to the concept of systemic alignment—for coaching to be coherent, school leaders need to develop coherent approaches to system change.

While research on coach learning and adaptation in complex systems is sparse, a study by Gallucci and colleagues (2010) examined coach practice and development in a school undergoing district-wide reform, focusing on a single case of a literacy coach in an intermediate school and his opportunities for professional learning. They use the Vygotsky Space model to analyze the coaches' learning and change over time through interviews, observations of conventional coaching activities, and coach professional development opportunities. The authors hypothesized that many interdependent organizational factors—particularly the district's coordinated messaging about a shared vision of literacy reform, the engagement of staff as learners, and the development of supports that remove coaching barriers, such as limited time (p. 951)—contributed to the coach's opportunities to develop expertise. Though Gallucci and colleagues (2010) focused almost solely on the instructional activities of coaching, excluding the political dimensions of the role, they noted the critical intricacies of organizational processes that can support or hinder coach development. Most significantly, the authors discussed the importance of attending to coach learning as they develop expertise in their position, noting that coaches do not function as reform-delivery pipelines, but instead learn to develop expertise over time. While their findings focus on coach professional development and not system learning, the authors argue for studies of coaching that balance detailed analysis of coaching activities with the organizational factors that shape them.

The literature that grapples with coaching contexts sketches a dense and tangled web of influences, suggesting that continued descriptive work that illuminates coaches' negotiation of systemic change is necessary to developing a more robust understanding of their function. Woulfin (2014) specifically argued that more research is needed on the effect of macro-level policy on coaches' work with teachers, also explaining that their boundary spanning position affords them access to many learning resources. Since coaches shape teachers' interpretations of policies and

facilitate changes in their instruction, coaching positions create a link between policy and practice (Woulfin, 2014). Woulfin's (2014) point reflects the broader argument for simply needing to know more about coaching across the literature; virtually every article reviewed noted the severe limitations in our currently understanding of the coaching role.

2.4 Discussion

Although this literature review reveals that the most salient characteristic of the coaching literature is the limited understanding of this complex role, the research does provide integral insights into coaching work. We know that coaches' work should be well-defined and bounded, that policy and instructional aims should be coherent and aligned, and that the ways principals and district administrators conceptualize and frame coaching can be very influential. We also know that coaches mediate policy messages, and their interpretations are influenced by many different factors. Coaches can be very effective in persuading teachers to change both superficial and deep elements of their practice, though the best ways to conduct, structure, and support this process remains unknown. Furthermore, coaching can be instrumental in shifting stubborn norms in schools by deprivatizing practice and facilitating collaboration. Coaches also have to balance multiple and often conflicting demands—they are expected to improve teaching through professional development and enact policy change through the way they conduct these professional learning activities. Coaching is also highly malleable and variable, depending on contextual demands and factors that are both seen and unseen; though there is a high degree of consensus in the coaching literature that context matters, an understanding of *how* different contextual factors shape and interact with coaching activities is still developing.

Situating coaching as a boundary spanning activity in a complex system may help to activate this existing knowledge in an effort to better describe coaching work—analyzing boundary spanning activities and practices within a broader understanding of systems in educational organizations could allow patterns of influence to emerge across variable contexts. The complexity of coaching and variability of school systems makes even descriptive claims about the nature of the role difficult to maintain. However, explicating system interactions and dynamics—such as, for instance, how a principal's framing of coach roles relates to coaches' policy negotiation processes—could bolster the field's understanding of coaching work.

One problematic characteristic of the extant coaching literature is the general framing of coaching interactions as one-directional; while coach learning is sometimes discussed in terms of coach training and development (e.g., Gallucci et al., 2010), investigations of coaching interactions are usually limited to the coach's delivery and mediation of policy and instructional messages. While some researchers have considered the role of responsiveness in professional feedback and teacher learning (e.g., Myung & Martinez, 2013), coaches' methods of learning and adapting in the boundaries are not addressed in the research. This also relates to the untapped potential of considering coaches as learning resources in systems. Coaches deliver and mediate information, but they also receive information, from teachers at the micro-level and district administrators (and sometimes, even state departments of education) in the macro-level. It seems likely that this multidirectional flow of information is another important factor in coaches' work. While the effect of this dynamic is unclear, considering the dialogical qualities of boundary-spanning coaching can be a meaningful frame for investigating this process.

Examining how coaches learn from their activities is also important for leveraging coaching positions as tools for collective improvement; systems remain viable, strong, and

responsive only if they remain flexible, and they remain flexible by continuously learning and absorbing information (e.g., Miller & Page, 2007). Coaches connect policy with practice through their boundary spanning work with teachers and administrators, so considering what coaches learn and how they adapt their activities to that information may help to clarify their function as change mechanisms. Variability in coach roles and responsibilities will pose a challenge in this respect, but clearly delineated information about contextual conditions and structures could alleviate this issue. Furthermore, coaches could also be examined as learning resources with regard to their capacity to feed information about teaching and learning back into the broader system. Reconsidering coach roles in this fashion would require careful rationalization and messaging, however, since trust has been identified as a key characteristic of innovating and improving schools (e.g., Bryk & Schneider, 2002), so it could be harmful if teachers perceived coaches to be informants. However, if situated in terms of collaborative learning and policy enactment, coaches could also be invaluable resources in terms of their potential to bridge the gaps between teachers, districts, and states.

Ultimately, coaching interventions hold great promise, but a more robust understanding of how and why these interventions work, and under what conditions, is necessary for reliably applying coaching to foster improvement and change. The research suggests that coaching interactions and activities still largely form a black box—the paucity of studies that account for what happens during coaching conversations indicates that additional detailed, situated studies of coaching practices are necessary for developing a robust understanding of its effects. Horn & Little's (2010) examination of teacher professional communities' "episodes of pedagogical reasoning" (p. 189) provides a blueprint for the kind of fine-grained discourse analysis that may be necessary to understand these dynamic interactions and their implications. This kind of work

will also be particularly data-intensive, since accounting for even one coach's practices requires deep understanding of the systems and subsystems that influence coaching activities. Furthermore, if understanding variation is indeed necessary for understanding the coaching role, comparative, situated accounts are needed in order to learn from that variation.

Given what the literature shows, it is clear that the field still lacks a body of research on coaching that is at once situated, descriptive, and focused on variation. Studies of coaching need to account for the complexity of coaching microprocesses and address the many interacting factors and forces that shape coaches work. This requires considerable data on coaching interactions as well as school, district, and state contexts. One thing to note about the extant literature on coaching is that few studies examine representations of one-on-one interactions between coaches and teachers over time; Coburn & Woulfin (2012) discuss the discursive details of coaching conversations, but the bulk of studies of coaching are based solely on interviews with coaches, principals, or district personnel. This suggests that, when designing research on coaching, extensive and multi-leveled data that accounts for macro-framing and micro-interactions could start to address the gaps in the literature.

2.5 Conclusion

This discussion of coaches as boundary spanners in complex systems represents a conceptual contribution to the field's nascent understanding of the role and function of coaches in policy implementation. While many scholars acknowledge coaches' conflicting roles and liminal positions, there is a surprisingly limited understanding of the significance of these attributes and their relationship to policy implementation, educational system structures, and implementation

outcomes. Investigating this more deeply offers not only a conceptual refinement in terms of how we understand and analyze coaching, but can also contribute to answering questions of system design and optimization in educational organizations.

Prominent scholars argue for integrity, rather than fidelity, of implementation in education (e.g., Bryk et al., 2015), suggesting that understanding how coaches accommodate and negotiate local systems as they enact policy initiatives is essential to understanding what works in coaching. Studying coaches' implementation of a reform initiative across a diverse set of contexts gives insight into specific coaching mechanisms that are effective, and also into environments that support effective coaching. Furthermore, understanding coaching in contexts traditionally viewed as difficult, particularly in urban schools, is a critical matter of educational equity, since coaches are disproportionately assigned to these schools, which are often already overburdened with improvement mandates and accountability pressures. The research design for this project specifically attends to contrasting cases, making variation a key object of study.

Finally, the substantial investment in coaching positions across the country suggests an urgent need for a deeper understanding of how coaching can be leveraged as a policy mechanism as well as an instructional support. As of now, the field lacks clarity on how to use coaches as policy brokerage *resources* in complex systems. As the role of coaches in policy dissemination and implementation becomes more pronounced with the rollout of the CCSS and other standards-based reforms, we need to understand their potential as learning resources and critical system actors. This is particularly vital in terms of learning how best to serve vulnerable students and communities.

2.6 References

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, *333*(6045), 1034-1037.

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, *81*(2), 132-169.

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

Atteberry, A., & Bryk, A. S. (2010). Centrality, connection, and commitment: The role of social networks in school-based literacy. *Social network theory and educational change*. (pp. 51-76). A. J. Daly (Ed.) Cambridge, MA: Harvard.

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays by MM Bakhtin* (M. Holquist, Ed.; C. Emerson & M. Holquist, Trans.). Austin, TX: University of Texas.

Bertrand, M. & Marsh, J. (2015). Teachers' sensemaking of data and implications for equity. *American Educational Research Journal*, *52*(5), 861-893.

Biancarosa, G., Bryk, A. S., & Dexter, E. R. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *The Elementary School Journal*, *111*(1), 7-34.

Blarney, K. L., Meyer, C. K., & Walpole, S. (2009). Middle and high school literacy coaches: A national survey. *Journal of Adolescent and Adult Literacy*, *52*(4), 310-323.

Blazar, D., & Kraft, M. A. (2015). Exploring mechanisms of effective teacher coaching: A tale of two cohorts from a randomized experiment. *Educational Evaluation and Policy Analysis, 37*, 542-566.

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.

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

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *The Elementary School Journal, 111*(3), 430-454.

Camburn, E. M., Kimball, S. M., & Lowenhaupt, R. (2008). Going to scale with teacher leadership: Lessons learned from a district-wide literacy coach initiative. In Mangin, M. M., Stoelinga,

S. R. (Eds.), *Effective teacher leadership: Using research to inform and reform* (pp. 120-143). New York, NY: Teachers College Press.

Clarke, A., & Collins, S. (2007). Complexity science and student teacher supervision. *Teaching and Teacher Education*, *23*, 160–172

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis, 23*(2), 145-170.

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

Coburn, C. E., Choi, L., & Mata, W. (2010). I would go to her because her mind is math": Network formation in the context of a district-based mathematics reform. In A. J. Daly (Ed.), *Social network theory and educational change*, (pp. 33-50). Cambridge, MA: Harvard University Press.

Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational Evaluation and Policy Analysis, 30*(3), 203-235.

Coburn, C. E., Russell, J. L., Kaufman, J. H., & Stein, M. K. (2012). Supporting sustainability: Teachers' advice networks and ambitious instructional reform. *American Journal of Education, 119*(1), 137-182.

Coburn, C. E., & Woulfin, S. L. (2012). Reading coaches and the relationship between policy and practice. *Reading Research Quarterly, 47*(1), 5-30.

Daly, A. J., Finnigan, K. S., Jordan, S., Moolenaar, N. M., & Che, J. (2014). Misalignment and perverse incentives: Examining the politics of district leaders as brokers in the use of research evidence. *Educational Policy*, 0895904813513149.

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

Domina, T., Lewis, R., Agarwal, P., & Hanselman, P. (2015). Professional sense-makers instructional specialists in contemporary schooling. *Educational Researcher*, 0013189X15601644.

Fligstein, N., & McAdam, D. (2012). *A theory of fields*. Oxford, UK: Oxford University Press.

Foster, D., & Noyce, P. (2004). The mathematics assessment collaborative: Performance testing to improve instruction. *Phi Delta Kappan, 85*(5), 367.

Fusarelli, L. D. (2002). Tightly coupled policy in loosely coupled systems: Institutional capacity and organizational change. *Journal of Educational Administration, 40*(6), 561-575.

Galey, S. (2016). The evolving role of instructional coaches in us policy contexts. *The William and Mary Educational Review, 4*(2), 11.

Gallucci, C., Van Lare, M. D., Yoon, I. H., & Boatright, B. (2010). Instructional coaching building theory about the role and organizational support for professional learning. *American Educational Research Journal, 47*(4), 919-963.

Goldstein, J. (2007). Debunking the fear of peer review: Combining supervision and evaluation and living to tell about it. *Journal of Personnel Evaluation in Education, 18*(4), 235–252. doi:10.1007/ s11092-006-9022-3

Holland, J.H. (2014). *Signals and boundaries: Building blocks for complex adaptive systems.* Cambridge, MA: MIT.

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher, 33*(8), 16-30.

Horn, I. S., Kane, B. D., & Wilson, J. (2015). Making Sense of Student Performance Data Data Use Logics and Mathematics Teachers' Learning Opportunities. *American Educational Research Journal*, *52*(2), 208-242.

Horn, I. S., & Little, J. W. (2010). Attending to problems of practice: Routines and resources for professional learning in teachers' workplace interactions. *American Educational Research Journal, 47*, 181–217.

Huguet, A., Marsh, J. A., & Farrell, C. C. (2014). Building Teachers' Data-Use Capacity: Insights from Strong and Developing Coaches. *Education Policy Analysis Archives*, *22*(52), 2-31.

Innes, J. E., & Booher, D. E. (2010). *Planning with complexity: An introduction to collaborative rationality for public policy*. New York, NY: Routledge.

Knight, J. (2004). Instructional coaches make progress through partnership: Intensive support can improve teaching. *Journal of Staff Development, 25*(2), 32–37.

Lave, J., & Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge, UK: Cambridge University Press.

Little, J. W., & Horn, I. S. (2007). Normalizing problems of practice: Converting routine conversation into a resource for learning in professional communities. *Professional learning communities: Divergence, depth, and dilemmas*, 79-92.

Mangin, M. (2009a). To have or not to have? Factors that influence district decisions about literacy coaches. *Literacy Coaching Clearinghouse*.

Mangin, M. M. (2009b). Literacy coach role implementation: How district context influences reform efforts. *Educational Administration Quarterly, 51*(2). 179-213.

Mangin, M. M., & Dunsmore, K. (2015). How the framing of instructional coaching as a lever for systemic or individual reform influences the enactment of coaching. *Educational Administration Quarterly*, 0013161X14522814.

Mangin, M. M., & Stoelinga, S. R. (Eds.). (2008). *Effective teacher leadership: Using research to inform and reform* (pp. 77-98). New York, NY: Teachers College Press.

Marsh, J. A., McCombs, J. S., & Martorell, F. (2010). How instructional coaches support data-driven decisionmaking policy implementation and effects in Florida middle schools. *Educational Policy*, *24*(6), 872-907.

Matsumura, L. C., Garnier, H. E., & Resnick, L. B. (2010). Implementing literacy coaching: The role of school social resources. *Educational Evaluation and Policy Analysis, 32*(2), 249-272.

Matsumura, L. C., Garnier, H. E., & Spybrook, J. (2012). The effect of content-focused coaching on the quality of classroom text discussions. *Journal of Teacher Education, 63*(3), 214-228.

Matsumura, L. C., Garnier, H. E., & Spybrook, J. (2013). Literacy coaching to improve student reading achievement: A multi-level mediation model. *Learning and Instruction, 25,* 35-48.

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life.* Princeton, NJ: Princeton University Press.

Mudzimiri, R., Burroughs, E. A., Luebeck, J., Sutton, J., & Yopp, D. (2014). A look inside mathematics coaching: Roles, content, and dynamics. *Education Policy Analysis Archives*, *22*(53), 1-32.

Murray, S., Ma, X., & Mazur, J. (2009). Effects of peer coaching on teachers' collaborative interactions and students' mathematics achievement. *The Journal of Educational Research, 102*(3), 203-212.

Myung, J., & Martinez, K. (2013). *Strategies for enhancing the impact of post-observation feedback for teachers.* Carnegie Foundation brief. Retrieved from: http://www.carnegiefoundation.org/resources/publications/strategies-enhancing-impact-post-observation-feedback-teachers/

Neufeld, B., & Roper, D. (2003). *Coaching: A strategy for developing instructional capacity, promises and practicalities.* Washington D.C.: The Aspen Institute.

Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal, 46*(2), 532-566.

Obera, S., & Sloan, M. (2009). The evolving role of a mathematics coach during the implementation of performance standards. *The Professional Educator, 33*(2), 1-33.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376-407.

Penuel, W. R., Riel, M., Krause, A., & Frank, K. A. (2009). Analyzing teachers' professional interactions in a school as social capital: A social network approach. *Teachers College Record, 111*(1), 124-163.

Porter, R. E., Fusarelli, L. D., & Fusarelli, B. C. (2015). Implementing the Common Core: How educators interpret curriculum reform. *Educational Policy, 29*(1), 111-139.

Powell, D. R., Diamond, K. E., Burchinal, M. R., & Koehler, M. J. (2010). Effects of an early literacy professional development intervention on head start teachers and children. *Journal of Educational Psychology, 102*(2), 299.

Russell, J. L., & Bray, L. E. (2013). Crafting coherence from complex policy messages: Educators' perceptions of special education and standards-based accountability policies. education policy analysis archives, 21, 12. Retrieved 2 November 2013, from http://epaa.asu.edu/ojs/article/view/1044

Sailors, M., & Price, L. R. (2010). Professional development that supports the teaching of cognitive reading strategy instruction. *The Elementary School Journal, 110*(3), 301-322.

Louis, K. S., & Leithwood, K. (1998). From organizational learning to professional learning communities. In K. Seashore Louis & K. Leithwood (Eds.) *Organizational learning in schools* (pp. 275-285). New York, NY: Taylor & Francis.

Smith, M.S. & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum & testing*. *Politics of Education Association Yearbook* (pp.233-267). Washington, DC: Falmer Press.

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.

Stoelinga, S. R. (2010). At the intersection of principal and teacher instructional leadership: The case of Donaldson Elementary. *Journal of Cases in Educational Leadership, 13*(2), 21-28.

Van Keer, H., & Verhaeghe, J. P. (2005). Comparing two teacher development programs for innovating reading comprehension instruction with regard to teachers' experiences and student outcomes. *Teaching and Teacher Education, 21*(5), 543-562.

Walpole, S., & McKenna, M. C. (2004). *The literacy coach's handbook: A guide to research-based reform.* New York, NY: Guilford.

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *21*(1), 1-19.

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

Woulfin, S. (2014). Charting the research on the policies and politics of coaching. *Education Policy Analysis Archives*, *22*, 50.

Woulfin, S. L., & Rigby, J. G. (2017). Coaching for coherence: How instructional coaches lead change in the evaluation era. *Educational Researcher*, *46*(6), 323-328.

3.0 Dissertation Paper 2:

Coaching in Context: Exploring Conditions That Shape Instructional Coaching Practice

Accepted as part of *Teachers College Record* special issue, forthcoming 2019

3.1 Introduction

Educational policy initiatives are increasingly relying on instructional coaches to improve classroom teaching in the standards-reform era (Domina et al., 2015; Neufeld & Roper, 2003; Woulfin & Rigby, 2017). Coaches are versatile professional development actors in education, in that they can fill a variety of instructional support and policy implementation roles (e.g., Coburn & Woulfin, 2012). Though the details of their roles vary, instructional coaches generally aim to develop teachers' expertise in content areas and mediate policy as it is translated into the classroom, simultaneously mentoring individual teachers and enacting collective reforms (Coburn & Woulfin, 2012; Domina et al., 2015; Knight, 2004; Mangin & Dunsmore, 2015). Despite the recent proliferation of the coaching role, and the optimism about its potential to improve teaching and learning, the immense variability across coaching positions, coupled with the heterogeneity of state, district, and local school organizations, pose conceptual and methodological challenges for researchers analyzing coaching and its effects.

Coaches are also central actors for *crafting coherence* from multiple, often conflicting, policy messages in educational organizations, a critical process that has been explored in the implementation literature (e.g., Honig & Hatch, 2004; Penuel, Riel, Krause, & Frank, 2009; Russell & Bray, 2013; Spillane et al., 2002). This process involves intensive sensemaking, a

sociological concept that describes the active construction of meaning through individual and social processes of cognition, negotiation, organization, and interpretation (Coburn, 2001; Weick, 1995). Due to their inherent proximity to instructional sensemaking, connections to building and district leadership, and capacity building roles, coaches are uniquely positioned to impact policy uptake and instructional coherence.

Coaches' inherent connectedness also highlights another key aspect of their work—that it is shaped by the structure of the complex organizational system in which the coaching takes place. The extensive variation in these contextual conditions suggests that coaching will work differently in different sorts of school and district environments. Educational researchers have established numerous ways in which organizational context shapes implementation (e.g., Datnow et al., 2002; Spillane et al., 2002), and studies of education reform efforts and specific instructional coaching initiatives suggest that organizational and social factors are especially critical to these processes (Coburn & Russell, 2008; Woulfin & Jones, 2018; Coburn, Russell, Kaufman, and Stein, 2012). This work suggests that coaches' embeddedness in networks of both social and technical resources matters to understanding the effect that they have on policy change and instructional improvement. This is an important research and design consideration as instructional coaches become increasingly common mechanisms of reform implementation and improvement in school districts across the country.

So, given that nested and multilayered social, technical, and organizational factors shape the practice and impact of instructional coaching, understanding the combinations of factors that are most consequential is a key question for both policy implementation and instructional coaching research. Accordingly, in this study we sought to understand the interconnected contextual conditions that interact with instructional coaching practice to shape its effect on teaching and

learning. To do this, we consider coaching practice as inherently imbricated with the context in which it takes place, and we make the social, technical, and organizational systems in which coaching work is embedded key foci of our study. Our findings support the argument that coaching work cannot be extricated from the surrounding system factors that shape it; on the contrary, we found that supports for and barriers to robust coaching practice are context-specific phenomena, suggesting that any coaching intervention ought to be tailored to its specific organizational location. We build on the findings of other researchers (e.g., Kraft et al., 2017) who posit that leveraging coaching for implementation and improvement efforts requires a deep understanding of *how* coaching plays out within a particular system, and argue that different combinations of contributing factors can affect coaching in different ways, thereby contributing to the theory that a nested systems perspective on instructional coaching is key to understanding its effects.

3.2 Conceptual Framework

3.2.1 Instructional Coaching

Building on Coburn & Woulfin (2012), who found that coaches play both educative and political roles in schools, this study explores the complexity of coaching work and the many conditions that either support or constrain it. Coaching activities take place in a boundary between state and district-level policy and classroom instruction, and they are extremely heterogeneous; one district may use coaches primarily to mentor novice teachers, while another may use a literacy coach to build will to adopt a new schoolwide initiative , a third may focus coaching efforts on achievement data use and interpretation, and yet another may dedicate a mathematics coach to

only intermediate grades to build instructional capacity in that specific content area—in short, there are endless configurations of coaching responsibilities. Coaches report feeling more professionally supported when their roles are well-defined and there is alignment between district priorities and coaching work (Mangin, 2009a; Mangin & Dunsmore, 2015), suggesting that developing knowledge about organizational coaching supports that are most effective and meaningful could have important practical implications in terms of facilitating coaching work across contexts.

While the field has yet to reach a consensus on the definition of effective, quality instructional coaching, an emerging evidence base suggests that coaching can facilitate change and improvement in teaching and learning. Studies of coaching effects suggest that, under certain conditions, coaching has had a positive impact on high leverage outcomes, like instructional quality and student learning, particularly when it is content-focused (Allen et al., 2011; Biancarosa, Bryk, & Dexter, 2010; Blazar & Kraft, 2015; Bryk, et al., 2015; Campbell & Malkus, 2011; Foster & Noyce, 2004; Matsumura et al., 2010; Matsumura, Garnier, & Spybrook, 2012; 2013; Neumann & Cunningham, 2009; Powell et al., 2010; Sailors & Price, 2010). Although the literature shows the promise of coaching interventions, definitive causal findings about coaching are elusive, since many studies lack critical details about organizational context, implementation processes, and coaching models or frameworks (Biancarosa, Bryk, & Dexter, 2010; Blazar & Kraft, 2015). Researchers often note that variability in coaching program design and implementation as a critical challenge in this effort (Campbell & Malkus, 2011; Mangin & Dunsmore, 2015; Mudzimiri et al., 2014; Obara & Sloan, 2009), and the distance between coaching work and student outcomes poses a significant measurement challenge (Blazar & Kraft, 2015; Mangin, 2009a; 2009b).

Despite these challenges, coaching programs and interventions continue to proliferate, and the research highlights the promising potential of instructional coaching under specific conditions. In a meta-analysis of 60 causal studies of coaching, Kraft and colleagues (2017) found significant positive effects on teaching and more modest positive effects on student achievement. However, they also noted many remaining questions about the nature of quality coaching, observing that their ability to analyze specific coaching programs, teacher and coach backgrounds, and other factors was constrained by a lack of information across these studies. This study also reiterates a now common imperative in the coaching literature to consider the social, professional, and organizational conditions that shape coaching work and any associated outcomes (Biancarosa, Bryk, & Dexter, 2010; Blazar & Kraft, 2015). Nevertheless, the overarching idea that coaching can create robust instructional and achievement gains is compelling, and researchers have shown that well-placed and supported coaches have the capacity to bolster student achievement, deepen teachers' understanding of their work, create opportunities for deep and meaningful professional interactions, accelerate uptake of new programs and instructional practices, and support data-driven decision making (Campbell & Malkus, 2011; Coburn & Russell, 2008; Coburn, Russell, Kaufman, & Stein, 2012; Coburn & Woulfin 2012; Huguet et al., 2014; Mangin & Dunsmore, 2015; Marsh et al., 2009; Matsumara et al., 2010).

3.2.2 Contextualizing Coaching Practice

The literature shows that one constant in coaching is the high degree of variation in coach roles and contexts (e.g., Coburn & Woulfin, 2012; Galey, 2016; Marsh et al., 2008; Mangin & Dunsmore, 2015), suggesting that understanding coaching practice requires also understanding role structures and organizational conditions that shape coaches' work. Research on both

instructional coaching and instructional change is increasingly attending to the organizational and social factors that shape, facilitate, and constrain the complex web of interactions that comprises coaching work (Bryk & Schneider, 2002; Coburn, Mata, & Choi, 2010; Coburn & Russell, 2008; Coburn, Russell, Kaufman, & Stein, 2012; Opfer & Pedder, 2011; Spillane et al., 2002; Woulfin & Rigby, 2017; Woulfin, 2018; Woulfin & Jones, 2018). This literature points to the complex systems and social networks that shape critical elements of coaching practice, emphasizing the interdependency of organizational processes, and the critical importance of considering not only the outcomes of implementation, but also the contextual factors that shape them.

In a study of teachers' social networks, Coburn and colleagues (2010) propose a "theory of organizational embeddedness" that is critical to understanding the flow of resources and relationships between people in public schools, noting that these organizations are "situated in complex policy and institutional environments that are likely to penetrate the technical core of schooling" (p. 313). Further, the complexity of the systems in which coaches are embedded is a key concept in analyzing and understanding coaching work and its function in schools and districts; as Bryk and colleagues (2015) have argued, performance is a product of the interactions between people, tools, materials, and processes that compose a complex system, and these systems may react unpredictably to attempts to change them (Miller & Page, 2007). Opfer and Pedder (2011) also argue that understanding teacher learning, which can include instructional coaching processes, requires accounting for the many interdependent forces that shape outcomes.

Furthermore, in navigating variable school and district contexts, coaches must also navigate variable policy environments, where, as schools respond to mounting pressures to adopt policies that improve teaching and learning, they are often tasked with connecting state- and district-level policies with teacher practice (Woulfin, 2014). In this sense, coaches occupy a crucial

boundary space in school systems, where meaning and enactment are negotiated and organizational change is produced (Akkerman & Bakker, 2011; Fligstein & McAdam, 2012). Given the complexity of school's organizational contexts, analyzing the activities in which coaches engage as they occupy these liminal spaces should also take structural constraints and facilitators into account. Consequently, understanding coaching requires understanding both specific practices, the structural contexts in which those practices are embedded, and varied implementation processes across system levels and diverse district contexts. As such, we built this study on the foundational idea that coaching work inherently, continuously interacts with the complex systems in which it takes place, and that understanding and analyzing coaching requires attending to the organizational and social structures shape it. Accordingly, this paper focuses on the following research question: *What contextual conditions facilitate and constrain coaching, and why?*

Given that analyzing coaching practice requires understanding these contextual factors, our analytical approach focuses on using comparative case logic to uncover patterns and themes in the relationships between coaching practice and the systems in which it takes place across a variety of diverse coaching cases. We designed our study to attend to the *quality* of coaching practice, which we call coaching rigor, which is a composite construct that reflects the quality of coaching conversations with teachers. To approximate the relationship between coaching rigor and coaching context, we also considered salient contextual factors at multiple system levels that emerge from the coaching literature. These include *macro level system factors,* such as concentrated poverty and racial composition at the district level, since the literature has established that inequitable conditions in schools influence the implementation processes and outcomes (e.g., O'Day & Smith, 2016). We also consider district-level factors that are specific to coaching, in particular *district*

instructional priorities and *district performance pressures*, since previous research on coaching has identified district-level norms as important supports for coaching initiatives (Coburn & Russell, 2008; Gallucci et al., 2010; Marsh, McCombs, & Martorell, 2010). Additionally, since the coaching literature has also posited that school-level factors and role construction can be consequential for coaching outcomes (e.g. Mangin 2009a), we explore these considerations by examining the details of coach role definitions and coaches' reports of the tasks and duties that detract from their time spent on coaching activities. Our exploratory analysis of the patterns across these key constructs from the literature also led us to consider social and structural factors that emerged as additional key factors across our cases, particularly coaches' access to tools, expertise, and supportive colleagues, as well as district instructional coherence.

3.3 Methods and Data

To understand these constructs and answer our research question, we investigated how a group of coaches implemented a mathematics coaching model, focusing specifically on the relationship between *coaching contexts* and *coaching practice.* To understand this dynamic, we used data from a large scale, three-year coaching project, where partners from a state department of education and a university-based research and development team built and refined a mathematics instructional coaching model to support teaching for conceptual understanding aligned with rigorous college- and career- ready mathematics standards. The coaches we studied spent two years learning and enacting a coaching framework that specified three key coaching practices: conducting deep and specific conversations about instruction, establishing both mathematical and pedagogical goals, and providing evidence-based feedback. In the first two years

of the project, professional developers trained 32 coaches in 21 districts in this coaching model, which coaches then implemented in their home schools and districts. These district contexts varied widely and provided an excellent opportunity to understand how coaching works under different contextual conditions.

3.3.1 Case Study Design

Our research design aimed to systematically capture the interplay of coaches' practice with their school and district contexts. To do this, we conducted a case study that drew on both traditional case based qualitative methods (Yin, 2003; Miles, Huberman, and Saldana, 2013) and qualitative comparative analysis, a mixed method grounded in set theoretical logic (Ragin, 2009; Rihoux and Ragin, 2008). The goal of this analysis is both exploratory and descriptive: exploratory in the sense that we aimed to unearth themes and patterns to illuminate the interplay of coaching with the complex organizational systems it inhabits, and descriptive in the sense that we use our data to illustrate themes that emerge across both individual and group level cases.

3.3.1.1 Qualitative Comparative Analysis

We used qualitative comparative analysis (QCA) as a tool to explore the relationships between coaches' contexts and their coaching practice. QCA is particularly well suited to exploring coaching because it allows us to make systematic comparisons between complex, multivalent cases to develop theories about how coaching works under different conditions (Rihoux & Ragan, 2008). By examining the interactions between coach contexts and coaching activities, we leverage these comparative cases to learn about how coaches negotiate different kinds of situations.

QCA is a relatively novel method in educational research, though it has been used in several other studies where researchers aimed to understand the complexity of educational environments (e.g., Coburn, Russell, Kaufman, and Stein, 2012; Trujillo & Woulfin, 2016). QCA examines the relationship between sets of conditions (which one might refer to as independent variables in a quantitative study) and how those conditions combine to lead to a given outcome (or not) (Ragin, 2006a; 2006b; Rihoux & Ragin, 2009). Given QCA's ability to describe and investigate different combinations of conditions, this method is particularly well-suited to exploring our research question about how complex coaching contexts influence coaching practice. QCA is also helpful because it works best with a small to medium number of cases (often between 8 and 150), allowing the researcher to analyze a larger amount of cases than would be possible in traditional qualitative designs (Ragin, 2006a; 2006b).

There are two types of QCA analysis: crisp set, wherein conditions (or, variables, in common research language) and outcomes are all designated as being in or out of a given set, and so conditions and outcomes are translated to binary code, and fuzzy set, which we primarily used in this study, where conditions and outcomes are coded with interval scales to represent the extent to which they are part of a group (Ragin, 2006a; 2006b). So, fuzzy sets involve degrees of membership in a category; for example, within the category of "satisfied customers" in a study of consumer satisfaction at Target, a customer who rated their local Target store with enthusiastic praise in a customer survey might receive a score of 1, as they are fully in the set of satisfied customers; whereas a customer who left negative feedback in open-ended responses and gave the store low-to-average scores in the survey might receive a score of .2, as they are more out of the group than within it, but showed slight satisfaction in their average responses; and a third customer who gave weak praise in open-ended questions but marked average survey responses might receive

a .6, as they are more in the set than out, but they did not show enthusiastic satisfaction in the manner of the first customer, and their middling survey responses suggested ambivalence. The researcher must use their understanding of the data and relevant theories to select set conditions and develop scales, which is an intensive and iterative process (Ragin, 2006a; 2006b).

3.3.1.2 Outlier Analysis

Social scientists have argued that even when QCA does not highlight one neat set of causal pathways, the method can also help researchers systematically understand, compare, and learn about cases. When tidy causal explanations do not emerge from a QCA analysis:

> non-conforming cases may be even more important than typical cases as they can be of great help in the understanding of causal complexity. Unlike in other research methods, where researchers neglect deviant cases or outliers as unavoidable nuisance, QCA takes into consideration even a combination of conditions that explains only a single case. (Berg-Schlosser et al., 2009, p. 1465)

Existing QCA outlier analysis techniques are based in Yin's (2003) replication logic, a tool for selecting further cases for comparison to enhance a case study's potential for making causal claims (Nair & Gibbert, 2016). Based on this idea, strategies for holistically comparing QCA cases using theoretical replication logic include two types: 1) comparison of cases that show the same conditions, but different outcomes, and 2) comparison of cases with the same outcomes but different conditions. In the first type, comparisons help to highlight factors that may contribute to the disparities in the outcome. In the second type, comparisons help to explore alternative explanations and build stronger theories (Nair & Gibbert, 2016). We focus on these comparative strategies in our interpretation of our QCA results.

This approach to QCA was particularly powerful for us, since our results highlighted many pathways and interactions between conditions rather than one neat set of clean, causal pathways. The purpose of this kind of holistic comparative method is to sharpen the understanding of the plausibility of causal relationships between variables, expand the existing theory, or even develop new theories about the phenomenon in question (Nair & Gibbert, 2016). Accordingly, the analysis we discuss here is intended to refine existing theories about how to support coaches across complex and varied contextual conditions, and to build new theory around how coaching works in widely divergent circumstances.

3.3.1.3 Comparative Case and Group Level Analysis

To further explore emerging theories about the interplay between coaching and its contexts, we built on our QCA outlier analysis with an additional wave of case based qualitative analysis (Yin, 2013; Miles, Huberman, & Saldana, 2013). We maintained the comparative, case-based logic using traditional qualitative methods, distilling additional qualitative data into coach-level case summaries, which we organized by broad themes embedded within system levels (e.g., facilitators of coaching at the school level, barriers to coaching at the district level, etc.). Then, we conducted a cross case analysis to develop second pass codes. We then used those second pass codes to organize a matrix, where codes were formed rows and groups of coaches formed the columns. This helped us comparatively identify patterns in the data and consider the patterns and other emerging indicators from our QCA analysis in additional detail (Bush-Mecenas and Marsh, 2018; Miles, Huberman, & Saldana, 2014).

3.3.2 Study Context and Participants

For four years preceding the coaching study, Tennessee's state Department of Education took ambitious actions to create a common vision of instructional excellence, which set the foundation for coaching work statewide. However, state-provided training mostly focused on teachers and principals, leaving a support gap for instructional coaches, despite the state's use of Race to the Top funds to dramatically increase the number of coaches across the state. Because of this need, the Tennessee Department of Education (TDOE) was interested in building and exploring instructional coaching programs. Given this interest, TDOE partnered with researchers and professional developers from the University of Pittsburgh and the Institute for Learning to build, execute, refine, and study the coaching model that we discuss here.

Coaches were selected in a competitive process which included written applications and scenario-based coach interviews. Through this process, the project partners selected 32 coaches from 21 districts across Tennessee to participate in the study. These coaches were strategically selected to show variation within a set of desired parameters, including sufficient mathematics knowledge and opportunities to engage in intensive one-on-one coaching. We intentionally selected coaches that met these requirements but also represented diversity in coaching experience and capacity as well as contextual conditions. All of the accepted coaches were asked to change their existing practice substantially to participate in the project coaching model. At the start of year 1, each of these 32 coaches selected two partner teachers who taught mathematics in grades 3 through 8 to participate in the study. For year 2, each coach was asked to select 1 new partner teacher and keep 1 of their 2 original partner teachers.

3.3.3 Data

The dataset we used allowed us to explore the nature of coaches' interactions with teachers as well as the contextual conditions that shape these conversations. These data represent multiple levels of the systems in which coaches work, so they position us to understand the dynamics that influence coaching, which is key to gaining knowledge about how coaching functions in complex systems. The following are the main data sources that provide insight into the coaches' contexts:

- Coach context surveys. The Coach Context Survey was administered to coaches at the end of project years 1 and 2 and included items measuring perceived support for coaching in the school/district, coaching role definition, and previous teaching and coaching experience. The survey provides a standardized way to assess coaches' perception of key dimensions of context. See measures section for more detail.
- Coach context essays. At the outset of years 1 and 2 of the project, coaches were asked to write a 2- to 3-page essay in which they described their coaching role and were asked to reflect on their school and district context. These essays were collected from all coaches in the study. The essay provides additional insight into coaches' perceptions of the context factors that shape their coaching practice.
- Partner teacher interviews. At the end of year 1 of the project, we conducted semi-structured interviews with the coaches' partner teachers. The protocol contained questions about teachers' experiences with their coach, the extent to which coaching influenced their teaching, the relational dynamics between the coach and teacher, and the school and district context in which the coaching took place. The interviews provide an additional triangulating perspective on coaches' contexts and their work with the partner teachers.

Transcripts of coaching conversations are the primary data sources for examining coaching practice:

- Videotaped coaching conversations. Each coach was asked to videotape coaching conversations with each of their two partner teachers during 3 coaching cycles in year 1 and 2 cycles in year 2. Each cycle included 2 coaching conversations: a pre-lesson planning conference and a post-lesson feedback conference. The conference lengths ranged from 10 minutes to over an hour. All video recordings were transcribed. The transcripts provide a close up look at the micro-processes of coaching.

3.3.4 Data Limitations

Due to the design of the broader project, we did not have access to any data that directly captured the point of view of school or central office leaders. While we could approximate the leadership environment from teacher and coach perceptions, leadership approaches are critical components of coaching uptake and implementation. Understanding these perspectives is an important consideration for future studies of coaching systems and should be a key methodological consideration in future studies.

3.3.5 Measures

To conduct a QCA analysis, researchers must identify an outcome and sets of conditions that may contribute to that outcome. This is an iterative and intensive process and is heavily shaped by the data, existing theory, and the research questions being pursued. Below, we describe the construction of the outcome and the conditions that undergird our QCA analysis.

3.3.5.1 Outcome Measure

Coaching rigor

In this QCA, our outcome is rigorous coaching, which serves as a proxy for quality of coaching interactions. Theoretically, coaching quality, as a latent construct, encompasses a variety of things, some of which are fairly clear cut (e.g., content accuracy) whereas others are more difficult to pin down (e.g., rapport between coach and teacher). In an earlier analysis activity for this project, we coded coaching interactions during pre-lesson planning conferences based on depth and specificity of their conversations, and aggregated those codes to create coaching rigor scores. Deep and specific conversations were those that took up substantive discussion of mathematical content, pedagogy, and student reasoning about mathematics. Then, these coaching rigor scores were used in growth models tracking the improvement of coached teachers, and we found that this aggregate rigor score was predictive of the rate of teaching growth over time. Specifically, teachers who worked with coaches with higher rigor scores had steeper teaching growth trajectories than teachers who worked with coaches with lower rigor scores. However, it is important to note that teachers who worked with the coaches with lower rigor scores also improved, but at a more modest rate than the teachers who worked with the more rigorous coaches. This concept of rigorous coaching practice is a core element of the coaching model that is the backbone of this coaching project, and it shaped the training that we provided to coaches and refined over time (see also Russell et al., 2018, under review).

To generate these rigor scores, we averaged coaching rigor scores across 5 timepoints over the course of 2 years for 3 partner teachers per coach. We averaged the coaching scores for the partner teachers that each coach worked with because there was limited variation between teachers working with the same coach, and due to inconsistency in the time that each coach worked with

each teacher (i.e., coaches selected two partner teachers in year 1, and then kept 1 of those partner teachers constant and chose a new one in year 2). We thought the aggregate score would be a robust way of representing the coach's practice (i.e., this is a measure of coaching practice across time and teacher). Across our 32 coaches, these rigor scores were distributed relatively normally across the sample. Given this distribution and our sample size, we decided to divide the distributed scores into quartiles to create a simple four level scale for this outcome measure, which allowed us to avoid becoming overly fine-grained in our analysis, categorize coaches based on the distribution of rigorous practices, and to create continuity with a quartile grouping strategy that has been useful in other models for this project. Also, because of the exploratory nature of this analysis, this scaling is largely descriptive and meant to build additional theory around how to measure and understand complex concepts like coaching quality or rigor. So, to scale this rigor outcome score, we used this common four level set scale, that provided easy to understand categories anchored in the distribution of coach rigor (Ragin, 2006a). The lowest level is coded as 0 to signify being out of the set (i.e., not showing rigorous coaching practices), and the medium-low measure is coded as .33 to signify being more out of the set than in. Medium-high is set at .67, and coaches in the high category received a score of 1. We classified the coaches in the medium-high and high categories as showing a more rigorous practice over time, and coaches in the low and medium-low levels as engaging in a less rigorous practice over time.

For this outcome measure, it is important to note that our sample restricted the range of our analysis of coaching, since coaches across the rigor scale helped teachers improve, but at different rates. So, we did not gain insight into contextual conditions that did not allow for uptake of the coaching model. Instead, the rigor scale allowed us to analyze the degree to which coaches were able to engage in the uptake of these rigorous conversations. We treat this as a limitation in our

study, although we also argue that our approach to analyzing the relationship between coaching rigor and coaching conditions offers a starting point and a foundation for future studies of these phenomena. Further research would benefit from examining situations where coaching was not taken up at all to understand the conditions that pose insurmountable barriers for the uptake of quality coaching practices. That said, this study can offer signals for where to start those investigations of a broader range of coaching uptake.

3.3.5.2 District-Level Conditions: Context Descriptors

Poverty rates

High rates of student poverty and a lack of district resources are perennial challenges in educational improvement initiatives; as such, a socioeconomic status condition was critical to include in this QCA analysis. We used Tennessee's state-level mean rate of economically disadvantaged students as an anchor to create the scale for this condition. Statewide, Tennessee's average rate of economically disadvantaged students is 35.1%. It is important to note that, unlike many other states, Tennessee's economic disadvantage measure is not wholly based on federal free and reduced lunch rates, as one usually sees in state accountability measures. Tennessee recently changed the definition of economically disadvantaged students to students who are "directly certified to receive free lunch without an application," which includes: students who receive SNAP benefits, families participating in the TANF program, students who are homeless, in Head Start, migrants, runaways, foster children, or otherwise certified as economically

disadvantaged by state/local officials.[2] As such, this is a much narrower group than those who receive free or reduced lunch. For the QCA scale, we translated the percentage of economically disadvantaged students in a district to an interval scale (from 0-1) using the state average rate of economically disadvantaged students and measuring how many standard deviations each district was from the state mean.

District racial demographics (from Tennessee Department of Education data)

We know that schools serving students of color and poor communities are most likely to be assigned coaches (Domina et al., 2015), so understanding how this iteration of the coaching intervention fits into these communities is critical to understanding its quality. Accordingly, we include a condition that signals the extent to which a district serves students of color. This condition in the QCA analysis is a fuzzy interval scale, based on the percentage of students of color served by the district.

District alignment with coaching model (survey data) and district performance evaluation pressures (survey data)

Pulled from two coach context survey items ("My district's priorities align with the Math Instructional Coaching Model" and "My district will judge my coaching performance by the value-added scores of the teachers with whom I work."), these two alignment conditions signal concepts that are important in the coaching literature and so we wanted to know how this fit into the formula

[2]https://gallery.mailchimp.com/b28b453ee164f9a2e2b5057e1/files/ED_Definition_for_Accountability_FAQ__01.pdf

of key ingredients for coaching quality. Both conditions were constructed as a fuzzy set to reflect the survey responses (strongly agree, agree, disagree, strongly disagree).

3.3.5.3 School Based Conditions: Coaching Distractions and Role Characteristics

School based coaching role (survey data)

We wanted to create a condition that would show whether a coach is based in the district office, which usually means that the coach works across the districts' schools, or is based in an individual school building, which can still signal that a coach works with a variety of teachers, but they occupy a different kind of organizational position than district-based coaches. As such, we created this condition as a crisp set. Either the coach is school based (1) or they are not (2).

Conditions that detract from time spent on coaching responsibilities

We know that coaches have to navigate a variety of conflicting demands, so including the contextual conditions that distract coaches is an important part of this analysis, and can give us clues about the most salient organizational factors that constrain (or facilitate) coaching. We used coaches' survey responses about contextual conditions to construct fuzzy sets, scoring "disagree" and "strongly disagree" responses as 0, "agree" responses as 0.8, and "strongly agree" responses as 1. We constructed the following conditions based on these survey items that asked coaches to rate the extent to which the following factors detracted from the time they spent coaching:

- Managing response to intervention (RTI) data
- Test prep
- Administrative duties
- District responsibilities

Initially, we also planned to use urbanicity as a condition, based on the NCES locale category designations (rural, fringe; rural, remote; town, distant; town, remote; suburb, small; suburb, large; city, small; city, midsize; city, large). However, in early stages of the QCA analysis, we found that coaches in urban and suburban districts were so different than coaches from rural and remote districts that they warranted separate groups based on urbanicity. We collapsed these categories into two separate sets: suburban and urban coaches, and rural and remote coaches. The rural and remote set included all rural and town classifications, and the urban and suburban set included all suburban classifications as well as, small, midsize and large cities. We made these classification choices because we are interested in comparing across the large, complex environments that researchers typically point to when they talk about urban education—small cities signal a smaller degree of complexity than their midsize and large counterparts, so we categorized them with suburban schools, because they are mostly within the classification of urban. As for the rural and remote category, small remote towns tend to face the same kinds of environmental constraints as schools we think of as rural in that they are small, often resourced-strapped systems, so they are also in the rural category and are not counted as in the urban set for this condition. So the 32 coaches in our sample were divided into two different urbanicity groups based on the urbanicity classification of their district, and separate QCA analyses were conducted for each group.

- Urban and suburban coaches (n=21)
- Rural and remote coaches (n=11)

3.4 Analysis

We used fsQCA software for our analysis, and built models using the conditions detailed above. For each sample (urban and suburban coaches; rural and remote coaches) we ran two separate QCA models, one based on the district-level context descriptors and the other based on school-level time pressures (described above). We separated the analysis into these two distinct groups of conditions because a broader QCA analysis that included all of the conditions contained too much variation for the QCA program to run. Additionally, for each subsample we ran QCA analyses for the presence of rigor and lack of rigor (in QCA logic parlance, ~RIGOR) to see if there were any conditions either led to coaching rigor or that seemed to be significant obstacles to rigorous coaching practices.

Using the fsQCA software, we developed truth tables (one for each of the eight models described above) that displayed all of the combinations of conditions contained in each case, and whether and to what degree each set of conditions led to the outcome of interest (high or low levels of coaching rigor, depending on the model) (Ragin, 2006b). This allowed us to use the software to generate solution terms, which show each combination of conditions that lead to the outcome in question.

After generating the QCA solution terms and finding that they did not lead to a neat set of causal pathways, we built case ordered matrices, where coach cases formed the rows and solution paths formed the columns, to holistically compare the QCA cases and draw cautious conclusions about the role of contextual factors in coaching practice. We used these emergent insights to structure our second analytical pass, employing traditional qualitative methods to dive more deeply into our case data and build more nuanced theories about the interplay between coaching rigor and coaching context. We started this process with within-case, coach level analysis, constructing a

case summary for each coach by organizing qualitative data from partner teacher interviews and coach context essays under leveled thematic categories: facilitators for rigorous coaching at the school and district level, and barriers to rigorous coaching at the school and district level. Next, we looked across these case summaries to find themes that emerged across coaches and explore those themes within and across each urbanicity group. We used these emerging themes to distill case summaries into a cross case matrix, where cross case themes (or codes) formed the rows and coach urbanicity and rigor categories formed the columns.

The QCA analyses generated many pathways that led to coaching with more or less rigor (Tables 3.1, 3.2, 3.3, 3.4, and 3.5). For most of the QCA results, coverage scores, which approximate how much of the sample is covered or explained by a given outcome, were low, suggesting that there were many paths to the desired outcome (in this case, coaching rigor) (Ragin, 2006a; 2006b). On the other hand, consistency scores, which gauge the relationship between the solution path and the outcome in question, were often quite high, suggesting the empirical relevance of the pathways that did emerge (Ragin, 2006a; 2006b; Trujillo & Woulfin, 2013). Given the proliferation of pathways, we approached our analysis as a systematic, holistic comparison method, since the QCA did not produce a definitive set of causal solutions. Furthermore, it is important to note that this QCA analysis generated a wide variety of pathways for both samples of coaches, signaling that coaches can do rigorous work under a wide array of contextual conditions. As such, these findings are part of a broader case study approach. This is in line with techniques for using QCA as a tool for systematic comparison and investigation as part of a broader case-based analysis strategy (Rohlfing & Schneider, 2013; Schneider & Rohlfing, 2016). Below, we share results from our QCA analysis, which are organized by the two groups of coaches that we investigated: urban and suburban coaches, and rural and remote coaches. Tables 3.1, 3.2, 3.3, 3.4,

and 3.5 show the solution terms that the QCA analysis identified, the coverage and consistency scores for each solution term, and the number of coaches included in each pathway.

Table 3.1 Truth Table Solutions for Conditions That Support High Coaching Rigor – Urban & Suburban Coaches – District Level Factors

Pathway (Solution Term)	Coverage	Consistency	# Coaches included
High percentage minoritized students x Value added performance evaluation pressure	.64	.78	14
High percentage minoritized students x District priority alignment with coach program x District deployment of coaches NOT aligned with coach program	.29	.89	3
District priority alignment with coach program x District deployment of coaches not aligned with coach program x No value-added performance evaluation pressure x No economic disadvantage	.17	.71	2

Table 3.2. Truth Table Solutions for Conditions That Lead To Low Coaching Rigor – Urban & Suburban Coaches – District Level Factors

Pathway (Solution Term)	Coverage	Consistency	# Coaches included
No performance evaluation x No economic disadvantage	.45	.80	3
No performance evaluation x District deployment of coaches aligned with coach program	.45	.74	4

Table 3.3. Truth Table Solutions for Conditions That Support High Coaching Rigor – Urban & Suburban Coaches – School Level Factors

Pathway (Solution Term)	Coverage	Consistency	# Coaches included
No administrative duties x No district responsibilities x Not school based	.07	.72	1
No administrative duties x No district responsibilities x Too many teachers	.17	.80	1
Administrative duties x No RTI duties x District responsibilities x Too many teachers	.10	1	1

Pathway (Solution Term)	Coverage	Consistency	# Coaches included
Administrative duties x District responsibilities x Too many teachers x School based	.23	.88	4
No administrative duties x No RTI duties x District responsibilities x No test prep responsibilities x School based	.06	.80	1

Table 3.4. Truth Table Solutions for Conditions That Support High Coaching Rigor – Rural And Remote Coaches – School Level Factors

Pathway (Solution Term)	Coverage	Consistency	#Coaches included
Test preparation responsibilities x Too many teachers assigned x School based	.4	.74	2

Table 3.5. Truth Table Solutions for Conditions That Lead To Low Coaching Rigor – Rural And Remote Coaches – School Level Factors

Pathway (Solution Term)	Coverage	Consistency	# Coaches included
RTI duties x District responsibilities x No test preparation responsibilities x Too many teachers x Not school based	.18	1	1
No RTI duties x District responsibilities x Test preparation responsibilities x Too many teachers x Not school based	.18	1	1
Administrative duties x RTI duties x District responsibilities x No test preparation responsibilities x Too many teachers	.13	1	1
Administrative duties x RTI duties x District responsibilities x Too many teachers x Not school based	.1	1	1
Administrative duties x District responsibilities x Test preparation responsibilities x Too many teachers x Not school based	.1	1	1

Following the QCA phase of our analysis, we used the case summaries and cross-case matrix discussed above to deepen and explore our theoretical conclusions about patterns across

our coach cases. Our case summaries highlighted several key ideas at play for coaches in both urban and suburban contexts: coherent instructional systems, autonomy and flexibility in coaches' roles, access to outside coaching expertise and professional development, opportunities to collaborate with peers, school and district-level administrator support, and coaching programs and tools at the district level. We synthesized these ideas with the emerging themes from the QCA results to develop the following theories about how coaching rigor interacts with school and district contextual factors.

3.5 Findings

3.5.1 Contextual Factors Influencing Urban & Suburban Coaches

3.5.1.1 District Level: Accountability for Teacher Performance, Coherent Instructional and Coaching Systems

For urban and suburban coaches, district contextual factors appear to have more explanatory power than school-level distractions and time constraints. District level performance evaluation pressures in particular seemed to have significance as a contextual factor that usually leads to coaching rigor in urban and suburban schools. In our QCA results for district level factors, both solution paths that led to coaching that was less rigorous included a lack of performance evaluation pressure (see Table 3.2). This performance evaluation condition was present for 10 of the 14 coaches with a rigorous practice in this sample, and 4 of the 7 coaches who lacked rigor in their coaching, suggesting that it may also have a positive association with rigorous coaching practices (see Table 3.1 and Figure 3.1). Of those 4 coaches who lacked rigor in their practice

despite performance evaluation pressures, 1 coach was near retirement and very superficially engaged in the coaching project, so the relationship between their district context and their rigor is likely not very strong. Another of these 4 coaches was in their first year as a math coach in their school, where coaching had only existed for 3 years total. These results suggest that a lack of value added performance evaluation pressure in itself is not enough to lead to a lack of rigor in coaching, however, it does seem likely that it is an important factor in how coaches approach their work. These patterns signal that a press for performance improvement at the district level that is tied to teacher performance may support quality coaching work.

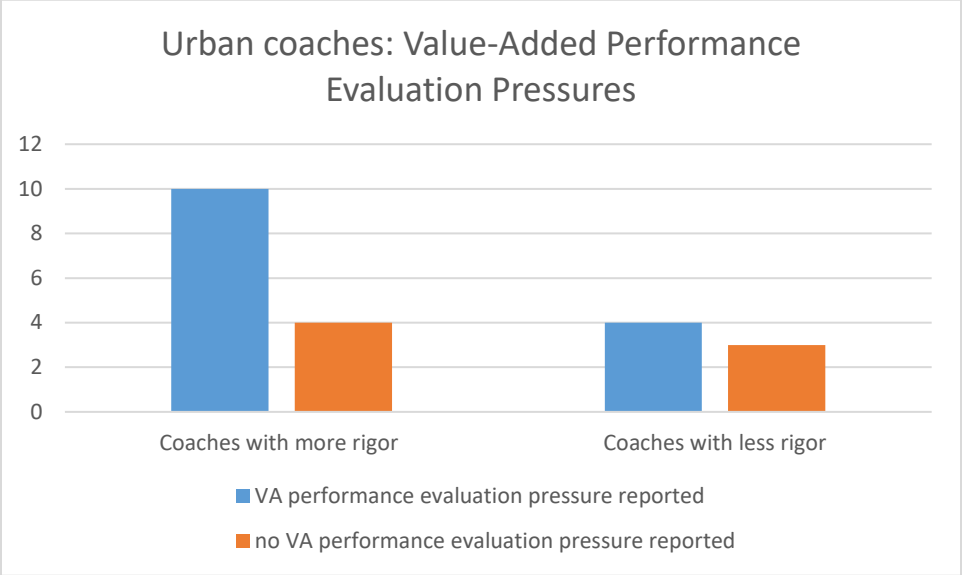

Figure 3.1. Urban and Suburban Coaches – Survey Responses: Value Added Performance Evaluation Pressures

Upon further inspection, this element of coaching context appears to slightly more complicated. First, it is important to note that this survey item reflected coaches' perception of how they would be evaluated, so some coaches within the same district perceived this differently. The 3 coaches who strongly disagreed with the statement "My district will judge my coaching performance by the value-added scores of the teachers with whom I work" all fell into the rigorous

practice category. Two of these coaches were in urban/suburban districts, and 1 was in a rural/remote district. It is important to note, however, that the urban/suburban coaches in this group had district colleagues in the study who answered this question differently, suggesting that these districts actually do press coaches to improve their teachers' value-added scores, or at least communicate that stance in some way. However, this idea of a press for performance improvement was present in nearly all of the coaches' discussions of their context in annual reflective essays about their work: the vast majority of all coaches in both urban and rural groups noted that standardized test scores were critical in their districts and central to the ways in which their teachers' performance and improvement would be evaluated by district leadership.

Our qualitative exploration of coaches' and teachers' description of their context suggested that another important district level factor that differentiated coaches with higher rigor from those with lower rigor, particularly in the urban and suburban group, was district-level instructional coherence and messaging. Nine of the thirteen coaches who had higher rigor in the urban and suburban set talked about vertical instructional alignment, opportunities to shape instructional messages, and/or coherent district wide instructional philosophies as key factors in their coaching practice. For instance, this is how one more rigorous coach wrote about the process of maintaining vertical instructional alignment across their partner teachers' schools in their annual context essay:

> both principals will meet with my supervisor on a regular basis so she can keep them informed [about] the curriculum and pacing guide and what we are to provide teachers in our building… I meet with both groups of administrators to map out a plan for both schools and discuss what teachers they want me to support with coaching. My teachers know that what I tell them about their instruction is what they will hear from the administration.

Of the five coaches in this set who showed less rigorous coaching practices, only one coach noted that there was standards-based instructional alignment in their district. This suggests that a degree of coherence in instructional messaging at the system level may be a key support for rigorous coaching practices, and clear district-level performance improvement expectations are likely part of a district-level instructional philosophy.

A third important district factor that emerged from the qualitative exploration of our coaches' discussions of their contexts was the structure of the coaching system in the district itself, and whether or not a coaching program, framework, or philosophy was present at the district level. Only one coach in the group that showed less rigorous practices noted an established coaching program at the district level; the other four talked about highly variable roles and a decentralized coaching system in their districts, though several did note that a coaching model was superficially adopted by the district office. On the other hand, nearly half (6) of the higher rigor coaches in urban and suburban schools noted a centralized coaching approach or philosophy at their district. One higher rigor coach talked about the evolution of the coaching program in their district, which had been in place, in one way or another, since 2005. Two other coaches in our project also came from this district, and both engaged in more rigorous coaching practices. Interestingly, another higher-rigor coach actually wrote about the adoption of district coaching program between years 1 and 2 of the project: "my role has changed tremendously since last year due to the addition of a coaching program… coaches are expected to be the catalyst for the change our school system desperately needs." Their description of the degree of change is another piece of evidence that supports the impact of having a district-level coaching system. This may indicate that the absence of a coherent coaching framework poses an additional challenge for coaches in large, complex systems.

3.5.1.2 Autonomy to Negotiate School-Level Conditions

For school level conditions in this group, there were many different pathways that emerged from QCA models (see Table 3.3), suggesting that school-level conditions interact in variable ways to support or hinder rigorous coaching. Some patterns challenged assumptions that we made about contextual conditions that support coaching rigor. For instance, a group of 4 coaches who engaged in rigorous practice over time fell into a solution path that combined 3 conditions that might be considered distractions or competing demands that may limit time for coaching (administrative duties, district responsibilities, and too many teachers assigned) with a school-based role. In 3 of the 5 solution paths that led to rigor, having too many teachers to coach was present, suggesting that coaches found adaptive ways to overcome this challenge. Our data suggests that at least some of these coaches were actively working to find ways to maximize their time during the course of the project; in fact, half of the coaches (2 out of 4) who fell into these solution paths were also using an inquiry tool (the plan, do, study, act cycle) to find better ways to manage their time.

Our qualitative analyses offered additional insight into this pattern that suggested that school level factors mattered less for urban and suburban coaches; one important idea that emerged from these coaches' context descriptions was that of flexibility and autonomy in their role. Over half of the more rigorous coaches in this group indicated that they had decision-making power in their role that allowed them to navigate challenges as they saw fit. For instance, one of these more rigorous coaches noted "there is a high level of trust given to me by the administration" and several others in this high rigor group talked about having the freedom to choose which teachers they work with and how they engage with them. This may explain why school-level distraction factors were less important to this group of coaches if their roles provided enough flexibility to negotiate these issues and creatively address them. Further, in the group of coaches who demonstrated less

rigorous coaching practices, two of the five described very rigidly proscribed and compliance oriented role structures, and another noted that coaches are "owned by the school," and so their role is determined by the principal, who was not bought in to standards-based instruction. While these are only emerging signals, they may suggest that coaches may benefit from some degree of autonomy to make informed and intentional decisions about how to navigate common contextual challenges.

3.5.2 Contextual Factors Affecting Rural & Remote Coaches

3.5.2.1 Decentralized Systems and Decreased District Significance

The QCA results for rural and remote coaches were significantly different than, and in some cases, nearly opposite from, what we saw for the urban and suburban sample, suggesting that necessary support conditions for coaching can vary considerably depending on environment. For instance, we noticed that evaluation pressure may be associated with higher quality coaching in urban and suburban schools. Almost all (4 out of 5) of the coaches with rigorous practices in the rural/remote sample did not report feeling value-added evaluation pressure in their work, whereas only 4 of the 14 coaches practicing with rigor in the urban/suburban sample reported a lack of value-added performance evaluation pressure (see Figure 3.2). So, evaluation pressures are likely working differently in different kinds of schools.

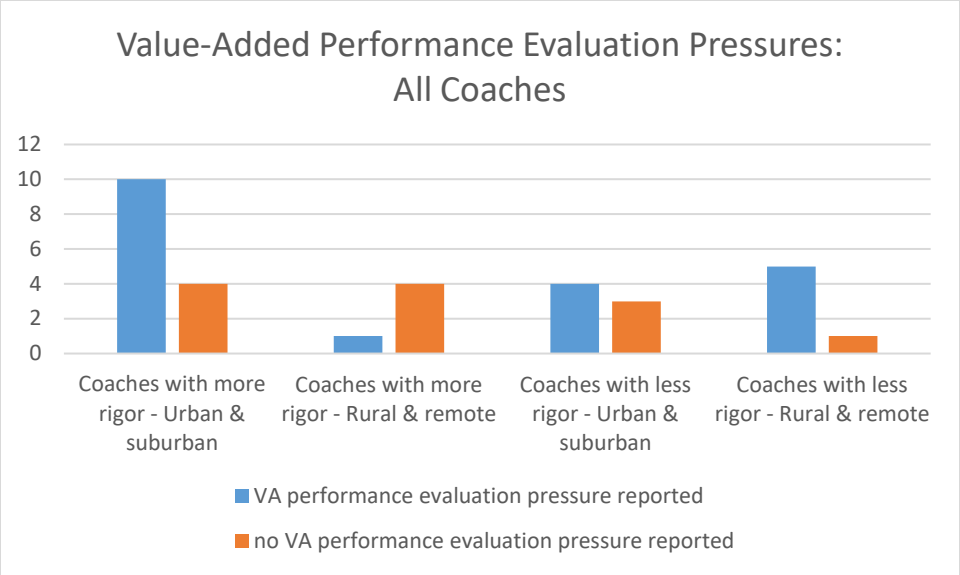

Figure 3.2. Value Added Performance Evaluation Pressures – Survey Responses: All Coaches

This pattern suggests that, in general, when districts, particularly urban districts, press coaches to improve the performance of their teachers, it may contribute to more quality coaching interactions, and perceived value-added evaluation pressures appear to be less consequential in rural and remote districts than they are in urban and suburban districts. This may also reflect the broader pattern of district factors being less important in rural and remote school systems in general.

If we view this performance improvement factor as embedded in a broader system of instructional messaging, our qualitative explorations of the context data offer further evidence to support this pattern. Cross-case analyses showed that few coaches in the rural and remote group worked within a system with clear, coherent, centralized instructional messaging; half of the higher rigor coaches in this set (n=3) noted that they emphasized coherent instructional messaging in their coaching practice, but none of these coaches talk about vertically aligned instructional messaging, which we often saw with the higher rigor coaches in the urban and suburban set. One coach, for instance, explained their messaging work in this way:

> The message I am communicating doesn't coincide exactly with the message other administrators are communicating about math instruction. However, most of the administration and supervisors in my district support and respect my opinion… Administrators/supervisors perceive that the way they were taught is the best or only way to teacher. Helping spread awareness of mathematical best practices is an issue in my district. District personnel do not get too involved in sending messages to teachers about math instruction.

This pattern may be a reflection of the more decentralized nature of dispersed rural and remote school systems, or it may show a lack of an instructional foundation that contributes to the practical difficulties that coaches in rural and remote districts face.

3.5.2.2 School Level Challenges, Supportive Colleagues, and Coach Role Structures

This notion that rural systems are more diffuse and so coaching is less structured by centralized mechanisms is also supported by other results of our analysis: for instance, according to our QCA results, in general, school level time constraints were more significant for rural and remote coaches than district level factors. Our initial truth table analyses of district-level factors for rural and remote coaches did not show strong enough patterns to move to the solution term stage (i.e., there were too few conditions that were associated with the outcome, coaching rigor, among this set of factors). However, for rural and remote schools, school-based time constraints showed clearer patterns than they did for urban and suburban schools, suggesting that the school environment may be more significant for coaches in small, more dispersed rural education systems.

This idea is also supported by the ways in which rural and remote coaches talked about district-level coaching resources and frameworks. Most of these coaches talked about a lack of tools, resources, or foundation for coaching at the district level, regardless of whether they

demonstrated high or low rigor in their coaching practice. However, many of the more rigorous coaches in these settings (n=4) talked about supportive *people* at the district level as key resources to their practice; three of these coaches noted that they collaborate with district level staff to structure their coaching and seek out coaching development resources beyond their district context. One coach in the higher rigor group explained this dynamic in the following way:

> To be honest, I am provided little guidance by my supervisors. They trust me to be in classrooms and to help my math teachers to the best of my ability. I depend greatly on the advice and help of the other math coaches in the county. They provide a great support system because we often face the same battles and go through the same struggles.

This coach's case suggests that opportunities for authentic collaboration with peers may help rural and remote coaches navigate the lack of centralized guidance in their school systems.

For rural coaches, not being school based**,** in combination with a variety of other school-level factors, was present in 4 out of 5 solution paths that lead to less coaching rigor in rural and remote districts. In other words, all but one of the coaches who enacted more rigorous practices in the rural/remote group were school based, while 4 out of 6 of the coaches who showed less rigor were not (see Figure 3.4). The one coach who was not school based but engaged in rigorous practice across time describes participating in targeted coaching for which they developed a plan collaboratively with district office staff and school administrators. They determined these groups of teachers to focus on based on performance scores from the previous year. So, while this coach faced many obstacles, this kind of targeted coaching may have helped to overcome the challenge of being spread thin in a rural or remote district. As with the urban and suburban coaches, rural and remote coaches may have an easier time navigating these challenging school level conditions if they have flexibility and decision-making power baked into their role; all of the more rigorous

coaches in this group signaled a degree of autonomy in their position by indicating that they collaborate with school leaders to shape their work (n=2), by stating that directly that there is flexibility and decision making power in their role (n=2), or by noting that they built the coaching position themselves, from scratch (n=1).

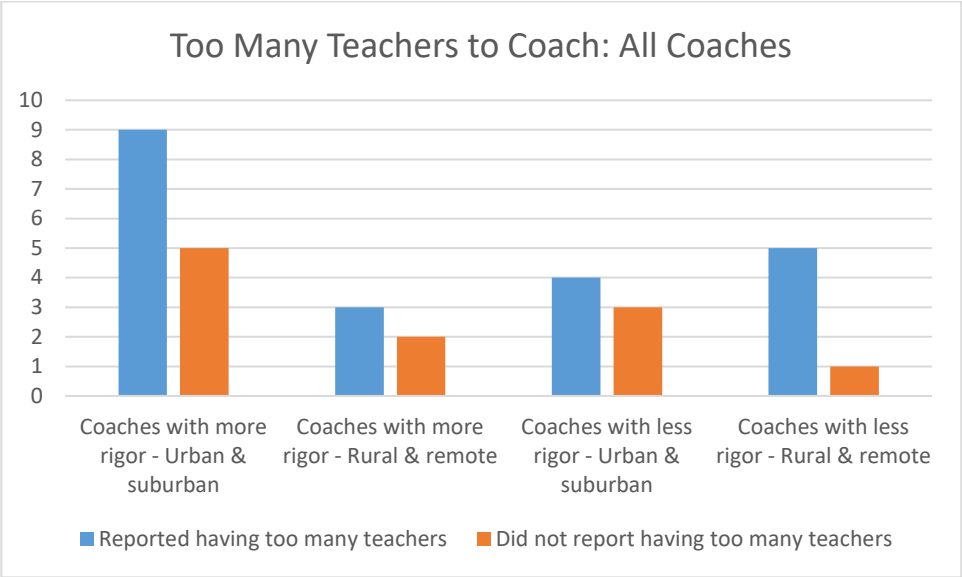

Figure 3.3. Too Many Teachers to Coach – Survey Responses: All Coaches

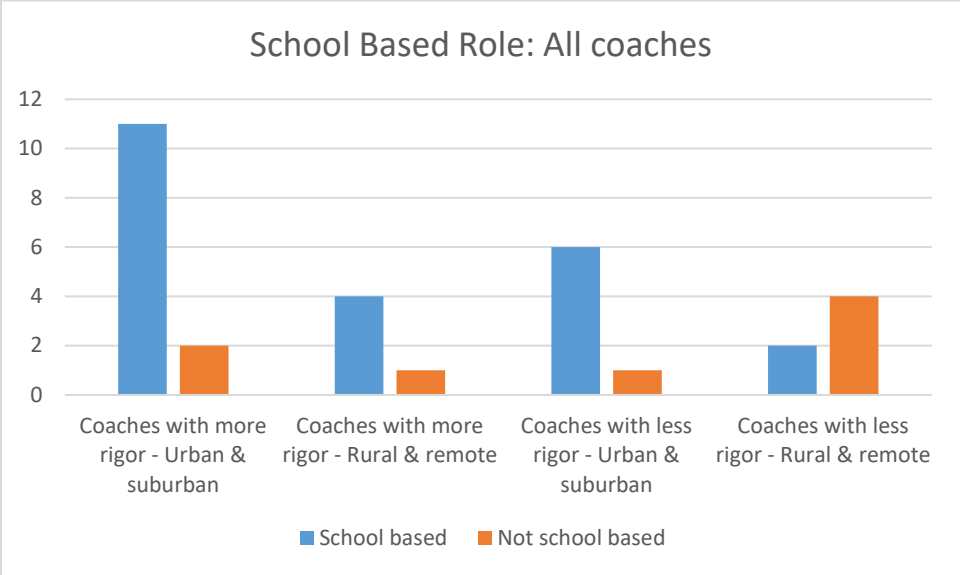

Figure 3.4. School Based Role – Survey Responses: All Coaches

For rural and remote coaches, having too many teachers assigned appeared in every school-level pathway to less rigorous coaching practices (see Table 3.5 and Figure 3.3). Four of the six coaches with less rigor fit into these pathways, and of the remaining 2, one of them reported the number of teachers they were assigned to as detracting from their time spent coaching (but they did not fit into any of the identified pathways). So, 5 out of 6 coaches who did not show rigor in their practice reported this condition of having too many teachers assigned to them in their work. The outlier in this group—the coach who showed lower levels of rigor but did not have too many teachers assigned to them—is also school-based, suggesting that some other condition led to their struggle with rigorous coaching. In other case study data, this coach describes a challenging environment where teachers do not engage in high cognitive demand work and resist the district tool for increasing teaching performance because they see it as "scripted." Their district was also struggling to move toward evidence-based teaching practices. One could imagine that this would make deep, specific, rigorous coaching quite difficult to enact. This coach also notes that their coaching position is new and was developed as part of a school improvement plan, suggesting that there may be many difficult factors in their school context.

3.5.2.3 Cross-Cutting Contextual Supports: Opportunities for Collaboration and Access to Expertise

While important differences distinguish the urbanicity groups from one another, a few commonalities that cut across both groups emerged from the qualitative analysis that followed our QCA work. In particular, the value of opportunities for collaboration and access to outside expertise was important for both rural/remote and urban/suburban coaches, though the patterns around these concepts did vary between the two groups.

For urban and suburban coaches, the majority of more rigorous coaches talked about cooperative relationships with colleagues, opportunities to collaborate with peers, and the availability of either district-based or outside professional development as important supports for their coaching practice. In contrast, most of the lower rigor coaches do not discuss the presence of these supports in their context; in fact, they more often talk about being monitored by their supervisors at the district level, or having access to training, but only inconsistently. For instance, one coach in the lower rigor group described decision-making about their coaching work in this way:

> I am expected to be able to coach all kindergarten through fifth grade teachers in all subject areas… our district has set forth this expectation, and it is upheld by the school principal… my building principal will also identify other areas or teachers that may need coaching.

The higher rigor coaches, on the other hand, note regular opportunities for district-level professional development, access to state and regional coach trainings, and opportunities to independently pursue professional learning and conferences. Over half of these rigorous coaches (n=8) talked about opportunities to collaborate with other coaches, or to collaborate with school administrators and/or district staff about their coaching work. For example, one higher-rigor coach wrote about meeting with the math curriculum specialist and all other school-level math coaches in their district about once a month to go on "instructional rounds" their buildings together and learn from what they saw. Almost none of the less rigorous coaches in urban and suburban schools mentioned these factors.

For rural and remote coaches, a similar pattern emerged from our qualitative analysis; all of the coaches who showed more rigorous practice across time talked about either a supportive

and/or collaborative relationship with either a school level administrator or a district based staff member, and most talked about working with other coaches or similar professionals (e.g., instructional supervisors) as important supports for their work. For example, when writing about the supervisor of federal programs and supervisor of instruction at their district, one coach noted: "we work together as a team to make sure we are supporting each other and sending out a unified message from our office." Only two of these rigorous rural and remote coaches talked about reliable access to professional development or outside professional learning. Conversely, most of the coaches who demonstrated less rigor in their practice talked about a degree of isolation in their work, often because there were too few peers in their school system with whom to collaborate. These coaches almost never talk about supportive colleagues at any level of their system as resources, which is a stark contrast to their counterparts in the more rigorous group.

3.6 Discussion

The patterns we discuss provide some compelling signals about the relationship between coach context and coaching practice, and some clues about system design principles that could facilitate robust coaching work. First of all, we saw significant differences in the influence of contextual conditions on coaching between our two urbanicity categories (rural and remote schools vs. urban and suburban schools). This discrepancy suggests that coaches in different kinds of systems require different kinds of conditions and structures to feel supported in their work and do their jobs well. While there are certainly some problems that may always make coaching difficult, we think that these findings indicate that there is no one-size-fits-all approach to coaching, and

that context ought to be a key design principle for implementing coaching interventions and programs.

While too many distractions and many teachers to work with are likely challenges for most coaches, the QCA results suggested that this obstacle was much more difficult to overcome for coaches in rural and remote districts. This may be attributable to geography; coaches in rural districts likely have more distance to cover when they travel between schools to coach many different teachers, for instance. Another factor here is that, since rural districts are less populous, rural educators have many different kinds of responsibilities, and a large coaching load may place more strain on a rural coach who wears multiple hats. This could be compounded by the comparative lack of flexibility and autonomy in rural coaches' roles; without decision making power in their role, they may have fewer opportunities to negotiate these challenging school level conditions than their peers in larger urban and suburban systems. The layered interactions between all these factors emphasize the necessity of taking a systems view of instructional coaching, following Opfer & Pedder's (2011) argument that teacher learning, too, cannot be understood without also accounting for the complex system in which it takes place. Furthermore, while the implementation literature has explored teacher autonomy and how it structures and constrains learning interactions in schools (e.g., McLaughlin & Talbert, 2001), this idea has not been investigated deeply in studies of coaching. Our findings suggest that this is an important dimension of coaching to consider, and future research would do well to attend to this notion.

We also found that some urban and suburban coaches facing the conditions we most readily classify as difficult—particularly having too many teachers to coach and other demands on their time—appear to be able to adapt to negotiate these circumstances. Since these coaches, as a group, appear to have more autonomy in their roles and more opportunities to collaborate with peers,

those supports may help to balance out the burden of having too many responsibilities, since they are more likely than rural and remote coaches to be able to shape and co-construct their work. Although district level factors appear to be less important for rural and remote coaches, the general lack of centralized coaching resources and programs for coaches in these systems, coupled with the lower level of rigor in rural and remote coaches overall, may suggest that the absence of autonomy to adapt to their circumstances is a hindrance to their coaching work. This is an important finding for the adaptation literature, which has shown previously that implementation processes are more complex than finding promising interventions and asking practitioners to implement them (Coburn, 2003; Elmore, 1996); rather, implementation research has found the processes of "mutual adaptation" tend to characterize educational reforms, wherein both the innovation itself and educators' work practices and routines likely undergo adjustments to facilitate uptake across contexts (Berman & McLaughlin, 1978; Coburn, 2004; Fishman, Penuel, Allen, Chang, & Sabelli, 2013; Honig, 2006).

Despite these differences between coaching in different context, there were also supports for coaching that appeared to matter regardless. In particular, the emergence of collaboration and access to expertise as universal positive factors for these two groups indicates that social networks may be just as important for coaches as they are for teachers (Coburn & Russell, 2008; Coburn, Russell, Kaufman, & Stein, 2012). This also offers insight into design considerations for coaching roles; while coaches are sometimes framed as independent change agents, this finding suggests that, in fact, they benefit from access to collaboration and social capital as much as any other educational professional. Since coaches are already positioned as critical brokers in their system, and their role is inherently cooperative, this coaching support is often overlooked in the literature. However, this study suggests that collaboration and professional learning are critical elements of

coaching practice, and may actually facilitate coaches' efforts to overcome other challenging conditions that are often present in school and district contexts.

3.7 References

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333, 1034-1037.

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, *81*(2), 132-169.

Berg-Schlosser, D., De Meur, G., Ragin, C., & Rihoux, B. (2009). Qualitative comparative analysis (QCA) as an approach. In B. Rihoux & C. Ragin (Eds.), *Configurational Comparative Methods*. Thousand Oaks, CA: Sage.

Berman, P., & McLaughlin, M. (1978). *Federal programs supporting educational change: A model of education change, Vol. VIII: Implementing and sustaining innovations*. Santa Monica, CA: Rand.

Biancarosa, G., Bryk, A., & Dexter, E. R. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *Elementary School Journal*, *3*(1), 7-34.

Blazar, D., & Kraft, M. A. (2015). Exploring mechanisms of effective teacher coaching: Results from two cohorts of an experimental evaluation. *Educational Evaluation and Policy Analysis* 37, 542-566.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. (2015). *Learning to improve*. Cambridge, MA: Harvard Education Press.

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

Bush-Mecenas, S., & Marsh, J. A. (2018). The DIVE approach: Using case-ordered meta-matrices and theory-based data displays to analyze multiple case study data. In *Complementary research methods for educational leadership and policy studies* (pp. 33-56). Palgrave Macmillan.

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *The Elementary School Journal*, *111*(3), 430-454.

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis, 23*(2), 145-170.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3-12.

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

Coburn, C. E., Mata, W. S., & Choi, L. (2013). The embeddedness of teachers' social networks: Evidence from a study of mathematics reform. *Sociology of Education*, *86*(4), 311-342.

Coburn, C. E., Russell, J. L., Kaufman, J. H., & Stein, M. K. (2012). Supporting sustainability: Teachers' advice networks and ambitious instructional reform. *American Journal of Education, 119*(1), 137-182.

Coburn, C.E., & Russell, J.L. (2008). District policy and teachers' social networks. *Educational Evaluation and Policy Analysis. 30*(3), 203-235.

Coburn, C. E., & Woulfin, S. L. (2012). Reading coaches and the relationship between policy and practice. *Reading Research Quarterly, 47*(1), 5-30.

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

Domina, T., Lewis, R., Agarwal, P., & Hanselman, P. (2015). Professional sense-makers instructional specialists in contemporary schooling. *Educational Researcher*, 0013189X15601644.

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

Fishman, B. J., Penuel, W. R., Allen, A. R., Cheng, B. H., & Sabelli, N. O. R. A. (2013). Design-based implementation research: An emerging model for transforming the relationship of research and practice. *National Society for the Study of Education*, *112*(2), 136-156.

Fligstein, N., & McAdam, D. (2012). *A theory of fields*. Oxford, UK: Oxford University Press.

Foster, D. & Noyce, P. (2004). The mathematics assessment collaborative: Performance testing to improve instruction. *Phi Delta Kappan, 85*(5), 367-374.

Galey, S. (2016). The evolving role of instructional coaches in US policy contexts. *The William and Mary Educational Review, 4*(2), 11.

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

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher,33*(8), 16-30.

Huguet, A., Marsh, J. A., Farrell, C. C. (2014). Building teachers' data-use capacity: Insights from strong and developing coaches. *Education Policy Analysis Archives*, *22*(52), 1-28.

Knight, J. (2004). Instructional coaches make progress through partnership: Intensive support can improve teaching. *Journal of Staff Development, 25*(2), 32–37.

Kraft, M. A., Blazar, D., & Hogan, D. (2016). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 0034654318759268.

Mangin, M. (2009a). To have or not to have? Factors that influence district decisions about literacy coaches. *Literacy Coaching Clearinghouse*.

Mangin, M. M. (2009b). Literacy coach role implementation: How district context influences reform efforts. *Educational Administration Quarterly, 51*(2). 179-213.

Mangin, M. M., & Dunsmore, K. (2015). How the framing of instructional coaching as a lever for systemic or individual reform influences the enactment of coaching. *Educational Administration Quarterly*, 0013161X14522814.

Matsumura, L. C., Garnier, H. E., Resnick, L. B. (2010). Implementing literacy coaching: The role of school social resources. *Educational Evaluation and Policy Analysis*, *32*(2), 249-272.

Matsumara, L. C., Garnier, H. E., & Spybrook, J. (2012). The effect of content-focused coaching on the quality of classroom text discussions. *Journal of Teacher Education*, *63*(3), 214-228.

Matsumura, L. C., Garnier, H. E., & Spybrook, J. (2013). Literacy coaching to improve student reading achievement: A multi-level mediation model. *Learning and Instruction*, *25*, 35-48.

Marsh, J. A., McCombs, J. S., & Martorell, F. (2010). How instructional coaches support data-driven decision making policy implementation and effects in Florida middle schools. *Educational Policy, 24*(6), 872-907.

McLaughlin, M. W., & Talbert, J. E. (2001). *Professional communities and the work of high school teaching*. Chicago, IL: University of Chicago Press.

Miles, M. B., Huberman, A. M., & Saldana, J. (2013). *Qualitative data analysis*. Sage.

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life*. Princeton, NJ: Princeton University Press.

Mudzimiri, R., Burroughs, E. A., Luebeck, J., Sutton, J., & Yopp, D. (2014). A look inside mathematics coaching: Roles, content, and dynamics. *Education Policy Analysis Archives,22,* 1–32.

Nair, L. B., & Gibbert, M. (2016). Analyzing inconsistent cases in Management fsQCA studies: A methodological manifesto. *Journal of Business Research, 69*(4), 1464-1470.

Neufeld, B., & Roper, D. (2003). *Coaching: A strategy for developing instructional capacity, promises and practicalities.* Washington D.C.: The Aspen Institute.

Neuman, S. B., Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal*, *46*(2), 532-566.

Obara, S., & Sloan, M. (2009). The evolving role of a mathematics coach during the implementation of performance standards. *Professional Educator*, *33*(2), n2.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376-407.

Penuel, W. R., Riel, M., Krause, A., & Frank, K. A. (2009). Analyzing teachers' professional interactions in a school as social capital: A social network approach. *Teachers College Record, 111*(1), 124-163.

Powell, D. R., Diamond, K. E., Burchinal, M. R., Koehler, M. J. (2010). Effects of an early literacy professional development intervention on Head Start teachers and children. *Journal of Educational Psychology*, *102*, 299-312.

Rihoux, B., & Ragin, C. C. (2008). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Vol. 51). Thousand Oaks, CA: Sage.

Ragin, C. C. (2006a). Set relations in social research: Evaluating their consistency and coverage. *Political Analysis, 14*(3), 291-310.

Ragin, C. C. (2006b). User's Guide to Fuzzy-Set/Qualitative Comparative Analysis 2.0. Tucson, Arizona: Department of Sociology, University of Arizona.

Russell, J. L., & Bray, L. E. (2013). Crafting coherence from complex policy messages: Educators' perceptions of special education and standards-based accountability policies. education policy analysis archives, 21, 12. Retrieved 2 November 2013, from http://epaa.asu.edu/ojs/article/view/1044

Russell, J.R. et al. (under review). Learning to leverage coaching for instructional coaching at scale: Adaptive integration in the Tennessee math coaching project. *American Educational Research Journal.*

Rohlfing, I., & Schneider, C. Q. (2013). Improving research on necessary conditions: Formalized case selection for process tracing after QCA. *Political Research Quarterly*, 220-235.

Sailors, M., & Price, L. R. (2010). Professional development that supports the teaching of cognitive reading strategy instruction. *The Elementary School Journa*l, *110*(3), 301–322.

Schneider, C. Q., & Rohlfing, I. (2016). Case studies nested in fuzzy-set QCA on sufficiency: formalizing case selection and causal inference. *Sociological Methods & Research, 45(*3), 526-568.

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.

Trujillo, T. M., & Woulfin, S. L. (2014). Equity-oriented reform amid standards-based accountability: A qualitative comparative analysis of an intermediary's instructional practices. *American Educational Research Journal, 51*(2), 253-293.

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

Woulfin, S. (2014). Charting the research on the policies and politics of coaching. *Education Policy Analysis Archives, 22*, 50.

Woulfin, S. L., & Jones, B. (2018). Rooted in relationships: An analysis of dimensions of social capital enabling instructional coaching. *Journal of Professional Capital and Community*, *3*(1), 25-38.

Woulfin, S. L., & Rigby, J. G. (2017). Coaching for coherence: how instructional coaches lead change in the evaluation era. *Educational Researcher, 46*(6), 323-328.

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

4.0 Dissertation Paper #3:

Mutual Adaptation for Continuous Improvement: Exploring Adaptive Integration of Improvement Science In Educational Organizations

4.1 Introduction

Over the past decade, continuous improvement research has garnered considerable attention from education policy researchers as well as government and private foundation funders. This growing interest has in part been spurred by the idea that learning "what works" in a particular educational organization is not enough, since the heterogeneity and variability of the education system in the United States necessitates specific, contextualized solutions to pressing policy problems. So, while randomized control trials and other classical intervention approaches remain popular, government officials and researchers alike are increasingly exploring methods for generating flexible, "sticky" implementation knowledge that supports scaling efforts across multiple, diverse contexts (Cohen-Vogel et al., 2015; Park et al., 2013; Redding et al., 2017; Tichnor-Wagner et al., 2017; 2018). Continuous improvement research has the potential to build this kind of knowledge through a focus on rapid, systematic inquiry cycles that develop the learning capacity of organizations. However, engaging in continuous improvement work is complicated and challenging, and often requires major cultural shifts within organizations. In education in particular, continuous improvement methods challenge entrenched norms around data and transparency, and can place significant capacity building demands on schools and districts.

An important aspect of continuous improvement research is the idea of *adaptation,* which is critical to building effective, scalable solutions to pressing educational quality and policy

problems. Implementation research has established that improving education is not a one-directional process of identifying good practices and enacting them in schools and districts; rather, a dynamic of "mutual adaptation" emerges from implementation efforts, wherein practitioners negotiate changes in their practice to accommodate interventions and also adjust the intervention to fit it to their context (Berman & McLaughlin, 1978; Coburn, 2004, Cohen et al., 2008; Honig, 2006; Russell et al., under review*a*; Tichnor-Wagner et al., 2018; Supovitz, 2008). Networked improvement communities (NICs) in particular center this process by encouraging *adaptive integration,* which encourages stakeholders to use systematic methods to learn their way into fitting a novel practice or process into their unique context (Hannan et al., 2015; Russell et al., under review*a*; under review*b*; Tichnor-Wagner et al., 2018). NICs bring many different stakeholders together to combine the rigor of improvement science methods with the power of collaborative networks to accelerate learning about how to address high-leverage problems in the educational system (Bryk et al., 2015; Russell et al., 2017). This collaborative approach to educational problem-solving, initially developed by the Carnegie Foundation for the Advancement of Teaching, has seen promising results over the past several years.

In this paper, we explore the case of one particular NIC and the processes of adaptive integration that emerged from its members' efforts to improve early literacy teaching and learning using improvement science methods. In particular, we consider the implementation lessons that can be drawn from educators' uptake and adaptation of *both* innovations to improve literacy teaching in their districts *and* continuous improvement methods. The burgeoning literature on continuous improvement in education has noted the promise and challenges of using this approach in educational organizations, and Tichnor-Wagner and colleagues (2018) have argued specifically that "opportunities for adaptation alone will not necessarily overcome conditions in the local

environment that can thwart adaptation," (p. 43), recommending that changes be made to continuous improvement processes *themselves* to support perpetual adaptation in educational organizations. Accordingly, we present an in-depth comparative case study of four school districts' uptake, adaptation, and application of improvement science methods, and explore the ways in which teams engaged with continuous improvement principles and strategies in different ways.

We argue that these variable approaches to the improvement methodology itself are a rational extension of the concept of adaptive integration, and that the complexity and heterogeneity of educational organizations demands this flexible approach to continuous improvement. Furthermore, we assert that a flexible approach creates the potential for greater equity of access to continuous improvement in schools; becoming a continuous improvement organization requires a considerable amount of capacity and resources, which can make it more difficult for marginalized, under-resourced schools and communities to engage in continuous improvement research (O'Day and Smith, 2016). However, this study shows that using improvement science methods in flexible ways that are responsive to contextual conditions and the specific needs of school populations can create powerful organizational learning and improvement despite a departure from orthodox continuous improvement approaches, indicating that a broader theory of *how* to do continuous improvement in education will likely benefit the most vulnerable schools and districts.

4.2 Conceptual Framework

4.2.1 Continuous Improvement Research in Networked Improvement Communities

NICs facilitate systematic inquiry through the use of continuous improvement methods: specifically, NICs draw on improvement science, a specific approach to continuous improvement, to discipline their inquiry and problem-solving efforts. Drawn from the work of Walter Shewhart and his apprentice, W. Edwards Deming, improvement science methods gained popularity in the manufacturing sector in the early 20th century before migrating to healthcare, largely through the work of the Institute for Healthcare Improvement, in the 1980s (Bryk et al., 2015; Cohen-Vogel et al., 2015; 2018; Langley et al., 2009; Deming, 2000; Scoville & Little, 2014; Tichnor-Wagner et al., 2018). This approach led to significant gains in industry and healthcare, which facilitated improvement science methods' spread to other sectors. Continuous improvement approaches are currently enjoying a surge of interest and investment in the educational sector, fostered by the work of the Carnegie Foundation for the Advancement of Teaching as well as several major funding streams from philanthropic organizations and government agencies (Cohen-Vogel et al., 2016; Tichnor-Wagner et al., 2018). This enthusiasm is also contributing to the field's learning about how to do this work in educational contexts, which involves translating processes and practices that work well in manufacturing and healthcare organizations and fitting them to schools, districts, and state departments of education. Implementation researchers and institutional theorists alike have often observed the distinctive organizational and institutional structures that characterize schools and districts (e.g., Coburn, 2001; Powell & Dimaggio, 1991; Meyer & Rowan, 1976), so adapting continuous improvement approaches to fit the peculiarities of educational organizations will likely be a natural part of this translation process.

While continuous improvement has been (and continues to be) understood and deployed in a variety of ways, NICs use this specific improvement science methodology drawn from Shewhart and Deming. Improvement science methods facilitate learning about the ways in which systems produce outcomes through the use of rapid inquiry cycles, or small, structured experiments that allow participants to introduce changes into their practice and test their effects (Bryk, Gomez, & Grunow, 2011; Bryk et al., 2015; Deming, 2000; Langley et al., 2009; Lewis 2015). Rapid inquiry cycles, also called Plan-Do-Study-Act cycles, use the logic of the scientific method to generate context-specific knowledge about which innovations produce desired outcomes and improvements (Bryk et al., 2015; Tichnor-Wagner et al., 2018). While randomized control trials can generate knowledge about whether a particular intervention works under a specific set of circumstances, improvement science aims to build knowledge about implementing innovations effectively, reliably, and with the potential for scaling them across contexts (Lemahieu 2011). The disciplined inquiry enabled through improvement science methods focuses on amassing an evidence base of context-specific problem solving strategies, and understanding the barriers and facilitators to quality improvement in work processes (Berwick, 2008; Bryk, Gomez, & Grunow, 2010; Park et al., 2013).

Though disciplined inquiry cycles are a key part of this methodology, continuous improvement is not simply about individuals conducting inquiries to create contextualized knowledge about problems of practice; the term also implies organizational characteristics, particularly an ongoing frequency of quality improvement efforts, deep and extensive integration of the approach at different organizational levels, and the framing of problems of practice as system characteristics (Bryk, 2009; Park et al., 2013). These organizational elements of continuous improvement imply that disciplined inquiry work is deeply contextualized and sensitive to the

organizational context in which it is taking place. So, understanding individual learning about improvement also requires attending to the organizational conditions and structures that support and shape these individual level processes, and considering the interactions of the many processes, tool, and people that come together to form a system (Argyris & Schon, 1996; Bryk et al., 2015; Clark & Collins, 2007; Holland, 2014; Innes & Booher, 2010; Miller & Page, 2007; Opfer & Pedder, 2011; Seashore Louis & Leithwood, 1998).

NICs embed this disciplined inquiry into the social organization of a network that brings together many perspectives and stakeholders to support cooperative knowledge generation about complex problems. Through this networked structure, NICs connect practitioners, researchers, and designers to accelerate the testing and refinement of innovations that address the problem the network is trying to solve, and to build knowledge about how to contextualize and integrate new learning into the many types of contexts network participants bring to the table (Russell et al., under review*b*; Russell et al, 2017). This social structure is built to "catalyze the type of community that can solve complex problems" (Russell et al., under review*b*, p. 2) by providing diverse contexts in which to test promising interventions and bring together educators with diverse perspectives and expertise.

This social infrastructure can support the challenging shifts that educational professionals are often required to make in NICs; although schools and districts typically emphasize professional autonomy and the use of data for accountability, NICs value collaboration, transparency, evidence-based practices, and the use of data for both learning and improvement (Russell et al., 2017). Schools and districts are notoriously resistant to making these kinds of cultural shifts (e.g., Lortie, 1975; Coburn, 2001), so this culture-building function of a NIC should not be taken lightly. Taking up these cultural components requires a significant amount of trust and openness across NIC

members (Bryk & Schneider 2002; Russell et al., 2017). These cultural shifts are also important for moving towards continuous improvement proficiency; the continuous improvement literature notes that using data and evidence for learning purposes is not typical in educational organizations (e.g., Park et al., 2013), and so building this capacity comprises a key part of the uptake process in continuous improvement research. Engaging authentically in these data-driven learning processes requires a willingness to surface and confront problems head-on, which is also a normative shift for most educational organizations.

4.2.2 Integrity of Implementation and Adaptive Integration

In addition to the significant cultural shifts around data use and transparency, continuous improvement research also departs from the status quo in implementation research in its emphasis on adaptive integration and integrity of implementation. Adaptive integration stresses the idea of integrity of implementation, where implementers must remain true to the core principals of an innovation or intervention, while also systematically refining it to align with specific contextual conditions (Lemahieu 2011; Russell et al., under review*a*; under review*b*). NICs support adaptive integration by explicitly encouraging participants to learn about how to solve problems in their context, and to use that knowledge to integrate reforms into their system; this philosophy presumes that integrity, rather than fidelity, is the key to creating sustainable and scalable improvement (Lemahieu, 2011).

Continuous improvement researchers have typically talked about adaptive integration in terms of taking tools, practices, and processes developed in one context and systematically adapting them, using continuous improvement methods, to fit into others (Hannan et al., 2015; Russell et al., under review*a*; under review*;* Tichnor-Wagner et al., 2018). This is a particularly

powerful notion when applied to instructional interventions, which are often difficult to implement at scale due to the widely divergent structures and capacities in different educational organizations (Atteberry & Bryk, 2010; Bryk, Gomez & Grunow, 2011; Bryk et al., 2015; Russell et al., under review*a*; Penuel, Fishman, Cheng & Sabelli, 2011). This perennial challenge in education policy implementation is one of the reasons continuous improvement approaches have sparked such excitement in the field, since they provide an intentional, disciplined approach to accommodating the inherent variation in educational organizations. Past continuous improvement research efforts in education have shown promising results in this regard; for instance, researchers have shown that adaptive integration can support the use of novice teacher feedback processes, coaching models, and social and emotional learning programs in K-12 spaces (Hannan et al., 2017; Russell et al., 2017; Russell et al., under review*b*; Tichnor-Wagner et al., 2018).

However, variation in organizational capacity to take up continuous improvement methods is also part and parcel of this work; given the complex and counter-cultural nature of *using continuous improvement methods themselves,* it follows that using this methodology will also require intentional, adaptive approaches across variable local contexts. To be sure, this adaptive process also requires attention to integrity of implementation; as Lemahieu (2011) and others (Hannan et al., 2015; Tichnor-Wagner et al 2018) have explained with regards to adaptively integrating innovations, engaging in this work systematically requires understanding the indispensable technical core of the thing to be adapted, and then making intentional changes around that core to tailor it to local conditions. Using improvement methods with fidelity has, perhaps paradoxically, been presumed to be the foundation of this process, since continuous improvement tools help users engage in systematic learning about adaptation. However, Tichnor-Wagner and colleagues (2018) observed that, while continuous improvement approaches can

anticipate and support recursive, iterative implementation processes, the straightforward uptake of continuous improvement methods may not be enough to overcome the many conditions that emerge to pose barriers to implementation with integrity. Instead, they argue, continuous improvement methods themselves ought to be adapted to support "ongoing adaptation in the face of [implementation] hurdles" (p. 43). Beyond this discussion, this dimension of continuous improvement research in education has been largely neglected in the literature so far.

To address this gap, our study centers this question of how educators adapt continuous improvement processes to most effectively address pressing problems and respond to variable conditions within their local context. As researchers have explored the potential of continuous improvement research to support educational policy and innovation implementation, we have often discussed continuous improvement work in a way that highlights the load bearing conditions (Iriti, Bickel, Schunn, and Stein, 2015) that educators, schools, and districts should possess to engage in the innovations associates with the methodology; essentially, in establishing what this work looks like, how it plays out, and its potential for scalable improvement, researchers have focused on delineating and describing the conditions that support, or hinder, the uptake of continuous improvement methods (e.g., Hannan et al., 2015).

Here, we expand this work, drawing on Iriti and colleagues' (2015) conceptualization of load bearing conditions to move beyond the initial conditions that must exist in a context to enable the traditional use of improvement science methods and consider the essential elements of the methods themselves and how they might be adapted for use in a wider variety of contexts. Given the migration of improvement science methods from industry and healthcare into school districts, it seems logical to assume that some methodological adaptation will need to occur. In particular, a flexible approach to improvement science uptake may support a more robust and inclusive way of

doing continuous improvement in educational organizations, particularly in districts and communities that have been exploited by historic, systematic disinvestment, since adopting continuous improvement methods in schools is fairly resource intensive. As such, we argue that exploring ways to take up continuous improvement methods with *integrity* in schools and districts, and attending to how educational organizations might adapt them based on their local contextual conditions, is an important part of building a theory of how to approach continuous improvement work in education. Furthermore, exploring this notion will contribute to the field's knowledge of how to foster greater access to the promise of continuous improvement across various contexts with different capacities and resources.

4.3 Methods

4.3.1 Study Context and Participants

For two years preceding this study, seven rural districts in a Southern state collaborated with a state department of education to form a NIC focused on learning how to improve early literacy teaching outcomes in contexts across the state. With technical support from the Carnegie Foundation for the Advancement of Teaching, the state Department of Education started this ambitious NIC to address low statewide literacy achievement levels and understand changes and innovations that could improve early literacy teaching and learning at scale.

In the spring of 2016, a group of seven districts began attending network convenings to learn continuous improvement methods and collaborate around solving problems of practice with the network Hub (composed of the leading state DOE and with support from staff from the

Carnegie Foundation) and support professionals working in the state DOE's regional offices. The authors were developmental evaluators for the NIC, and so worked closely with the Hub to understand their goals and aspirations for the network, and also engaged regularly with network members to keep a pulse on their experiences and provide summative feedback to the Hub throughout the project. Over the next two and a half years, participants continued to attend network meetings three times per calendar year. In between these face-to-face convenings, network members took part in coaching webinars led by the Hub that provided opportunities for tool and routine development related to literacy teaching and learning, continuous improvement coaching, and cross-district collaboration.

During these convenings and webinars, network members received training and coaching in using improvement science methods, particularly ways to deepen their understanding of a problem of practice with root cause analysis and the use of PDSA cycles to systematically test ways to address the problem. The Hub also encouraged network members to test specific kinds of tools and innovations during year 1, in particular, a literacy behavior protocol observation tool, which aimed to help practitioners build a common understanding of key student literacy behaviors, and a Google Doc meeting template to refine communication and coordination about teaching and learning between teachers, instructional assistants, and literacy specialists. District leads and school teams were permitted to pursue individual change ideas as well, though they were encouraged to engage with these shared tools to accelerate learning about these specific interventions across the NIC. This first phase of the NIC concluded in the summer of 2018, when the state DOE added additional districts to the network to explore scaling across the state. Although scaling is a critical component of networked improvement communities, here we focus on the first phase of the network's development, specifically the work of four Cohort 1 districts, to explore

the process of taking up continuous improvement approaches to support improvements in teaching and learning.

4.3.2 Case Study Design

We used a comparative case study approach (Yin, 2003; Miles, Huberman, & Saldana, 2014) to explore the following research question: *How did four case study districts in the network use and adapt continuous improvement methods to navigate complexity in their organization?* The goal of our qualitative study was twofold: we aimed to *describe* participants' use of continuous improvement methods to add to the still developing literature on using this approach in schools, districts, and other educational organizations, and to generate theory by exploring themes and patterns in district teams' uptake, application, and adaptation of the methods to attenuate a pressing problem of practice (improving early literacy teaching and learning).

Using a comparative case study approach was particularly well suited to answering this research question because multiple case study designs emphasize understanding the *process* of implementation and unpacking possible explanations for how and why that process unfolded in the way it did (Bush-Mecenas & Marsh, 2018). In this project, network participants engaged in several layers of implementation work—they took up new instructional practices and organizational routines to support literacy while simultaneously learning and enacting continuous improvement methods—so we aimed to capture those interconnected processes in this case study. Collaborative problem solving research is inherently a complex process that involves many perspectives and interlocking organizational mechanisms, so exploring those dynamics and the ways in which they affect implementation is an important analytical goal. Furthermore, the notion of *context* is critical to networked improvement communities and continuous improvement

research (Bryk et al., 2015; Langley 2009; Tichnor-Wagner 2018), so we designed our case study to attend to these elements of the implementation process. One limitation of our approach is that we could not simultaneously explore the instructional impact of the literacy improvement strategies along with the implementation of improvement science methods; a lack of literacy outcomes data and scoping concerns necessitated that we focus primarily on the uptake and implementation of improvement science strategies.

4.3.3 Data Collection and Sampling

The data collected for this study were part of a developmental evaluation of the NIC that aimed to comprehensively document the initiation, development, and growth of the network over time, and to provide tailored formative and summative feedback to network leadership. For this study, we focused on the first cohort of the NIC, in which seven districts across two rural regions worked with the state DOE to improve early literacy teaching and learning. We selected two case study districts from each of the two participating regions of the state. To select our four case study districts, we used purposive sampling to identify the organizations that would offer the richest insights into our research question (Maxwell, 2004; Strauss & Corbin, 1994). The case study districts showed more robust and creative engagement with continuous improvement work than their other three Cohort 1 district counterparts, and so they provided the most useful insights into how schools engage in this kind of process. Furthermore, our four case study districts represent varied and challenging contexts in which to do continuous improvement work; three of the four districts served high populations of students growing up in poverty, and one district served an unusually high (>50%) proportion of Latinx students for schools in this region, most of whom were also English Language Learners.

We collected a wide range of data to build an understanding of the implementation experiences and perspectives of the school improvement teams in the four case study districts (Sycamore County, Elm County, Cherry County, and Cedar County). We worked with the improvement leads in each district to gain access to the following data sources:

- Two semi-structured interviews with each pair of district leads in each district, conducted at two separate timepoints in year 2. Our protocols asked them to reflect on their districts' use of continuous improvement methods, their engagement with the NIC, and their perception of these processes' effect on their early literacy teaching and learning systems. It is important to note that we could not get formal interviews with the Sycamore County district leads, but we did have relevant informal conversations with them at network convenings.
- Artifacts of each team's continuous improvement work, which varied slightly based on the documentation routines of each team, but always included some form of PDSA documentation as well as presentations that each district prepared in order to share what they were learning with educators in other districts during network convenings;
- Virtual observations of school improvement teams' meeting routines, which included:
    - 3 Google Hangout observations in Cedar County; 1 Google Hangout observation in Sycamore County
    - 1 videotaped improvement team meeting in both Cherry County and Elm County,
- Semi-structured interviews with school team members that asked about their experiences in the network and the supports and barriers that affected their participation, and
- Field notes capturing team participation in network-wide convenings.

One limitation of our data set was a relative lack of information about specific, measurable literacy outcomes in these schools. This was a network-level design issue, because neither the Hub nor the districts were tracking changes in leading indicators or imposing measures of literacy practices. Furthermore, the network's approach to school improvement focused on routines that support literacy teachers' work in years 1 and 2; subsequent work in the network gradually moved more toward focusing on literacy content. We think that linking the improvement of tools and processes with student outcomes is an important direction for future research, and that the work that we've done here illuminates some of the foundational elements of continuous improvement work in schools that can support more specified instructional aims.

4.3.4 Data Analysis

To explore the dynamics of continuous improvement uptake and adaptation across our case study sites, we proceeded with our analysis in two main phases; first, we explored within-case uptake of continuous improvement approaches, followed by a comparative cross-case analysis of district teams' adaptation of improvement science methods to fit their contexts. To develop a deep and clear picture of the improvement work happening in each case study district, we focused on building richly descriptive case analyses in the first phase. Then, we looked critically across the cases to compare teams' different approaches and understand how continuous improvement methods interacted with the many social, technical, and organizational conditions supporting and constraining this work. This layered analysis allowed us to understand the character of the improvement work happening in each school and notice the adaptive difference in districts' use of continuous improvement methods.

Within the first analytical pass, we used a two-pronged strategy: we compiled case summaries structured by simple codes often used in studies of continuous improvement uptake in education (successes in, and barriers to, continuous improvement work), and we synthesized teams' PDSA data into a single document that allowed us to observe teams' use of PDSA cycles over time. For the case summaries, we used our two broad codes to reduce and organize data from transcribed interviews with school teams and district leads, improvement meeting observations, and field notes from network-wide convenings. Once we organized these data, we identified themes that emerged within each broad category, developing a set of inductive codes specific to each district. Concurrently, we processed and synthesized district and school teams' PDSA data in a matrix that summarized each case study district's streams of PDSA testing over time. Columns were constructed using the change ideas (innovations that were being tested in the PDSAs) and rows were dates, and we filled cells with synthesized explanations of teams' PDSA forms. This matrix allowed us to observe how teams tested and refined change ideas over time and what they learned from these processes.

Following the within-case analysis, we moved to looking across cases to compare how each district interpreted and adapted continuous improvement methods for their local context. To do this, we first developed a cross-case matrix, where our four case study districts formed the rows and a set of four broad themes (continuous improvement outcomes, school team structure and distribution of expertise, district leads as brokers and sensemakers of continuous improvement, and adaptive approaches to using continuous improvement methods) from the first coding pass were the columns. We then summarized data from all of our qualitative sources (described above) and placed them in the appropriate cells to look across each districts' work around common implementation themes and interpret how each team adaptively integrated continuous

improvement into their district contexts. This helped us to understand the patterns across cases, in particular the relationship between district system structures, team leads' assumptions about school teams' capacity and resource availability, and the manner in which each district employed improvement science methods.

4.4 Findings: Local Approaches to Continuous Improvement Efforts

In this section, we describe the continuous improvement work that happened in each of our four case study schools, paying particular attention to: the structure and distribution of district- and school-based improvement teams; key characteristics of each district organizational system; foci of continuous improvement projects broadly, and PDSA work specifically; the continuous improvement outcomes each team reported; and the challenges they faced. These descriptive representations are important because they offer examples of various approaches to continuous improvement implementation processes in schools and districts, and they lay the foundation to understand how teams adaptively integrated continuous improvement methods to suit their purposes, a topic we turn to in the second half of our discussion of the findings.

4.4.1 Sycamore County

In Sycamore County, continuous improvement work took place across all four of the district's elementary schools. The district leads were central office staff tasked with federal programs, literacy instruction, and RTI coordination. They organized four small school-based teams, each of which was composed of 3-5 teachers and interventionists working with K-3

students. To coordinate at the district level, the four school teams, along with district leads (and sometimes regional office staff), participated in a Google Hangout about once a month, allowing school teams to communicate with and learn from one another. In an observed Hangout, the district leads led the conversation, giving each school team 5 minutes to report on their current work and the artifacts they were preparing. Leads offered positive comments, but did little explicit coaching about continuous improvement methods.

In year 1 of the project, Sycamore focused on refining huddle routines to facilitate communication and coordination between teachers, interventionists, and other professionals involved in literacy instruction, and then teams moved to testing more instructionally-focused change ideas, such as audits of students' time spent in text and questioning protocols, in year 2. In particular, their continuous improvement efforts around huddles at the school level highlighted problems with a key literacy teaching routine used district wide: small group, differentiated reading workshops. More generally, the PDSA cycles helped school teams make a case to district leaders that they needed more flexibility in the enactment of curriculum and instructional systems in order to experiment with systemic improvements.

School team members' talk about their PDSA cycles showed that their continuous improvement efforts ameliorated a key instructional problem in their district: a lack of communication and coordination between instructional staff teaching reading. At Sycamore, a district-wide reading initiative structured the school day, which included a "workshop" period. These workshops consisted of small ability groups taught by paraprofessionals and teacher assistants, with instructional activities pre-planned by the classroom teachers, however there were limited opportunities for communication among teachers and support staff about students' literacy development needs. The Sycamore teams tested the introduction of short "huddles" (10-15

minutes, depending on the school) during this workshop period to allow classroom teachers, paraprofessionals, and interventionists to share information about the students with whom they were working. In interviews, school team members talked about how the huddle routine helped illuminate and address problems with the workshop structures, particularly by fostering communication between classroom teachers and instructional support staff. For instance, when asked about evidence that huddles led to improvement in their school, one school team member explained, "we've noticed that, in our workshops, there's a lot of kids that needed to be moved [to another level] and we may not have even known that… because [before initiating the huddle routine] we didn't get to talk to their workshop leads." Several other teachers also noted that the huddles helped to address this problem by creating a reliable opportunity for all the adults responsible for literacy teaching activities to talk to one another and share data on their students.

Another important outcome of Sycamore's improvement work was the identification of the need for increased flexibility within a fairly rigid instructional system. Many of the teachers we interviewed talked about the tightly controlled environment in this district, and noted that they had to seek permission from the central office to do the kind of exploration involved in testing change ideas. A few of these school team members thought that improvement work was opening opportunities for innovation in their district. For instance, one school team member noted that, "I think by doing the huddles it made the central office more aware of some of the places they're going to have to be more flexible or reconsider what we're doing," and another team member said that, despite the fact that teachers generally had little control over things in this district, "our district has let us actually do something a little different, not a lot of change, but some of the things like huddles and different things that we've tried." At least in these schools, testing around huddles

planted a seed about the need for additional flexibility to explore potential improvements in the district's instructional system.

Although the Sycamore team arrived at important new knowledge about their instructional system, they also faced many challenges as they engaged in continuous improvement work. Most school team members talked about the challenge of finding the time to engage in continuous improvement work, and the difficulty of doing experimental inquiry cycles in a centrally controlled environment. Specifically, teachers mentioned the challenge of documenting PDSAs, the district's strict requirements about curriculum and pacing, and the many other concurrent initiatives as barriers to their engagement in continuous improvement. Most Sycamore teachers also noted their limited autonomy in the classroom, and talked about how the district's tightly controlled instructional environment constrained their ability to deviate from district mandated instructional routines and practices. However, the learning and change that emerged as a result of the continuous improvement efforts in this district suggests that using improvement science in difficult organizational contexts has the potential to lead to meaningful gains, even when continuous improvement strategies are not used with high degrees of fidelity.

4.4.2 Cherry County

In Cherry County, continuous improvement work was based in the small district's only elementary school, Cherry County Elementary. The district leads were a school administrator and a central office staff member who specialized in literacy. The school team composition changed from year 1 to year 2 due to teacher turnover and shifting improvement foci; in year 1, the team was composed of interventionists and classroom teachers, whereas in year 2, the team was composed of an interventionist, ELL (English Language Learner) specialists, and classroom

teachers. In year 1, much of Cherry County's improvement work revolved around using Google Docs for RTI data meetings and testing the literacy behavior protocols in their school. This meant that some of their improvement work was built into pre-existing data meetings, and some meetings happened independently of existing meeting routines. To help the school team members engage in these additional meetings, the district leads provided substitute teachers to support teachers' work beyond their regular classroom assignments. Meeting routines in year 2 were variable, due to the inquiry-driven nature of their improvement focus and the gradual rebuilding of the team. These meetings provided an opportunity for the team members to receive improvement coaching around their change ideas and inquiry work from the district leads.

In terms of PDSA testing, Cherry County's efforts to test and refine the literacy behavior protocol tool were the most extensive, and the testing process enhanced instructional alignment by facilitating communication across teachers and other instructional professionals. The district leads systematically tested several versions of the literacy behavior protocol with teachers and interventionists on their improvement team, making changes to the tool based on the feedback they received from team members as they used it in their practice. While the district leads were primarily responsible for documenting and guiding this process, the teachers and interventionists also contributed critical information. When school team members reflected on testing and refining the literacy behavior protocol in interviews, they noted that the conversations the protocol generated helped them to align their instruction, which in turn helped the students to make stronger connections in their learning. School team members and district leads reported that this increased instructional alignment supported a common vision of literacy instruction across Cherry County Elementary's complex literacy environment, reflecting the district leads' goal of integrating and crystallizing messaging about literacy improvement across this school.

As the Cherry County district leads pivoted to a narrower focus on English Language Learners, they directed their efforts to understanding the state of their current ELL instruction and support system and exploring possibilities for designing improvements. This led to crucial insights about where their current system was failing as well as strategies for reinforcing and integrating ELL supports. This work was specifically tailored to Cherry County's context, which was unique in the network and in their region. Cherry County's elementary school is over 50% Latinx students, many of whom were ELLs, and both district leads consistently noted that supporting these students was central to any literacy improvement effort in their district. To learn how to focus their ELL continuous improvement efforts, the Cherry County team used root cause analysis to unpack their ELL system and identify key areas of concern and weakness. We observed teams brainstorming the root causes of their ELL literacy achievement problem in a recorded improvement team meeting, during which the district leads led a conversation between classroom teachers, ELL specialists, and interventionists, who shared their ideas and insights about ELL students, how the school was supporting them, and what could be improved. Furthermore, we heard from both district leads and school team members that collaboration had been "an area of trouble" in the school in the past, but that these cooperative continuous improvement activities provided critical opportunities for authentic collaboration around a pressing problem of practice.

Cherry County was a very small yet fairly complex organizational context; many concurrent district-level policy changes were taking shape as they took part in continuous improvement. For example, their district was part of two separate collaborative reform projects in addition to the literacy NIC, and they were in the process of implementing a new portfolio system, assessment system, *and* a new ELL curriculum. While these changes in the district posed implementation and integration challenges, the district leads talked about intentionally using

improvement science methods to help them be systematic in the way that they enacted multiple reforms. For instance, in an interview one of the district leads explained how improvement science provided a helpful lens for strategically leading multiple, concurrent initiatives:

> We realized this improvement science has become a really helpful tool for us in that administrative role to really look deeply at what we're asking teachers to do, to really think carefully… it changed our perception as we began to think about what other changes we wanted to introduce. It gave us a tool to use to really help them to make the changes that we wanted to make, so that we knew we were doing it strategically, we knew we were measuring along the way if there was a problem.

Furthermore, while the small size of their school system certainly posed challenges with regards to time and resources, it may have also made the district leads' efforts to coordinate continuous improvement work more manageable. They talked about using continuous improvement methods to document and evaluate what they were doing across many reform efforts and anchor the changes they made in data and evidence. Their school team members talked about how "everything kind of aligns," so it seemed that efforts to make these efforts cohere and make sense of many initiatives teachers were effective. This suggests that the improvement work at Cherry County was supporting movement toward the district leads' dual target outcomes of building a common vision of literacy instruction across their literacy system as well as enhancing supports for ELLs.

4.4.3 Elm County

In Elm County, continuous improvement efforts were based in one small elementary school, Green Elementary. The district leads were the central office staff who specialized in literacy and teaching and learning systems. The school improvement team at Green Elementary

was composed of teachers, the school principal, and a reading coach. The team met twice a month to discuss their improvement work and current change ideas. In an observed school team meeting, we saw the district leads asking the school team members targeted questions that fostered continuous improvement philosophies and strategies. For example, a district lead asked a school team member "How will you know if this change has made an impact?" to guide that team member toward evidence-based evaluation of a change they made in their practice. Elm County added additional team members to their improvement team between year 1 and year 2, which they noted was difficult, since they had to do a lot of work to help those new team members to catch up on improvement science knowledge. However, both the district leads and the school team members mentioned that the opportunity for cross-role collaboration that these continuous improvement meetings created was valuable and meaningful. Comments from both district leads and team members in this district indicated that their goal for their continuous improvement work was to integrate many different literacy improvement efforts across their system.

Initially, the Elm County team focused on improving communication about reading instruction across a variety of people in both their school and district offices, which is important since many different educators tend to support struggling students as they learn how to read. One of the change ideas they tested around this goal was the use of a huddle routine to give teachers and interventionists a chance to communicate about specific students. The team tried these huddles in a variety of configurations—for example, focusing them on student work, exchanging post-it notes about students' reading behaviors, and doing an email huddle—and in each test, improvement in overall communication was noted. One school team member said that "PDSAs have given us a common language and common thought process… of how to think about solving these problems." This common language may have helped to facilitate the cross-role

communication that both district leads and school team members note is another critical support for literacy teaching and learning.

Elm County subsequently built on this work to foster a common, coherent language for literacy improvement across the school. An important component of this was their systematic analysis of running records located in student reading folders, which the district leads reviewed and tracked with a continuous improvement lens. The district leads zeroed in on these reading folders as a critical point for communication about reading through a root cause analysis activity. These folders, which follow students across grade levels, contain artifacts documenting students' literacy learning progress and benchmark assessments over time. The district leads realized that these were critical tools because they shape how teachers approach reading instruction for each student, since they illustrate their achievements as well as their areas for growth. After reviewing a sample of teachers' reading folders, the district leads determined that teachers' comments needed to focus more clearly on evidence-based literacy behaviors in order to communicate student progress and learning needs. Subsequently, they introduced a checklist of reading behaviors and aligned professional development activities related to those reading behaviors. In order to assess whether this change led to improvement, the district leads created a rubric of behaviors to evaluate the quality of teachers' comments on students' literacy behaviors. They found that, over the course of the year, teachers' comments had become more focused on critical literacy behaviors. One district lead noted, "the last time we checked folders, there was only one teacher in the entire school who did not have specific literacy behaviors. So, that was an amazing improvement."

While Elm County engaged in extensive work around improving literacy, systematically testing the small changes they made and documenting that process was a challenge for them. A school team member noted that small changes can feel counter-intuitive, since "as educators, [we

are] looking at the big picture and looking at the caveats… instead of just trusting the process." Another school team member mentioned that they felt they weren't being systematic about the changes they were making at their school and that they were using informal data for PDSAs. Other comments in interviews suggested that there was some tension in Elm County around improvement science methods more generally. One district lead noted that "our students are not crash test dummies" when talking about testing new ideas around literacy instruction. On the other hand, she also talked about how teachers were trying new things in their classrooms and sharing the results of those changes with the rest of the continuous improvement team. While Elm County struggled with systematicity and documentation of PDSAs, firsthand accounts of their literacy work over the past two years suggests that their continuous improvement work had a positive impact by fostering a more coherent instructional focus around literacy improvement.

4.4.4 Cedar County

In Cedar County, continuous improvement work was based in three of the districts' seven elementary schools, and led by district staff from the curriculum and instruction office. They facilitated and supported the three school-based teams, which included classroom teachers, interventionists, special education teachers, and school psychologists in year 1, and in year 2 were reconstituted with interventionists and classroom teachers. Each school-based team was made up of about 3-4 individuals. It was sometimes difficult for the school teams to find a reliable time where they could meet in person in addition to previously scheduled meetings, so they primarily kept in touch through Google Hangouts, emails, and occasional huddles. Regional office staff and district leads visited each school team about 2 times per semester, which provided the school teams with an opportunity to receive feedback and coaching around their continuous improvement work.

During these visits, district leads and regional officials helped school team members understand how to run PDSAs and how to integrate them into their practice in an authentic way; for instance, in a meeting observation, we noticed district leads and regional office staff asking school team members about the problems of practice they were most interested in learning about, and then helping them to plan for small tests of change based on the problems they explained.

To begin their continuous improvement work, teams in three schools tested changes to the Google Doc-based system they had been using to coordinate collaborative conversations among teachers, interventionists, and other instructional support staff about student progress in literacy. Cedar County had already established a Google Doc tool in the fall prior to joining the network, so, when they started their improvement work around this tool, school team members were already in the process of taking up and tweaking this routine to fit their own needs. Continuous improvement methods provided an evidence-based strategy for this work. When discussing this process in interviews, the Cedar County district leads hypothesized that the staff's prior familiarity with the Google Docs supported continuous improvement uptake, since the school team members had already had an opportunity to use the tool and consider how they might tailor it to their needs. They also noted that the team's familiarity with the tool was a helpful foundation as the school teams learned new improvement science strategies. Each school team ran a series of related PDSAs around their changes to the Google Doc tool and routines for use, such as adding new components to the Google Doc, sharing samples of meaningful comments about student literacy learning with teachers using the tool, and asking building administrators to endorse the Google Doc as an important tool for teaching practice. Over time, the school teams noticed that the quality and specificity of comments in the Google Docs increased overall. School team members also noted that, while the quality of comments increased, so did the data meeting efficiency and the quality

of collaboration between teachers; for instance, one interventionist noted that the improved commenting routines "streamlined" the process of communicating about RTI students and made the data team meetings more "interactive."

Concurrently, district leads coordinated literacy learning walks for all seven elementary administrators in the district, who visited each other's buildings to learn about literacy instruction. This routine served as an as an inquiry tool to learn about how to facilitate communication and enhance literacy expertise across different role groups in their districts. In the Literacy Learning Walk routine, teams of school administrators visited classrooms during literacy lessons. They focused on observing key instructional behaviors and choices, like time spent reading texts, text difficulty levels, and teachers' questioning sequences. These walks were followed by a group debrief where all participants reflected on what they learned about literacy instruction from their observation.

These literacy learning walks laid an important foundation for instructional enhancement across the Cedar County school system. The Cedar County team used these learning walks as an inquiry tool to learn about how to facilitate communication and enhance literacy expertise across different role groups in their districts (e.g., among administrators). The district leads explained how this was a critical step for continuous improvement in Cedar County, since several leaders in network schools were new to instructional leadership roles. The district leads mentioned the value of the substantive learning this routine generated for the school leaders; it gave them a framework to learn about critical instructional elements of literacy, like text complexity and questioning, and also provided them with material for engaging with the teachers they observed, which are all central elements of supporting instructional leadership. Building leaders planned to continue to

use the literacy learning walk routine and to find additional ways to bring teachers into the process and get them involved with this inquiry work.

In Year 2, the school teams' PDSA work tended to be more individualized and rooted in specific problems of practice, but data gathered in school team interviews indicate that many teachers focused on inquiries related to literacy teaching strategies, such as time spent in text and text-based questioning. For instance, several teachers on Cedar County's school team noted that they had "eye opening" experiences about the time students were spending reading texts as a result of their small tests of change. One teacher elaborated on how PDSAs affected their practice in this way:

> I'm doing something that's actually helping my teaching 'cause not only am I analyzing my instruction of how I'm asking the questions, I'm also helping those students to get better at answering a question because I'm taking what a change idea and I'm taking it in small steps so that in the end I can be more successful.

This school team member's perspective suggests that using the PDSA tool helped them to enhance their questioning skills by learning about how to analyze their own instructional practice. Both district leads and school team members talked about how their continuous improvement work created meaningful change in their professional practice, improved the quality of communication between teachers, interventionists, and other educators involved in literacy teaching, and fostered important collaboration across different roles in their district.

As with the other three case study schools, Cedar County enacted their continuous improvement efforts in a complex and shifting organizational environment as they adopted new assessment systems and became part of a new statewide literacy initiative while also taking part in this continuous improvement network. District leads noted that the challenges that come along

with these changes prompted them to think about how to use continuous improvement methods as a strategy to integrate various literacy improvement efforts in their district going forward. Another challenge that both school team members and district leads mentioned often was the time and effort involved in documenting PDSAs; they often talked about how difficult it was to balance documenting PDSAs with their other responsibilities. Nevertheless, the considerable improvement capacity that the district leads fostered across multiple school sites suggests that these tools and methods have the potential to continue to support integrative, systematic literacy improvement efforts in Cedar County going forward.

4.4.5 Brokering Continuous Improvement Uptake: District Leads' Adaptive Integration of Continuous Improvement Methodology

In each of our four case study sites, district leads enacted continuous improvement methods in strategic and highly contextualized ways. They occupied a critical role in the NIC, existing in the intermediary level of the network, between network leaders, the source of directives and training, and school-level improvement teams, who carried out the continuous improvement work on the ground to realize that vision. These meso-level actors are key mechanisms for negotiating and shepherding change in a system (Fligstein & McAdam, 2012; Holland, 2014), and district leads took up this role in diverse ways that reflected their distinct system structures, capacities, leadership styles, and aims for continuous improvement work. Their strategic brokering efforts offer important insights into the ways in which educators can take up and enact continuous improvement work specific to their context and goals in educational organizations.

Significantly, each district team retained some core elements of continuous improvement methods and philosophies, but made changes in the specific ways they enacted them to fit their

perceived organizational capacities and the problems they were trying to address. While the network trained district leads in improvement methods and directed their focus to early literacy improvement, district teams had considerable flexibility to design their own continuous improvement work routines and processes. Districts also brought their own specific concerns and goals to bear on this work, and, coupled with organizational structures and capacities, implemented continuous improvement practices adaptively in their home organizations. Comparing the four case study schools offered perspective on how continuous improvement approaches can take hold in schools and the kinds of strategies and adaptations that improvement leaders might use to fit continuous improvement methods into their system. Accordingly, in this section we discuss how district leads in each case engaged in these different strategies to lead continuous improvement work in their districts and the effect those approaches had on implementation.

4.4.6 Cherry County – Centrally Managed Continuous Improvement to Manage Competing Demands

In Cherry County, the district leads organized their continuous improvement work to control messaging and buffer their school team members from additional demands on their time in an already complex reform environment. They described their leadership style as "sneaky" in that they brought their school team members into the continuous improvement work in targeted but limited ways; these two high capacity district leads ran most of the improvement science work themselves, taking responsibility for most of their team's PDSA documentation. They reasoned that, given the many different initiatives happening simultaneously in their small district, their teachers had too much on their plates already, so they were reluctant to add another task to their workloads. In talking to their school communities, the district leads presented continuous

improvement methods, particularly the PDSA cycle and root cause analysis, as tools for improving practice, which generated enthusiasm and interest among the teachers who chose to take part in continuous improvement efforts. The district leads describe working with school team members to identify areas for inquiry that they already had in their practice, and then working together with them to shape PDSA cycles to fit into that work they were already interested in. In this sense, the Cherry County leads persuaded their school team members to engage with continuous improvement efforts by showing them real value of using these methods, but also buffered them from the more time consuming parts of this methodology. However, by heavily scaffolding the continuous improvement process, district leads limited their investment in building team capacity to engage in improvement work without direct support from district leads.

Though district leads at Cherry County did not engage in extensive capacity building around improvement science methods with their school teams, they did include them in a strategic way that facilitated authentic collaboration across roles around common problems; first, through their work with the literacy behavior protocol and meeting routines in year 1, and then through their exploratory inquiry efforts about ELL support systems in year 2. While school team members showed that they had limited knowledge of the specifics of continuous improvement methods in their interviews, they expressed gratitude and excitement about the opportunities for cooperative work that continuous improvement provided. For instance, one ELL teacher made these comments about why it was important for their team to come together around challenging issues in their school:

> we're looking forward to seeing how what we've come up with is getting to be implemented and I think that it's easy to forget if you are [an ELL] teacher, it's easy to forget some of the challenges that the regular classroom teachers have, and if you're a regular classroom

> teacher or content teacher, it's easy to forget that some of the challenges that we face in working together with classroom teachers. So I feel like this has definitely helped with cohesiveness and I do feel like that's an area of trouble and at our school – it's collaboration and being transparent and vulnerable about that.

Another teacher noted that continuous improvement meetings and activities were the only opportunity she had to meet with peers in her building. The district leads said that they were hoping to grow a more collaborative culture in their school, since past administrations had not fostered this kind of environment. School team members' interview comments suggest that this support of collaboration through continuous improvement was effective, and that they appreciated the opportunities to learn about practice improvement with other professionals in the building.

Cherry County's district leads conceptualized their improvement work around the context-specific, overarching goal of transforming their ELL support system and curriculum, and used exploratory inquiry methods in year 2 to unpack and understand their current ELL system. They used a root cause analysis diagramming activity to bring teachers, interventionists, and ELL specialists into this process collaboratively. We observed this process through a video recording and noted that the district leads worked cooperatively with their school team counterparts to deeply understand the instructional staff's experiences with the current ELL system, and to organize and record their concerns and unmet needs. Through this activity, the mixed-role group identified issues such as "trust and transparency," "planning and collaboration," and "teacher knowledge," which helped them think about how to focus their improvement work in the 2018-2019 school year, and to realize that teachers across the school would need additional ELL training. These insights laid key parts of the foundation for the district's ongoing work around ELL instruction and learning systems.

4.4.7 Elm County: Centrally Managed Continuous Improvement to Foster Instructional Coherence

The Elm County district leads were also concerned with the many initiatives happening in their school, and they too centrally managed continuous improvement efforts to support instructional coherence by integrating the many schoolwide literacy reforms happening in Green Elementary. This in turn supported their effort to create coherent messaging around literacy improvement in this school and, more broadly, across their district. The district leads talked at length about how their role was to help the instructional staff at Green Elementary understand how the many different strands of literacy work in their district fit together around a common improvement objective. "We do think integratively… we feel like that is our job, essentially to integrate all of the overwhelming amount of information we get," one district lead said in an interview. They went on to explain that they were working on "aligning all the people in the building and all of the initiatives in that building around our primary reading goal… we're working to frame all of the action steps of both people and initiative around that goal so that everyone is thinking in the same direction."

In a pattern that is perhaps related to this overarching integration effort, the Elm County team took up continuous improvement methods more broadly and conceptually than their counterparts in Cherry County and Cedar County; while they ran and documented some PDSAs around huddle routines, the bulk of their work with continuous improvement focused on using improvement science language and concepts to anchor their efforts to foster quality communication and evidence-based decision making across their school team members. For instance, we observed district leads continually using the language of "testing" in their meeting conversations, and often asking teachers, "how will you know if the change that you made let to

an improvement?" This broad focus on improvement concepts did not translate into capacity building for following the specific inquiry steps of the PDSA routine—in fact, many school team members talked about their struggle to collect data and run PDSAs in a systematic fashion. In this approach, the data collection and study elements of PDSAs were largely absent, which may have limited opportunities to systematically learn about specific change ideas in this district.

However, the Elm County leads were most concerned with building level coherence around literacy improvement, and their use of improvement methods to foster a common language, coupled with their data-driven efforts to refine key literacy teaching routines, did appear to make meaningful progress toward these goals. The district leads' work with the reading folders was the clearest example of this; after reviewing and analyzing reading folder data from a small sample, the district leads learned that they needed to focus their efforts on helping teachers learn the language of substantive reading behaviors in order to make these data meaningful and valuable. Their measurable results—the observation that, after introducing a literacy behaviors checklist tool and focusing literacy professional development on these behaviors, only one teacher in the school failed to use the specific reading behavior language—suggests that their use of continuous improvement concepts to build instructional system coherence was, in fact, supporting that goal. This outcome is particularly interesting since improvement science methods are typically focused on using fast, systematic changes to learn from small experiments, but here the district leads used the spirit of a PDSA—collecting data, making a decision about how to improve based on that data, and measuring the results of that decision—to work on a school level problem, subverting some of the axioms of PDSA testing but still achieving meaningful, measurable improvement outcomes.

4.4.8 Cedar County: Distributed Continuous Improvement for Individualized Instructional Support

In Cedar County, district leads focused on building capacity across multiple school teams to enable them to run concurrent small tests of change on problems of practice, which fostered parallel learning about instructional alignment across several different school contexts. The district leads persuaded their school team members to see the utility of continuous improvement methods by showing them how to fit PDSAs into questions they already had about their practice; in so doing, they demonstrated the authentic usefulness of this method and respected the teachers' needs and personal practice. The district leads talked about approaching continuous improvement leadership as "starters" and "support" for their school teams, offering coaching and sensemaking to help them construct meaningful PDSAs during improvement team coaching sessions, and also encouraging their school team members to take their change ideas and try them out independently.

Unlike the Cherry County and Elm County district leads, the Cedar County leads also focused on capacity building across their three school teams to help teachers and interventionists learn how to integrate continuous improvement methods—even the sometimes time-consuming work of documenting PDSAs—into their regular work. With the assistance of regional office staff, the district leads offered school team members specific improvement coaching as they built their PDSAs in year 2. As a result, school team members in all three sites took up PDSAs to explore individualized questions about their instructional practice. Many of the school team members talked about how rewarding this specific PDSA work was for them; for instance, one teacher noted, in response to an interview prompt about what had gotten them excited about the NIC: "Just some of inquiries we've been working about with text-based questioning and just finding out that maybe I'm not pushing my students enough to dig into the text even in first grade. So finding out that I

wasn't doing as good of job as I thought I was with that." This teacher's excitement about finding this specific area for growth was echoed by their peers in similar ways, and also reflects the supportive environment in which this testing was happening. This evidence suggests that Cedar County's district context and improvement strategies created the necessary supports for teachers to talk about their failures and struggles and what they learned from them.

4.4.9 Sycamore County: Distributed Continuous Improvement Uptake Leads to Learning Through Struggle

In Sycamore County, continuous improvement work was distributed in that it was spread across four schools, where district leads and school team members' attempts to engage in small tests of change uncovered critical system characteristics in their district. The four Sycamore teams faced some difficulties as they started to engage in PDSA work around huddles in year 1, as they found that the strict structure of another reading initiative in their district left little room for the experimentation and flexibility that continuous improvement work requires. Within this inflexible context, school team members struggled to systematically make changes, collect data about those changes, and document their PDSAs. However, our interviews with them showed that this struggle in itself led to some important learning for both the improvement team members and their school and district administrators, that may enable future continuous improvement efforts.

In particular, school team members talked about how continuous improvement efforts to take up and refine huddle routines opened up communication between classroom teachers and the school and district level administrators. They noted that the culture in their district was very top down and "do what you're told," but that bringing outside perspectives into their schools through using improvement science methods helped create more "wiggle room" within that culture. In

comments at the a network wide meeting following the second year of the NIC, one of the district leads noted that the district administration had been "frozen in fear" for the first year of the project, since they had just recently enacted another literacy improvement initiative that had shown modest gains, so they were loathe to change anything too soon. However, the NIC helped them see that they needed to adopt a growth mindset and engage in some experimentation to learn their way into improvement: "sometimes you have to fail to grow," they reflected.

In keeping with the district's typical way of approaching reform, the Sycamore district leads initially mandated that each of their four school teams join the network; school team members noted that their participation in the network was a top down decision. This leadership aligns with the district reform environment described by school team members and district leads alike; however, in interviews at the end of year 2 of the project, several school team members expressed optimism about the ways in which attempting to use improvement science methods had opened up communication between classroom teachers and building and district administrators. This, coupled with the powerful organizational learning about engaging in experimentation to find ways to grow and improve, suggests that even grappling and struggling with continuous improvement can create important organizational shifts for schools and districts that do not quite have the capacity or resources to quickly become a traditional continuous improvement organization.

4.5 Discussion

These adaptive approaches to continuous improvement uptake and enactment illustrate how continuous improvement leaders in each case study site interpreted and brokered

improvement science methods in ways that were informed by their goals, capacities, and school system characteristics. In this sense, districts' different implementation approaches reflect the concept of adaptive integration, where the technical core of a process, practice, or tool is preserved while users make systematic changes around that core to tailor it to their local context (Hannan et al., 2015; Russell et al., under review*a*; Tichnor-Wagner et al., 2018). Traditionally, continuous improvement methods are seen as only the vehicle for this adaptive integration process, but in this NIC, district teams adaptively integrated continuous improvement methods *themselves* in diverse and creative ways. Even in the case of Sycamore County, where this adaptive work was not necessarily planned, we saw that the *attempted* use of improvement science methods, and the lessons that teachers and administrators took from those attempts, led to powerful organizational learning about the state of their system, and created cultural changes in their school that may lay the foundation for future continuous improvement work. These findings suggest that a more expansive and inclusive approach to continuous improvement research in education, one where schools and districts can take up these methods adaptively, depending on their system structures and capacities, could lead to substantive improvements across contexts, even where the load-bearing conditions usually associated with continuous improvement, such as well-developed data and measurement systems (e.g., Langley et al., 2009), are absent.

It is important to note, however, that there are some challenges associated with adapting improvement science methods. First of all, adaptive integration involves identifying the essential technical core of an innovation so that the spirit that animates the design is retained across contexts. Since there was no specific planning for the adaptation of improvement science in this project, in some ways these adaptive approaches lacked the specificity and intentionality of true adaptive integration. However, we think that the district leads' adaptive approaches to improvement science

uptake provide critical learning for researchers in that the demonstrated that different ways of doing this work can support organizational change and improvement in unanticipated ways.

In this network, some districts' adaptive work retained the technical core of improvement work more than others; Elm County in particular neglected many of the systematic elements of the process, and despite their positivity about their work, we think it is important to note that retaining the evidence-based, systematic orientation of improvement science is an important element of this work. Sycamore County, on the other hand, was able to retain many of the key elements of improvement science work because they supported classroom teachers in their efforts to use PDSAs. Even in cases where those PDSAs could not gain enough traction to create measurable improvement, the consistent attempted uses of the routine itself helped to illuminate important organizational structures that were barriers to this work. This suggests that additional research on this adaptive approach to improvement science should explore and identify the technical core of this methodology to help support educators as they make systematic adaptations around it.

Another key piece of this process was the brokering and boundary-spanning role of the district leads. The implementation literature has shown that brokers and boundary spanners are critical actors in organizational change efforts (Akkerman & Bakker, 2011; Fligstein & McAdam, 2012; Holland, 2014; Miller & Page, 2007), and this study indicates that they play a consequential role in NICs as well. District leads' approaches to enacting continuous improvement work included several different strategies, which mirrored some familiar processes from policy implementation literature. In particular, we saw district leads engaging in sensemaking, framing, buffering, persuading, and mandating (Coburn, 2001; Coburn & Woulfin, 2012; Weick, 1976) to facilitate the uptake of both continuous improvement methods and innovations in their districts. The differences in the way that each district took up continuous improvement research were closely

tied to the philosophies and goals of the district leads, suggesting that NICs and other similar collaborative problem solving research efforts (Penuel et al., 2017) ought to closely attend to the placement and brokerage processes of key actors in their networks. This supports concepts in both network and implementation literature that emphasize a structural approach to distribution of expertise (Coburn, Mata, & Choi, 2010; Coburn & Russell, 2008; Penuel, Riel, Kraus, & Frank, 2009; Russell et al., 2015), and echoes the idea from systems theory that a few key actors who exist in the boundaries or meso-levels of a system have enormous potential to drive change (Akkerman & Bakker, 2011; Fligstein & McAdam, 2012; Holland, 2014; Opfer & Pedder, 2011).

Observing patterns across the four districts also showed two different ways of structuring continuous improvement—with *centrally managed* and *distributed* approaches (See Figure 4.1). Improvement leads in the two smaller school districts, Elm County and Cherry County, created centrally managed continuous improvement systems by taking on the most of the responsibility for PDSA and inquiry work, whereas in Sycamore and Cedar County, two much larger districts, continuous improvement activities were distributed in that they were spread across multiple schools and enabled more independent capacity building for the school level teams. These structures dovetailed with each districts' continuous improvement goals and outcomes—in the centrally managed improvement districts, the team leads were aiming for messaging coherence and a common language for literacy learning, whereas in the distributed districts, teams focused more on aligned instructional improvement tailored to individual schools and classrooms. While we were unable to secure an interview with the district leads in Sycamore and so cannot speak to their explicit goals for continuous improvement work, their school team members shared many insights in their interviews that allowed us to understand their perspectives about what was needed in their district. Their accounts show that their distributed uptake of continuous improvement

methods helped teachers communicate with building and district leadership and show the necessity of greater flexibility in their instructional environment, and in this case, continuous improvement helped teachers to gather evidence to show the need for organizational change.

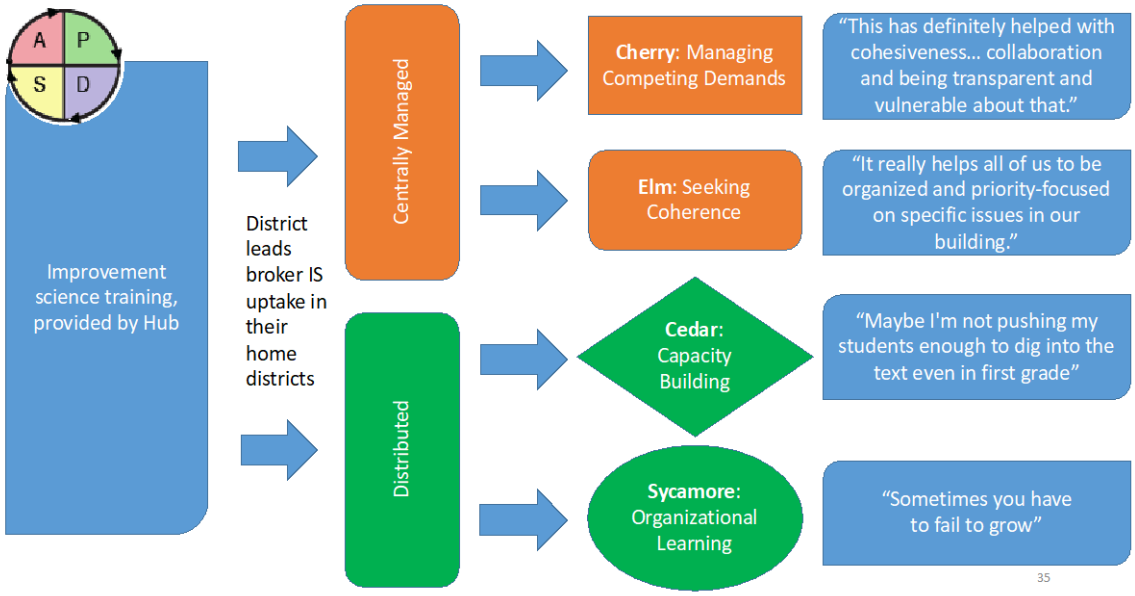

Figure 4.1. Adaptive Improvement Approaches

Across both the centrally managed and distributed districts, continuous improvement work supported movement towards greater instructional coherence, either through building and refining instructional guidance systems (Cherry County), establishing a common language and framework for literacy instruction (Elm County), or aligning literacy instruction around key tools and practices across a variety of contexts (Cedar and Sycamore). Research has shown that robust, aligned instructional systems can foster capacity for improving teaching and learning (Bryk et al., 2010; Honig & Hatch, 2004; Hopkins & Spillane, 2015). Furthermore, these efforts to increase coherence also created opportunities for informal interaction between teachers and other instructional professionals across all four case study districts, which the research has shown can also be an

accelerator for school improvement (Coburn & Russell, 2008; Penuel et al., 2017). These commonalities that emerged despite teams' very different approaches to continuous improvement work suggest that using CI methods in adaptive ways has untapped potential for generating powerful learning and improvement across local contexts.

Exploring the specifics of how each school district in the network took up continuous improvement methods also suggests that adaptive integration of these methods may be inevitable in educational organizations; the comparative findings from our case study were especially interesting since district teams were *not* instructed to take up improvement science adaptively. Rather, following traditional approaches to continuous improvement, the network Hub framed improvement science methods as a non-negotiable process that, when used with fidelity, would enable systematic integration of instructional innovations and changes. However, our cases show that district leads adapted continuous improvement methods to their contexts anyway, reflecting their ideas about school team and system capacities as well as their specific goals for continuous improvement in their district context. Since each adaptive approach led to meaningful improvement outcomes, we think that considering the potential of adaptive integration of improvement science methods is a key future direction for continuous improvement research in education. In particular, additional explorations of how to adapt continuous improvement methods, and what constitutes the indispensable, technical core of this methodology, will be of great value to the field.

4.6 References

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, *81*(2), 132-169.

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

Atteberry, A. and Bryk, A.S. (2010). Centrality, connection, and commitment: The role of social networks in school-based literacy initiative. In A. J. Daly (Ed.), *Social Network Theory and Educational Change.* Cambridge, MA: Harvard Educational Press.

Berman, P., & McLaughlin, M. (1978). *Federal programs supporting educational change: A model of education change, Vol. VIII: Implementing and sustaining innovations*. Santa Monica, CA: Rand.

Berwick, D. M. (2008). The science of improvement. *JAMA, 299*(10), 1182-1184.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. Halliman (Ed.), *Frontiers in sociology of education.* New York, NY: Springer Publishing.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. (2015). *Learning to improve.* Cambridge, MA: Harvard Education Press.

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

Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago.* Chicago, IL: University of Chicago Press.

Bush-Mecenas, S., & Marsh, J. A. (2018). The DIVE approach: Using case-ordered meta-matrices and theory-based data displays to analyze multiple case study data. In *Complementary research methods for educational leadership and policy studies* (pp. 33-56). Palgrave Macmillan.

Clarke, A., & Collins, S. (2007). Complexity science and student teacher supervision. *Teaching and Teacher Education, 23,* 160–172

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, *23*(2), 145-170.

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

Coburn, C. E., Mata, W. S., & Choi, L. (2013). The embeddedness of teachers' social networks: Evidence from a study of mathematics reform. *Sociology of Education*, *86*(4), 311-342.

Coburn, C. E., Russell, J. L., Kaufman, J. H., & Stein, M. K. (2012). Supporting sustainability: Teachers' advice networks and ambitious instructional reform. *American Journal of Education*, *119*(1), 137-182.

Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational Evaluation and Policy Analysis*, *30*(3), 203-235.

Coburn, C. E., & Woulfin, S. L. (2012). Reading coaches and the relationship between policy and practice. *Reading Research Quarterly*, *47*(1), 5-30.

Cohen, D. J., Crabtree, B. F., Etz, R. S., Balasubramanian, B. A., Donahue, K. E., Leviton, L. C., Clark, E. C., Isacson, N. F., Strange, K. C. & Green, L. W. (2008). Fidelity versus flexibility: translating evidence-based research into practice. *American Journal of Preventive Medicine*, *35*(5), S381-S389.

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

Cohen-Vogel, L., Cannata, M., Rutledge, S. A., & Socol, A. R. (2016). A Model of Continuous Improvement in High Schools: A Process for Research, Innovation Design, Implementation, and Scale. *Teachers College Record*, *118*(13), 1-26.

Deming, W. E. (2000). *Out of the crisis.* Cambridge, MA: MIT press.

DiMaggio, P. J., & Powell, W. W. (Eds.). (1991). *The new institutionalism in organizational analysis*. Chicago, IL: University of Chicago Press.

Fligstein, N., & McAdam, D. (2012). *A theory of fields*. Oxford, UK: Oxford University Press.

Hannan, M., Russell, J. L., Takahashi, S., & Park, S. (2015). Using improvement science to better support beginning teachers: The case of the building a teaching effectiveness network. *Journal of Teacher Education*, *66*(5), 494-508.

Holland, J.H. (2014). *Signals and boundaries: Building blocks for complex adaptive systems.* Cambridge, MA: MIT.

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

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, *33*(8), 16-30.

Hopkins, M., & Spillane, J. P. (2015). Conceptualizing relations between instructional guidance infrastructure (IGI) and teachers' beliefs about mathematics instruction: Regulative,

normative, and cultural-cognitive considerations. *Journal of Educational Change*, *16*(4), 421-450.

Innes, J. E., & Booher, D. E. (2010). *Planning with complexity: An introduction to collaborative rationality for public policy*. Routledge.

Iriti, J., Bickel, W., Schunn, C., & Stein, M. K. (2016). Maximizing research and development resources: identifying and testing "load-bearing conditions" for educational technology innovations. *Educational Technology Research and Development*, *64*(2), 245-262.

Langley, G. J., Moen, R. D., Nolan, K. M., Nolan, T. W., Norman, C. L., & Provost, L. P. (2009). *The improvement guide: A practical approach to enhancing organizational performance*. John Wiley & Sons.

LeMahieu, P. G. (2011). *What we need in education is more integrity (and less fidelity) of implementation*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

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

Lortie, D. (1975). *Schoolteacher: A sociological analysis*. Chicago, IL: University of Chicago Press.

Louis, K. S., & Leithwood, K. (1998). From organizational learning to professional learning communities. In K. Seashore Louis & K. Leithwood (Eds.) *Organizational learning in schools* (pp. 275-285). New York, NY: Taylor & Francis.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, *33*(2), 3-11.

Miles, M. B., Huberman, A. M., & Saldana, J. (2013). *Qualitative data analysis*. Sage.

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life.* Princeton, NJ: Princeton University Press.

Meyer, J. W., & Rowan, B. (1975). *Notes on the structure of educational organizations*. Palo Alto, CA: Stanford Center for Research and Development in Teaching.

O'Day, J. A., & Smith, M. S. (2016). Quality and equality in American education: Systemic problems, systemic solutions. In *The dynamics of opportunity in America* (pp. 297-358). Springer.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376-407.

Park, S., Hironaka, S., Carver, P., & Nordstrum, L. (2013). *Continuous improvement in education.* Carnegie Foundation for the Advancement of Teaching.

Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, *40*(7), 331-337.

Penuel, W.R., Peurach, D. J. et al., (2017). Defining collaborative problem solving research: Common values and distinctive approaches. Cedar Paper. Boulder, CO: University of Colorado Boulder School of Education, Research and Practice Collaboratory.

Penuel, W. R., Riel, M., Krause, A. E., & Frank, K. A. (2009). Analyzing teachers' professional interactions in a school as social capital: A Social network approach. *Teachers College Record*, *111*(1), 124–63.

Redding, C., Cannata, M., & Taylor Haynes, K. (2017). With scale in mind: A continuous improvement model for implementation. *Peabody Journal of Education*, *92*(5), 589-608.

Powell, W. W., & DiMaggio, P. J. (Eds.). (2012). *The new institutionalism in organizational analysis*. Chicago, IL: University of Chicago Press.

Russell, J. L., Bryk, A. S., Dolle, J. R., Gomez, L. M., Lemahieu, P. G., & Grunow, A. (2017). A framework for the initiation of networked improvement communities. *Teachers College Record*, *119*(5).

Russell, J.R., Correnti, R., Stein, M.K., Schwartz, N., Bill, V., Hannan, M., Booker, L. (under review*a*). Learning to leverage coaching for instructional coaching at scale: Adaptive integration in the Tennessee math coaching project. *American Educational Research Journal.*

Russell, J.R., Bryk, A.S., Khachatryan, E., Lemahieu, P.G., Peurach, D. Sherer, J.Z., Hannan, M. (under review*b*). The social structure of networked improvement communities: Cultivating the emergence of a scientific professional learning community. *American Educational Research Journal.*

Supovitz, J. (2008). Instructional influence in American high schools. In Mangin, M. M., Stoelinga, S. R. (Eds.), *Effective teacher leadership* (pp. 144-162). New York, NY: Teachers College Press.

Tichnor-Wagner, A., Wachen, J., Cannata, M., & Cohen-Vogel, L. (2017). Continuous improvement in the public school context: Understanding how educators respond to plan–do–study–act cycles. *Journal of Educational Change*, *18*(4), 465-494.

Tichnor-Wagner, A., Allen, D., Socol, A. R., Cohen-Vogel, L., Rutledge, S. A., & Xing, Q. W. (2018). Studying implementation within a continuous continuous-improvement process: What happens when we design with adaptations in mind? *Teachers College Record, 120*(5).

Scoville R, & Little K. (2014). *Comparing lean and quality improvement*. Cambridge, MA: Institute for Healthcare Improvement.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology. *Handbook of Qualitative Research, 17*, 273-85.

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

Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 1-19.

5.0 Conclusion

Taken together, these three papers demonstrate the organizational complexity that shapes education implementation research, and suggest that several different concepts, perspectives, and strategies can support researchers and policy-makers as they endeavor to create system-wide educational improvement. Paper 1 offers a theoretical argument to underpin the systems perspective that drives the empirical explorations in Papers 2 and 3, since, ultimately, both instructional coaching and networked improvement communities are system-focused interventions. This overarching conceptual focus highlighted several common themes in my findings, in particular: the value of opportunities for collaboration and the critical function of brokering and boundary spanning in organizational change efforts. Both empirical studies engaged deeply with the idea of adaptation based on local contextual conditions, mirroring the trend in educational research funding that asks not only what works, but how, why, and where different kinds of interventions impact educational outcomes (e.g. Bryk et al. 2015). Furthermore, the first paper highlights the ways these processes of adaptation are embedded in system structures, and together the three papers contribute to the implementation literature by elaborating a structural approach to implementation research and building theory about how this affects the dynamics of uptake and adaptation.

5.1 In Educational Improvement Efforts, Collaboration Matters

Although a key finding in Study 2 showed that the character of a school system—particularly, its urbanicity—can shape the conditions needed to enable robust coaching practices, there was one organizational resources that supported rigorous coaching across all kinds of systems: access to collaboration. While this looked a bit different based on whether coaches were urban or rural, the importance of opportunities to collaborate with peers and other supportive colleagues emerged as a common pattern across all coaches in all kinds of contexts. These findings extend the existing literature on teacher collaboration, which has shown that teachers' capacities for instructional improvement, innovation, and reform are bolstered by an environment characterized by trusting interactions and access to supportive peer groups (Bryk & Schneider, 2002; Coburn, Choi, & Mata, 2010; Coburn & Russell, 2008; Frank et al. 2011; Frank et al. 2004; Molenaar et al., 2011; Spillane et al. 2002).

In Paper 3, collaboration played a similarly critical role; while each district-based improvement team took up and adapted continuous improvement methods in slightly different ways, every single district used those methods to enable some kind of improved collaboration in their schools. Creating space for collaboration was also something that school-level team members, such as teachers and interventionists, were particularly excited about. This is especially important when considering how to generate the will to use improvement science in schools, which can be difficult given the additional work and capacity building it often requires. Since the teacher collaboration literature has shown that access to supportive collaboration can accelerate instructional improvement (Bryk et al. 2011; Bryk & Schneider, 2002; Coburn et al. 2012; Frank et al. 2004), expanding this notion to consider how access to collaboration may enable collaborative problem solving research in schools is a promising new direction for research.

Further, across Papers 2 and 3, participants engaged in collaborative processes to enable complex changes in their school systems: for coaches, this was working with teachers to create greater instructional alignment, coherence, and quality, and for district improvement leaders, this meant establishing continuous improvement processes and routines to learn about how to improve literacy in their schools and districts. The promising results of the collaborative efforts in these studies mirror other recent arguments in the collaboration literature that suggest that designing interventions that engage multiple perspectives and directly account for organizational structures will continue to support growth and improvement in educational research (Penuel et al. 2017). Furthermore, the conceptual framework developed in Paper 1 suggests that these interactions are multi-directional learning processes that can contribute to the strength of the overarching system, which is a key implication for future research. Further studies of collaborative implementation work could continue to explore the potential of these interactions as learning resources. Finally, together these studies start to explore ways to structure collaborative improvement efforts by explicitly attending to the boundary spaces where change is created.

5.2 Brokering and Boundary Spanning Are Essential Mechanisms for Change

Returning to the concept of brokering, these three studies show that attending to key brokers and boundary spanners in educational improvement efforts can be a viable implementation strategy, since these actors who span discrete system locations have considerable potential to create change and improvement. This idea is key to conceptualizing the power and potential of both instructional coaching and NICs' model of improvement science uptake; both processes designate a few key actors who bridge different organizational locations and engage in brokering

work to foster change and improvement. For coaches, this brokering and boundary spanning has been identified in the literature (Coburn & Woulfin, 2012; Woulfin, 2014) but the research has yet to fully explore how this can work in different contexts and initiatives.

Paper 2 starts to piece together the ways in which instructional systems and opportunities for collaboration shape these brokering processes, but as I have mainly focused on building theory about the critical dimensions of coaching context, more research is needed on the brokering function of instructional coaches. Coupled with the contextual specificity I argue for in Paper 1, future explorations of coach brokering that presume a theory of embeddedness have the potential to contribute crucial information to the field's understanding of instructional coaching and its effects. Additionally, this paper highlights the need for supporting coaches and other boundary-spanners with opportunities to collaborate to support the key change processes that they drive in systems.

Paper 2 also offered insights into the conditions that may support coaches as effective change and improvement brokers in complex systems; in particular, the finding that greater levels of autonomy and access to centralized coaching resources at the district level were often present for more rigorous coaches suggests that these conditions may attenuate other challenging circumstances that coaches often face in their work. Since these more rigorous coaches generated teaching improvement at faster rates, these conditions may be important supports for the challenging boundary-spanning work that coaches do. This finding ties in to other research on how networks and relationships embedded in school organizations can facilitate change (Honig & Hatch, 2004; Russell & Bray, 2013; Spillane, 2002), and emphasizes the importance of taking system structures into account in coaching research. This study also starts to build an

understanding of how to create coaching supports that are sensitive to contextual conditions, which is a key question in the coaching literature.

Furthermore, Paper 3 found that improvement leaders in a NIC also engage in coaching moves as they adapt and translate continuous improvement practices for their home schools and district; in our exploration of improvement leads' adaptive integration of continuous improvement work in their home organizations, we found that they engage in buffering, brokering, mandating, and sensemaking about continuous improvement in a way that is very similar to what Coburn and Woulfin (2014) observed about instructional coaching. This parallel between instructional coaching and other kinds of boundary-spanning roles in education points to another direction for future research that would explore these critical brokers in NICs and other collaborative problem solving efforts. It also connects to the arguments in the networks and collaboration research for taking a structural view of diffusion and distribution of expertise (Coburn, 2001; Coburn, 2006; Coburn & Russell, 2008; Coburn, Choi, and Mata, 2010), suggesting that future studies of continuous improvement research in education could explore strategies for leveraging organizational structures beyond NIC boundaries to facilitate adaptive integration of improvement science approaches. In particular, these studies could investigate ways to maximize the potential of key actors in organizations who move between multiple system levels and locations to negotiate changes and improvement.

5.3 Expanding the Notion of Adaptive Integration

Finally, both empirical studies in this dissertation take on the idea of adaptation within implementation processes; in the case of instructional coaching, I considered how various local

contexts interacted with individual coaching processes, and in the case of continuous improvement uptake, I found that adaptive approaches to improvement science methods themselves can support organizational learning across diverse contexts with differing capacities. This notion reflects the idea of "mutual adaptation" from the implementation literature—studies of implementation have often found that the adaptation process involved in the education reform are bi-directional in that educators tend to change both their existing work processes and routines and the innovation itself as they incorporate something new into their practice (Berman & McLaughlin, 1978; Coburn 2004; Honig, 2006; Spillane et al. 2002). This process, however, is not typical in education, particularly in efforts to make substantial changes to teaching practice (Coburn et al. 2001; 2004; Reiser et al., 2000).

The studies in this dissertation explore mutual adaptation deeply by focusing on new applications of the idea of adaptive integration. Paper 3 in particular contributes to the literature on adaptive integration by expanding our conception of the term—while, in the past, it has been applied to tools and processes that are adapted using improvement science methods (e.g. Bryk et al., 2015; Hannan et al. 2015; Cohen-Vogel et al. 2015), here I build on the gap identified by Tichnor-Wagner and colleagues (2018) to argue that continuous improvement methods themselves can have more utility when they are tailored to individual contexts. To more fully understand this finding, more research is needed on understanding the indispensable technical core of improvement science methods; in order to engage in adaptive integration of continuous improvement methods with integrity, the critical design elements of the method must be more further explored and clearly delineated. Networks that facilitate adaptations across contexts can help to build insight into these design elements by comparing different approaches to adaptation and evaluating the extent to which those approaches support improvement. This comparative

element of improvement networks can also help to elucidate strategies for tailoring continuous improvement methods to organizational structures, which is another key component of this adaptation process. Finally, the insights in paper 1 can contribute to building out the systems focus in continuous improvement research in education and provide a foundation to structure future studies of how continuous improvement processes interact with existing system structures and characteristics.

5.4 References

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, *81*(2), 132-169.

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333, 1034-1037.

Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, *13*(1), 1-14.

Biancarosa, G., Bryk, A., & Dexter, E. R. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *Elementary School Journal*, *3*(1), 7-34.

Blazar, D., & Kraft, M. A. (2015). Exploring mechanisms of effective teacher coaching: A tale of two cohorts from a randomized experiment. *Educational Evaluation and Policy Analysis, 37,* 542-566.

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*. New York, NY: Springer.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. (2015). *Learning to improve.* Cambridge, MA: Harvard.

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *The Elementary School Journal*, *111*(3), 430-454.

Cannata, M., Cohen-Vogel, L., & Sorum, M. (2017). Partnering for improvement: Improvement communities and their role in scale up. *Peabody Journal of Education*, *92*(5), 569-588.

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

Coburn, C. E., & Penuel, W. R. (2016). Research–practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, *45*(1), 48-54.

Coburn, C. E., & Woulfin, S. L. (2012). Reading coaches and the relationship between policy and practice. *Reading Research Quarterly, 47*(1), 5-30.

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

Fligstein, N., & McAdam, D. (2012). *A theory of fields.* Oxford, UK: Oxford.

Hannan, M., Russell, J. L., Takahashi, S., & Park, S. (2015). Using improvement science to better support beginning teachers: The case of the building a teaching effectiveness network. *Journal of Teacher Education*, *66*(5), 494-508.

Holland, J.H. (2014). *Signals and boundaries: Building blocks for complex adaptive systems.* Cambridge, MA: MIT Press.

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

Huguet, A., Marsh, J. A., Farrell, C. C. (2014). Building teachers' data-use capacity: Insights from strong and developing coaches. *Education Policy Analysis Archives*, *22*(52), 1-28.

Kraft, M. A., Blazar, D., & Hogan, D. (2016). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 0034654318759268.

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

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

Mangin, M. M., & Dunsmore, K. (2015). How the framing of instructional coaching as a lever for systemic or individual reform influences the enactment of coaching. *Educational Administration Quarterly*, 0013161X14522814.

McLaughlin, M. W. (1991). The Rand change agent study: Ten years later. In A. Odden (Ed.), *Education policy implementation* (pp. 143-155). Albany, NY: State University of New York Press.

Miller, J. H., & Page, S. E. (2007). *Complex adaptive systems: An introduction to computational models of social life*. Princeton, NJ: Princeton University Press.

Morris, A. K., & Hiebert, J. (2011). Creating shared instructional products: An alternative approach to improving teaching. *Educational Researcher*, *40*(1), 5-14.

Mudzimiri, R., Burroughs, E. A., Luebeck, J., Sutton, J., & Yopp, D. (2014). A look inside mathematics coaching: Roles, content, and dynamics. *Education Policy Analysis Archives*, *22*(53), 1-32.

Obara, S., & Sloan, M. (2009). The evolving role of a mathematics coach during the implementation of performance standards. *Professional Educator*, *33*(2), 11-23

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376-407.

Penuel, W.R., Peurach, D. J. et al. (2017). Defining collaborative problem solving research: Common values and distinctive approaches. White Paper. Boulder, CO: University of Colorado Boulder School of Education, Research and Practice Collaboratory.

Powell, D. R., Diamond, K. E., Burchinal, M. R., Koehler, M. J. (2010). Effects of an early literacy professional development intervention on Head Start teachers and children. *Journal of Educational Psychology*, *102*, 299-312.

Tichnor-Wagner, A., Allen, D., Socol, A. R., Cohen-Vogel, L., Rutledge, S. A., & Xing, Q. W. (2018). Studying implementation within a continuous continuous-improvement process: What happens when we design with adaptations in mind? *Teachers College Record, 120*(5), 1-52.

Russell, J. L., Bryk, A. S., Dolle, J. R., Gomez, L. M., Lemahieu, P. G., & Grunow, A. (2017). A framework for the initiation of networked improvement communities. *Teachers College Record, 119*(5), 1-51.

Sailors, M., & Price, L. R. (2010). Professional development that supports the teaching of cognitive reading strategy instruction. *The Elementary School Journa*l, *110*(3), 301–322.

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.

Woulfin, S. (2014). Charting the research on the policies and politics of coaching. *Education Policy Analysis Archives, 22*, 2-12.