

With Scale in Mind: A Continuous Improvement Model for Implementation

Christopher Redding, Marisa Cannata, Katherine Taylor Haynes

This research was conducted with funding from the Institute of Education Sciences (R305C10023). The opinions expressed in this report are those of the authors and do not necessarily represent the views of the sponsor.

This is an Accepted Manuscript of an article published by Taylor & Francis in Peabody Journal of Education on October 3, 2017, available at <https://doi.org/10.1080/0161956X.2017.1368635>.

The conventional approach to school implementation involves different sets of schools in designing, developing, and testing an innovation and yet another set of schools involved in implementation at scale. Decades of research suggest that this approach has yielded inconsistent and often disappointing improvements for schools most in need. More recent scholarship on school improvement suggests these activities should be integrated, although this presents challenges in how we evaluate implementation in particular schools. This paper presents a framework to conceptualize implementation when design, implementation, and scaling up are integrated activities.

Introduction

Despite decades of ambitious reform, high schools have generally been unable to improve students' academic performance, particularly among students from traditionally lower-performing subgroups (Becker & Luthar, 2002; Davison, Young, Davenport, Butterbaugh, & Davison, 2004; Lee, 2004; Mazzeo, Fleischman, Heppen, & Jahangir, 2016). Practices that may be effective in certain sites are spread inconsistently to new school contexts (Datnow, Hubbard, & Mehan, 2002; Fullan, 2000). Traditional implementation efforts have demonstrated the many challenges inherent in designing, implementing, and scaling up educational innovations. These challenges include the lack of teacher buy-in and participation (Datnow et al., 2002; Glennan, Bodilly, Galegher, & Kerr, 2004; Nunnery 1998), inadequate knowledge of the design (Spillane, 1999; Spillane, Resier, & Reimer, 2002) insufficient capacity to implement the more ambitious aspects of a design (Cohen, Peurach, Glazer, Gates, & Goldin, 2013; McLaughlin, 1987), adaptation to the point that the original design loses coherence (Desimone, 2002), and too little attention to the organizational context in which the practices are to be implemented (Bodilly, 1998; Elmore, 1996; Fullan, 2001; Stringfield & Datnow, 1998). Additional problems emerge when the goal of implementation is not just the adoption of a new program or practice, but to have the practice scaled up into all corners of a school (Elmore, 1996). Perhaps not surprisingly, then, research on how programs achieve their intended impact on outcomes when scaled up has

found that the positive impacts on outcomes diminish as the program is scaled (Chambers, Glasgow, & Stange, 2013).

In recognition of these challenges, the last decade has seen the proliferation of new approaches to achieving school improvement at scale. These approaches take many forms, including improvement science, research-practice partnerships, design-based implementation research, but what they share is a focus on researchers and practitioners working in partnership on shared problems and using continuous improvement methods to design, implement, study, and iterate on issues of school improvement (Bryk, Gomez, Grunow, & LeMahieu, 2015; Coburn & Penuel, 2016; Cohen-Vogel et al., 2015; Cohen-Vogel, Cannata, Rutledge, & Socol, 2016; Fishman, Penuel, Allen, & Cheng, 2013). In other words, these approaches recognize that design, implementation, and scale are not distinct phases of improvement work, but are interrelated. Successful scale up depends on characteristics of the innovation itself, the contexts in which it will be implemented, and the organization that designed the innovation, suggesting that how these components interact is critical for improvement (Cohen, Peurach, Glazer, Gates, & Goldin, 2013). Indeed, innovation designs vary in how “scalable” they are, and designers that want to be successful at scale need to design for scalability (Clarke & Dede, 2009). Likewise, as the innovation is implemented in more contexts, the designers can learn from how the innovation is adapted, allowing the innovation to evolve over time (Dede, 2006).

These new approaches to scale have contributed to our understanding of school improvement by highlighting the processes used to achieve scale, including the development of a culture of learning and continuous improvement (Fullan, 2016). For example, considerable research has been paid to the role of improvement infrastructure and how these structures, supports, or norms shape educational change (Hopkins & Woulfin, 2015; Sabelli & Dede, 2013).

The types of improvement infrastructure provided by reform organizations to implement their program can shape the extent to which educators engage with the program (Peurach, 2011; Peurach & Neumerski, 2015). Through developmental evaluation, researchers can explore aspects of the reform organization to understand the extent to which it has established processes and capacities to support improvement at scale (Peurach, Lenhoff, & Glazer, 2016).

Current efforts to scale up effective practices also emphasize the important role of local context and the need to engage local educators in developing the innovation and building understanding about it (Bryk, Gomez, Grunow, & Hallinan, 2011; Fishman et al., 2013). Indeed, this work emphasizes a relational approach to scaling up, recognizing that practitioners are not passive recipients of reform knowledge (Hartmann & Linn, 2008). Despite this emphasis on the need to consider context in scale up efforts, there is little conceptual understanding of what happens when reform designs meet schools, and the ways in which schools, and the educators working within them, change as a result of this reform engagement. This article reviews current research on implementation and develops a framework for evaluating implementation in the context of new approaches to improvement at scale. In doing so, we define implementation as the process of “scaling in” to a school, recognizing that part of achieving scale is enacting deep change in the beliefs, norms, and practices in schools (Coburn, 2003). By distinguishing scaling in to a school as distinct from scaling out to other schools and changing the institutional environment of the school, we emphasize that scaling up includes both horizontal and vertical aspects (World Bank, 2003). Further, shifting from a focus on implementation to scaling in to a school highlights how the process can evolve over time, perhaps deepening or establishing more sustainable infrastructure in the school.

In developing this framework of implementation in the context of continuous improvement approaches to scale, we aim to integrate literature around both implementation and scale, highlighting how traditional approaches to conceptualizing implementation need to shift to encompass a broader approach to scale. We also draw heavily from Coburn's (2003) framework in which scale encompasses *depth*, *spread*, *shift in reform ownership*, and *sustainability* to demonstrate how reconceptualizing implementation as scaling in is aligned with this broader understanding of scaling up. We illustrate this framework with examples from the National Center on Scaling Up Effective Schools (NCSU), that has developed an integrative model of design and development, implementation, and scale up that addresses the persistent research-to-practice gap. Without laying the groundwork for scaling up during initial implementation (scale-in), we assert that scale up will remain elusive. To realize the goal of scaling up practices within and across schools, a cohort of implementers must be involved in the design and development of an innovation to build the requisite will, beliefs, and capacities to enable implementation and scale up.

First, we briefly summarize our integrative approach of design and implementation and describe how we envision reaching scale. We then present three elements of our implementation framework: facilitating conditions, implementation supports, and evaluating the quality of implementation. These elements indicate the underlying school conditions that promote initial implementation, the structures put in place to support implementation, and how we propose to evaluate the quality of implementation. We draw on the school improvement and implementation literatures to situate our approach and describe how we conceive of changes in the implementation of an innovation contributing to a greater likelihood of achieving scale up of the innovation.

Overview of the National Center on Scaling Up Effective School

The National Center on Scaling Up Effective Schools (NCSU) develops a new model for school systems to scale up practices of effective high schools. Traditionally, the model for identifying, developing, testing, and then implementing an innovative practice at scale has separated each of these phases and the work has typically occurred in distinct locations. For example, one set of schools may be involved in identifying and developing an innovation, another set of schools involved in testing the innovation, and yet another set of schools involved in implementation at scale. In contrast, National Center on Scaling Up Effective Schools’s model of scale up situates all of these phases of the work within a single district context—and involves school and district personnel in each phase to ensure alignment with existing priorities and unique circumstances. National Center on Scaling Up Effective Schools is a partnership between research universities, education innovation development specialists, and school districts. Our work involves four phases: Research, Innovation Design and Development, Implementation, and Scale Up. In the Research Phase, we study both higher and lower value-added schools in the same district to identify the programs and practices that may explain the differences in their performance. The research findings become the “design challenge” that guides a collaborative design and development process. In the Innovation Design and Development Phase, a district wide team is established to take the design challenge and design an innovation that addresses the challenge. Through a continuous improvement process, school teams are established to test, refine, and further develop the innovation and adapt it to their unique school context. By successively enlarging the scope of testing, the Design Phase gradually evolves into the Implementation Phase. The final phase, Scale Up, occurs as the testing and implementation

process involves additional schools and the [BLINDED FOR PEER REVIEW] gradually transfers leadership to the district.

Conceptualizing Implementation as Scaling In To a School

The implementation process has been extensively studied both inside and outside education, and numerous frameworks for understanding and evaluating implementation have been developed (Berends, Kirby, Naftel, & McKelvey, 2001; Desimone, 2002; Durlak & Dupre, 2008; Meyers, Durlak, & Wandersman, 2012). This research has generally focused on characteristics of the teachers, schools, and districts in which implementation is occurring, characteristics of the structures and processes established to help personnel learn about the innovation they are implementing, and how the innovation shapes behaviors and practices in the implementing school.

In reconceptualizing implementation as a process of scaling in to a school, this framework draws heavily from existing research on implementation. As such, the major components of the framework include attention to the facilitating conditions, implementation supports, and indicators of quality implementation that leads to scale. In short, there are a set of *facilitating conditions* (will, beliefs, implementation capacity, and alignment with local context) that set the stage for successful implementation. These facilitating conditions exist prior to implementation, but also can be changed as the innovation is scaled in to the school. For example, alignment to local context may shift as both the context and the demands of the innovation evolve over time. As Dede (2006) defines scaling up as closing gaps between an innovation's demands and the organization's capacity, scaling in to a school should result in changes to the facilitating conditions. The *implementation supports* represent the structures and processes put in place specifically to support the implementation of the innovation. The

implementation supports include establishing implementation teams, developing an implementation plan, allocating resources, ongoing technical support, and engaging in a continuous improvement process. The third component of our implementation framework focuses on the *quality of implementation*. This includes examining how behaviors of individuals within the school change to be more aligned with the expected behaviors of the innovation. This part of the framework focuses on integrity and frequency of behaviors, the program reach, and participant responsiveness. Similar to the other two components, the deepening and spreading of these behaviors over time represents successful scale in to the school.

Ultimately, outcomes of successful implementation are improvements in proximal and distal student outcomes. Figure 1 provides a visual representation of how these components are related and are envisioned as achieving scale. The section below outlines each of these components and demonstrates how the evolution of these components in schools can be measured over time to examine how the innovation is scaled in to the school. For example, while will may be initially lacking when an innovation is first introduced, increasing will to engage with the innovation reflects shifting ownership of the innovation from an external organization to the school. Likewise, the integrity of innovation practices to the core design may be low at first, but may achieve greater depth of change over time.

While this framework draws on existing frameworks of implementation, there are a few notable ways in which our approach to implementation departs from these more conventional models. The largest difference pertains to the specificity of the implementation plan. Fixsen and colleagues' (2005) comprehensive review of implementation research links successful implementation with a clearly conceptualized program design and operationalized program components. This view contends that a detailed program treatment plan is linked to less variation

in how a program was implemented, a greater likelihood of detecting changes in student outcomes, and possibly greater program outcomes (Blakely et al., 1987; Dusenbury, Brannigan, Falco, & Hansen, 2003; O'Donnell, 2008; Weiss Bloom, Brock., 2014). As argued above, this view understates the influence of local organizational factors in supporting successful scale-in. We further contend that not addressing local contextual factors limits the chances of an innovation becoming embedded in the norms of a school, influencing the depth of teacher practice, and becoming sustainable practices within schools and even the district.

By opting for a more contextually based approach to implementation that seeks local expertise to inform the innovation design (Cohen-Vogel et al., 2016), this framework emphasizes a continuous improvement process that seeks to further develop practices related to the innovation as aligned in various ways with the innovation's broad initial goals. This focus shifts the implementation process towards a capacity-building approach, where school stakeholders, in partnership with researchers and program designers, hone in on the core set of practices that is aligned with the innovation's goals and the broader school culture. Implementation, then, is not just a technical practice, but a "profoundly developmental practice" where stakeholders learn (and unlearn) powerful educational ideas and practices (Elmore, 2016, p. 531). This framework also reflects how the process of scaling in an innovation to a schools takes place in a multi-layered system where the innovation itself may shift as it is co-constructed by implementers (Datnow & Park, 2009).

Facilitating Conditions of Successful Implementation

Belief, will, capacity, and alignment with local context comprise what we refer to as the *facilitating conditions* for successful implementation. The shortcomings of traditional educational implementation have been attributed to the absence of these elements. We

understand these facilitating conditions as the *willingness* to alter current practices in favor of new, promising approaches, *capacity* as the individual ability and organizational conditions to engage in the behaviors implementation requires, and *beliefs* as both the knowledge of the innovation and the feeling that the proposed innovation will address a problem. Finally, we emphasize *alignment with the norms, beliefs, and priorities of the partner schools and districts*, including other district and state policy mandates.

The initial motivation to address these facilitating conditions comes from the recognition of how pre-existing school practices shape implementation. Yet, successful implementation alone is insufficient to achieve scale up, as innovative educational practices are often isolated in “pockets of excellence” (Elmore, 1996, p. 1). In describing these facilitating conditions, we also describe how consciously addressing will, belief, capacity, and alignment with the local context ensures that school and district stakeholders also have the capacity and are willing to sustain implementation once external support is withdrawn.

Will

Will or buy-in have been discussed as aiding successful implementation at individual and collective levels. At the individual level, will has been emphasized in the education reform literature as a way to focus attention on individuals as integral participants in the change process (Cohen, Raudenbush, & Ball, 2002; Hess, 1999; McLaughlin, 1987). Will is generally defined as the attitudes and motivation that ground stakeholders’ response to external policies (McLaughlin, 1987). Datnow and Castellano (2000) describe the willingness to support reform as related to more fundamental conceptions of the value of education, teaching, and schooling. They describe that with schools that adopted the Success for All reform model, teachers and school staff had recognized the need to support students in new ways in order to raise student achievement. In

this framework, the definition of will is focused on the motivation and desire for stakeholders to assume responsibility for implementation and strive to do what the innovation requires or encourages. We argue that this emphasis on school stakeholder will is necessary to build the necessary ownership required to sustain the innovation. NCSU also attempts to develop will among the core implementation team at each innovation school. Involving them in the design, development, and implementation processes engenders buy-in and commitment. By consciously cultivating will among a core group of school staff, this conception of will moves beyond reform as compulsory to one in which teachers' and school leaders' ownership over design and implementation process lays the groundwork for scale up.

We also recognize a collective dynamic of will to indicate how the social and organizational context influences teachers and other school personnel's' receptiveness to reform and follow-through in implementing its practices. Scholars of comprehensive school reform (CSR) models have similarly recognized the need for collective will as one of the most important components in scaling up and sustaining innovations (Datnow, 2000; Desimone, 2002). Yet, CSR models often treat buy-in as a prerequisite for reform rather than something to be developed through implementation itself. For instance, Success for All would only work in schools where 80 percent of teachers had voted in favor of adopting the reform model (Vernez, Karam, Mariano, & DeMartini, 2006). By conceptualizing implementation as the process of scaling in to a school, stakeholder will is not just a pre-existing factor, but a factor that could evolve over time as greater support for implementation is developed. The potential for will to increase magnifies the chances that stakeholders will remain committed to implementing the innovation and may assume greater ownership, a component of scale.

Belief

A consistent theme in the educational implementation literature is the recognition that school stakeholders' knowledge of the innovation design does not equate with the reformers' initial intent, but is subject to teachers' understanding of how their students learn and their predisposition towards practices that are aligned with preexisting knowledge, beliefs, and experiences, and classroom norms (Coburn, 2004; Spillane, 2000; Spillane & Callahan, 2000; Spillane & Jennings, 1997). Spillane and colleagues (2006) illustrate, "[W]hen asked to interpret a proposed instructional practice...[a teacher] applies tacit knowledge about children and the discipline to mentally envision the situation and draw inferences about how effective that practice would be" (p. 51). This aspect of the research-practice gap is addressed in this framework by involving district and school stakeholders in the initial design of the innovation. As collaborators in the design, stakeholders develop an in-depth knowledge of the design and the research that informed it. As Elmore puts it, achieving scale is focused on "demonstrating that powerful ideas can work in diverse environments" (2016, p. 533). Scale in, then, is focused on the ideas of learning and helping educators develop a shared understanding of the theories of action behind these powerful ideas (Fullan, 2016; Gallagher, Malloy, & Ryerson, 2016).

By engaging in co-construction of the innovation, members of the school implementation team can develop an in-depth understanding of the innovation design. The presence of an improvement community within the school provides a more long-lasting support system than teachers would find with the implementation of an "off-the-shelf" reform model. Indeed, developing learning communities focused around the innovation practices are critical for sustaining the innovation in a school and achieving scale (Elmore, 2016; Newman, Zacamy, Lazarev, & Lin, this issue). As local experts, they can help develop a deeper understanding of

how the innovation design may lead to what Coburn (2003, p. 4) describes as “change that goes beyond surface structures or procedures (such as changes in materials, classroom organization, or the addition of specific activities) to alter teachers’ beliefs, norms of social interaction, and pedagogical principles.” We argue that knowledge of the innovation and belief in its efficacy are precursors to achieving this depth of practice.

Capacity

In the broader implementation literature, Wandersman and coauthors (2008) argue that “capacity is central to addressing the gap between research and practice” (p. 173). Capacity for scale-in entails both individual skill and effective organizational structures to make successful implementation possible. Individual capacity focuses primarily on the skills and expertise of teachers and administrators within the school. Developing their capacity is vital to creating sufficient expertise and knowledge of the reform to ensure sustainability of the reform—another component of Coburn’s conception of scale up—once external support is withdrawn. At the collective level, capacity refers to the human capital, structures, and culture that facilitate successful implementation. The implementation literature identifies several organizational characteristics that predispose some organizations to more successful implementation than others. These include strong leadership, professional capacity, and a supportive climate (Durlak & DuPre, 2008). More specifically, we use indicators of organizational capacity from the school improvement literature that include the history of collaboration in the school, stability of faculty and study body, and trust among school staff (Bryk and Schneider, 2002; Bryk et al., 2010; Johnson Kraft, & Papay, 2012; Murphy and Torre, 2014; Newmann, Smith, Allensworth, & Bryk, 2001). Gauging the level of organizational capacity can help determine whether certain schools may need additional technical support in order to implement successfully. For instance,

schools with high annual turnover may need to develop a particularly robust system of professional development to integrate new teachers. The level of organizational capacity may also inform the adaptations that a school implementation team chooses to make depending on the organizational culture of each school.

Ideally, both individual and collective capacity to enact the innovation are increased over time. This includes the individual capacity for educators to engage in specific innovation practices, but also the organizational capacity to support individuals and organize collective action. As an innovation is scaled in to a school, the organizational structure or routines should evolve to provide more supportive infrastructure for innovation practices (Sabelli & Dede, 2013). Through these mechanisms, greater scale is achieved as the organization has the individual and collective capacity to sustain innovation practices.

Alignment to Local Context

Local context often interferes with successful implementation as a result of changes in district and state policy context (Datnow, 2005), conflict between programs within a district (Berends et al., 2002; Datnow, McHugh et al., 1998; Stringfield, Datnow et al., 2000), and other unforeseen challenges brought by the local environment or individual actors (McLaughlin, 1987). Given this influence, scholars warn that policymakers often pay insufficient attention to local context (Bodilly et al., 1998; Elmore, 1996; Fullan, 2001; Stringfield & Datnow, 1998). Involving local stakeholders in the design *and* implementation process acknowledges the many ways that local context shapes what is actually embedded in the innovation schools. We use the language of alignment as a way to describe how the innovation design connects disparate school practices.

Alignment occurs in a few different ways. At the most basic level, alignment takes place when the connections between the innovation and school vision or priorities are made explicitly. In a school where the principal has previously emphasized the importance of teacher-student relationships, the introduction of an innovation with formal systems for these interactions could be easily aligned to current school practices. The introduction of the innovation may also serve as a focal point to unite individual efforts around a shared goal. For example, in the personalization of students' academic and socioemotional learning scenario, the innovation may provide structures and protocol to support teacher engagement in these practices, beyond a principal's encouragement. Another form of alignment pertains to the district context. A major challenge to implementing comprehensive school reform successfully was shifting state and district policies. Changes in state standards and the conflict between elements of a reform model and district policies undermined alignment (Datnow, 2000; Desimone, 2002; Vernez et al., 2006). By partnering with districts to design and implement reform meant to address an ongoing need in the district, the goals of the innovation may merge with the larger vision of the district. Coburn argues that "teachers and schools are more likely to be able to sustain and deepen reform over time when school and district policy and priorities are compatible or aligned with reform" (p. 7). In other words, alignment becomes a tool to achieve greater consistency among school and district policies.

At the school level, achieving alignment is placed in tension with the intended implementation of the design. Promoting alignment risks compromising central features of the design when educators alter the design to fit their local demands. This possibility is especially concerning given the high frequency with which schools would implement CSR practices to the point where they lost coherence (Desimone, 2002; Berends et al., 2002). To address this concern,

the final way to achieve alignment is adaptation, the process of altering an innovation's design to fit the local school context. In the broader implementation literature, adaptation is contrasted with fidelity, which is typically equated with more positive outcomes (Blakely et al., 1987; Dusenbury et al., 2003; O'Donnell, 2008). We argue that the complexity of school improvement precludes this approach. Instead, we recognize adaptation as potentially beneficial as an innovation can be productively tailored to a specific context (Datnow & Castellano, 2000; Debarger et al., 2013; Kirshner & Polman, 2013; McLaughlin, 1976; Penuel et al., 2011). Several scholars have described the process by which adaptation occurs including McLaughlin's (1976) discussion of *mutual adaptation*, Datnow and Park's (2009) conceptualization of *co-construction*, and Supovitz's (2008) theory of *iterative refraction*. These authors emphasize that adaptation to local context yields a variety of outcomes that are often only partially related to the original policy or program design. We draw on these conceptualizations of adaptation to frame how this process could be guided to increase alignment of the innovation design and existing school practices.

NCSU attempts to manage adaptation in a productive manner, treating adaptation as a tool to learn about local context and how elements of the innovation design are most beneficial for students of various grade and ability levels and across different schools. This process is described in greater detail below. In the context of the facilitating conditions, we note that the negotiation of program components and a schools' local context is rarely straightforward given the organizational complexity of high schools (Siskin, 2012). NCSU builds this knowledge and skill through the school and district team members' involvement in the design and ongoing development of the innovation. A goal of capacity-building among the implementation teams is to prepare members to manage what we refer to as *principled adaptations* made at the school

level. These adaptations are “principled” in that implementers are making decisions about what to put in place at their school, recognizing the goals of the innovation and the unique context and priorities of the school.

Implementation Supports for Successful Implementation

In addition to these facilitating conditions, successful scale-in requires several structural supports. From Meyers, Durlak, and Wandersman (2012) Quality Implementation Framework (QIF), we focus on elements related to delivery and support systems. These include establishing implementation teams, developing an implementation plan, dedicating resources towards the goals of the innovation, and ongoing technical support. While the QIF also includes process evaluation and a supportive feedback mechanism, our framework encompasses these aspects into a continuous improvement process, the final implementation support.

Establish Implementation Teams

School implementation teams tend to take one of two forms. Most commonly, teachers are grouped into grade or subject teams to receive training and ensure the curriculum sequence is aligned across school organizational units. A complementary approach is to utilize a school leadership team, often consisting of the principal, teacher leaders, and possibly even community members (Vernez et al., 2006). NCSU’s implementation teams are more closely related to this latter approach, although we place a large emphasis on teacher leadership, which has been linked to greater reform adoption with other school reform (Cohen et al., 2013; Datnow & Castellano, 2000). The principal is not a member of the team, but helps to recruit which teachers are on the team. The principal should also endorse the process and provide the necessary time, resources, and structural supports to facilitate scaling in. The implementation teams also differ from more traditional implementation models in that their work begins with the design work not

implementation. As we argued above, this initial involvement aims to help to ensure the innovation is aligned to the school context and addresses school needs, thereby engendering their buy-in and commitment and developing their capacity to implement the design. In this way, this framework reflects a design-based implementation approach, where participants focus on both the innovation design itself, and the process of implementing the innovation in particular contexts (Penuel, Fishman, Cheng, & Sabelli, 2011).

The school implementation teams include stakeholders who show promise in being able to train their peers how to implement the innovation and champion the work within the school by inspiring and leading others imbed its associated practices (Meyers, Durlak, & Wandersman, 2012). The implementation team uses their knowledge of the design to oversee implementation and translate their intimate knowledge of *why*, *what*, *when*, and *where* of the innovation design to the practices needed for school members to achieve self-efficacy, proficiency, and correct application of the innovation (Meyers, Katz, et al., 2012). By establishing an implementation team that has capacity and authority to lead the school's engagement with the innovation, the innovation is made more sustainable.

Develop Implementation Plan

Program implementation often relies on a plan for implementation that clearly specifies the timeline of activities, required tasks, and the resources and roles of the people to execute them (Meyers, Katz, et al., 2012). The main factor that distinguishes between plans is the level of specificity (Weiss et al., 2014). Peurach and Glazer (2012) emphasizes the importance of including codified routines, either more restrictive step-by-step procedures or open routines with general frameworks that determine the course of action and the frequency with which they are

enacted. Given the continuous improvement approach of NCSU, the implementation plans are revised over time as routines are tested and refined.

Meyers and colleagues (2012) also note that when developing the implementation plan, the implementation team should predict any challenges to effective implementation and document strategies to address them proactively. While we recognize the importance of anticipating challenges, conceptualizing implementation as scaling in to the school, there is a recognition that both the school and the innovation will evolve as the new practice is adaptively integrated into the school (Hannan, Russell, Takahashi, & Park, 2015). In this way, the innovation itself becomes more scalable as the design evolves to accommodate school conditions that are less than ideal (Clarke & Dede, 2009). The adaptive integration of the innovation into the school fosters the shift in ownership of the innovation from an external organization to the school itself.

Allocation of Resources

Resources refer to the time, materials, money, personnel, and expertise required to successfully carry out the implementation plan. In the context of the implementation plan, these resources are not viewed as part of the broader organizational capacity of a school—as the implementation literature seems to suggest (Meyers et al., 2012; Wandersman et al., 2012). Instead, resources are viewed in relation to their potential utilization for the productive enactment of the innovation. For instance, within a more conventional approach to implementation, the existence of experienced school staff who had previously implemented other reforms as an asset to successful implementation. In our framework, we would only view these teachers as an asset if they used their expertise and experience in the school to implement the innovation successfully. In other words, a school may have ample organizational capacity, but

for some reason does not make use of this capacity for scaling in the innovation. This becomes more apparent when considering financial resources; a school may have a general budget that would be sufficient for supporting the innovation, but only if these funds are made available for the innovation are they considered a resource.

Beyond their initial availability, resources often shift once implementation begins. Maintaining consistent resources has been a particular challenge in previous school improvement efforts. For instance, with the New American Schools, support was often only provided to the superintendent or an individual principal, rather than the central office or school staffs. Districts and schools often failed to dedicate organizational resources for schools (Berends et al., 2002). The allocation of resources is important as schools need to establish a sustainable system of supports for the innovation; the infrastructure that is enacted through dedicated time, organizational structures, and resource allocation will help to sustain the innovation as it is scaled in (Sabelli & Dede, 2013).

Ongoing Technical Support

Providing ongoing technical assistance is an essential support strategy for successful implementation (Desimone, 2002; Meyers, Katz, et al., 2012; Rowan, Correnti, Miller, & Camburn, 2009). Technical assistance helps practitioners handle the inevitable and unforeseen problems that develop once implementation has begun. Such supports may include further training, practice in administering more challenging components of the innovation, data analysis from practitioner feedback, or eliciting more support or resources for either the implementation team or other school stakeholders. By reconceptualizing implementation as scaling in to a school, ongoing technical support needs come in many forms, and the identification of such needs arise in multiple ways. One, with the recognition that implementation is as much about the

process of scaling in as the innovation itself, the teams need support for both the innovation itself and the process of engaging in continuous improvement. Two, as school teams engage in continuous improvement, themselves identify areas where technical support is needed. Three, research partners can provide another perspective on accomplishments and areas for further improvement through analysis of various forms of data. For example, in NCSU, researchers conduct fieldwork in participating schools and prepare reports back to those schools that summarize what was observed about how the innovation has been scaled in to their school.

Technical support is designed to build capacity among local leaders to ensure that when external partners withdraw technical support, these local experts can provide support for their peers in the district. Coburn (2003) describes this shift in reform ownership as a key element of scale up. When teachers, administrators, and other district staff no longer have to rely on external technical support, the innovation design, development, and implementation process can become self-sustaining.

Continuous Improvement Process

A critical feature of conceptualizing implementation as scaling in to a school is the way in which local context informs the practices that are embedded in schools. The initial design includes a combination of non-negotiable elements and those which are allowed to change, members of the school implementation teams and other early adopters engage in a continuous improvement process where the design is further developed as the team aligns the design to their unique school context. Yet, adaptation does not occur in isolation, but in the context of an improvement community where school teams bring evidence of what they've accomplished to share and discuss with others focused on the same problem of practice (Cannata, Cohen-Vogel, & Sorum, this issue; LeMahieu, Grunow, Baker, Nordstrum, & Gomez, 2017). In this way,

improvement communities bring discipline to the process of adaptive integration (LeMahieu et al., 2017).

NCSU frames the continuous improvement process around the "Plan, Do, Study, Act" (PDSA) cycle, where feedback is integral to short cycles of testing the effectiveness of targeted practices (Langley et al., 2009). PDSA is a model for organizational improvement that requires identifying the aim of a particular improvement, testing the change idea, and monitoring whether the observed changes led to the intended improvement. These brief and small-scale actions help inform decisions about the changes that merit being implemented at a larger scale. By starting small and allowing members of the implementation team to test discrete practices within the larger design, they are able to refine the innovation before involving a broader population of school stakeholders. PDSA can also be used to identify productive adaptations, previously unforeseen barriers to implementation, and to guide implementation in a growing number of organizational contexts, such as additional grade levels or classrooms. Ultimately, the improvement process builds an evidence base at each school that the district design team can use to inform plan to scale out to additional innovation schools.

Evaluating the quality of implementation

Reconceptualizing implementation as the process of scaling in to a school requires focusing on how innovations are co-constructed as local actors make sense of the innovation's demands and adapt them to their context (Datnow & Park, 2009). The continuous improvement process uses local knowledge and evidence base developed from testing specific practices as tools to increase the efficacy of the innovation design in each school. Yet, by undertaking a continuous approach to implementation, challenges arise in how best to determine the extent to which the innovation results in deep and consequential change in school and classroom practices

(Coburn, 2003). This framework focuses on four elements from the implementation literature. In their review of implementation effects across over 500 studies, Durlak and DuPre (2008) identify eight aspects of implementation: fidelity, frequency of exposure, quality of delivery, participant responsiveness, program differentiation (Dane and Schneider, 1998), program reach, adaptation, and monitoring comparison group conditions. Of these elements, we focus on the four that map onto two critical elements of Coburn's (2003) conceptualization of scale up: *depth* and *spread*. Three of these include frequency, participant responsiveness, and program reach. The fourth indicator in our framework, *integrity*, reframes fidelity to capture how innovations may be adapted to context. We focus on these elements to document the quality of initial implementation as well as track scale up. We focus first on the elements related to depth (*integrity* and *participant responsiveness*) before describing the elements related to spread (*frequency* and *program reach*).

Integrity

In Coburn's (2003) framework, depth is achieved only when an initiative or new program has altered teachers' beliefs about how students' learn, the norms of social interaction between teachers and students, and the underlying pedagogical principles that guide teachers' practice. Depth implies that the innovation design not only had superficial implementation or that it changed general school structures, but aligned teachers' beliefs and norms of practice with the innovation design. To achieve this alignment, the innovation must first be implemented with integrity. Integrity generally implies that the core components of the innovation were implemented as planned, reflecting traditional ideas of fidelity of implementation (Dane & Schneider, 1998; Durlak & PuPre, 2008). *As planned* tends to refer to specific program components, a curriculum, or scripted practices. In our framework, these plans include both the

correspondence with practices in the schools' implementation plans but also the core components that form the basis of the innovation.

We measure integrity at two levels, both in relationship to how school practices correspond with each school's implementation plan and the more generalized core components that informed the initial design. This contextualized conception of integrity is in sharp contrast to an approach to implementation that assumes that a program is implemented with integrity (or fidelity) only when it corresponds with a clearly planned treatment (e.g. Cordray & Pion, 2006; O'Donnell, 2008). Our definition of integrity implies that a school can adhere to the core components of the innovation design, even with adaptations that are different across schools, as long as those adaptations were described in their implementation plan and aligned with the goals of the design. For example, in NCSU's work in Fort Worth, Texas, a core innovation component includes the explicit teaching of the ideas of growth and fixed mindset to students. A common curriculum was developed, piloted, and revised by the DIDT. However, each school enacted this curriculum in different ways. One school had teachers throughout the school teach common lessons throughout the second day of school. Another school established an advisory period through which the curriculum was taught.

While the focus on strict fidelity may be applicable to some innovations, the complex process of integrating new practices into a school requires a more complex view of how a teacher's practices correspond with the initial design. We argue that circumventing the emphasis on strict fidelity allows for a broader focus on the quality of implementation. This shift in focus more closely resembles Meyers and colleagues' (2012) Quality Implementation Framework. They write: "Our focus is on quality implementation—which we define as putting an innovation into practice in such a way that it meets the necessary standards to achieve the innovation's

desired outcomes” (p. 465). This approach also reflects similar new approaches to scaling up, which suggest the goal of implementation is adaptive integration, or the “[integration of] a standard work process into new contexts” (Hannan et al., 2015, p. 496). Similarly, the scaling up literature emphasizes the need to attend to not only the “what” of the innovation practices, but the “why” and theory of action behind the practices, so that educators can guide adaptations in ways that retain integrity to the innovation’s core ideas (Bradach, 2003; Thompson & Wiliam, 2008). Evidence of scale up arises when teachers’ move beyond procedural elements of the design to having the underlying goals influence their beliefs about their teaching practice or the norms that guide how they interactions with their students and colleagues. This shift in not only practices, but beliefs about learning, is critical to achieving depth of change (Coburn, 2003; Cohen, 1990).

Participant Responsiveness

Participant responsiveness focuses on the extent to which school stakeholders are influenced by the innovation design (Dane & Schneider, 1998; Meyers et al., 2012). Responsiveness includes how school stakeholders feel and act in relation to the innovation. These responses include indicators such as levels of participation and enthusiasm. For school staff, this responsiveness includes how much they express support for the goals of the innovation design as well as the extent to which they are active participants in implementation. These measures of teacher behaviors are particularly important given the tendency for reforms to leave routine teacher practices unchanged. These measures would indicate how many teachers engage in specific practices over a pre-specified time period. Related to indicators of will described above, qualitative measures indicate the extent to which their practices are done out of compliance or commitment to the innovation’s goals. As implementation progresses and the

level of responsiveness to the innovation grows, we would have evidence that teachers' responsiveness increased their depth of practice.

A separate set of indicators is used to indicate how responsive students are to the innovation design. These include the level of engagement during delivery of the innovation practices and evidence that students have adopted behaviors outlined as goals of the innovation. Evidence of scale comes when responsiveness is translated into new patterns of interaction between students, teachers, and other school personnel in ways that support the proximal outcomes in the innovation's theory of action.

Frequency

Spread, in Coburn's (2003) scale up framework, is related not only to the expansion of an innovation to new classrooms or schools, but how teachers in those classrooms also experience the change in beliefs and practices described in regard to depth. Similar to Dane and Schneider (1998), the focus on frequency quantifies the level of students' exposure to the innovation. An example of a measure related to frequency would be the number of times a teacher is observed using one of the innovation's practices in a given class period. Since the implementation plans specify the expected frequency with which practices are enacted, evaluation of frequency examines the extent to which implementers are enacting the practices with the expected frequency delineated on the implementation plan, which also established another measure of integrity. Frequency leads to scale when, over time, stakeholders should increase the frequency of innovation practices as the norms of the innovation become embedded in their daily routines.

Program Reach

Given that non-implementation is the all-too frequent norm in educational implementation (Berman & McLaughlin, 1976), we also consider program reach. Durlak and

DuPre (2008) define program reach as the proportion and representativeness of program participants that implement an innovation. Examining the extent to which school personnel and students receiving the innovation are representative of the larger school is especially important during early phases of implementation, when only a subset of teachers may implement certain practices. It is likely that the early adopters may be more effective teachers and early evidence of effectiveness may not be attributable to the innovation but the teachers' pre-existing capacities. In measuring program reach, the representativeness of students also helps to understand how outcomes are due to changes in school practices or attributes of the exposed students due to their grade or ability level. The specific measures related to program reach allow for the documentation of how the number of students introduced to the innovation increase over time, indicating how the innovation spreads throughout a school. Program reach incorporates spread in the Coburn framework to indicate the innovation-related norms that spread within classrooms or school.

Conclusion

Given substantial challenges in traditional approaches to implement educational innovations at scale, new approaches to scaling up are gaining prominence (Bryk et al., 2015; Cohen-Vogel et al., 2015). Recent scholarship on these new approaches have focused on how they enact a research-practice partnership, establish communities of improvement across contexts, and place new demands on innovation developers (Coburn & Penuel, 2016; LeMahieu et al., 2017; Peurach, 2011). Less attention has been placed on understanding what it means to implement an innovation in a school under this new approach to scale. By reconceptualizing implementation as the process of scaling in to a school, we offer this framework as a way to evaluate the school-level enactment of an innovation as it is goes to scale. To achieve scale, new

approaches to scaling up involve local stakeholder not only in implementation, as is the case with traditional reform efforts, but in the design of an innovation and its adaptation to individual school contexts. Implementation is no longer a process distinct from achieving scale, but the two processes are interconnected as the innovation is scaled in to particular schools, scaled out throughout a district or other network, and scaled up by developing system-wide infrastructure to support continued engagement.

Further, conceptualizing implementation as scaling in, reframes the core elements of an implementation framework as evolving over time. While many of the components of the framework presented here reflect components of existing research on implementation, one goal was to illustrate how these components are not static, but can deepen and shift over time. For example, educator will and capacity have long been noted as important facilitating conditions to successful implementation (Berends, Bodilly, & Kirby, 2002; McLaughlin, 1990; Spillane, Reiser, & Reimer, 2002). The framework presented here reflects this research and further argues that these facilitating conditions may change over time; initial resistance to an innovation may shift to greater support as stakeholders experience early success. Likewise, initial enthusiasm may fade if the innovation does not meet expectations. This evolution over time can occur in all components of this framework, and the extent to which this evolution reflects a shift toward the ideas and practices of the innovation, this reflects more robust scaling in to the school. Achieving improvement at scale, then, requires attention to both the numbers of schools and districts impacted by an innovation (McDonald, Keesler, Kauffman, & Schneider, 2006), but also attention to the process of achieving depth of change and internal ownership within specific schools (Coburn, 2003).

References

- Becker, B. E., & Luthar, S. S. (2002). Social-emotional factors affecting achievement outcomes among disadvantaged students: Closing the achievement gap. *Educational Psychologist*, 37(4), 197–214.
- Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American Schools after a decade*. Santa Monica, CA: RAND. Retrieved from <http://www.rand.org/>
- Becker, B. E., & Luthar, S. S. (2002). Social-emotional factors affecting achievement outcomes among disadvantaged students: Closing the achievement gap. *Educational Psychologist*, 37(4), 197–214.
- Berends, M., Bodilly, S., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American Schools after a decade*. Santa Monica, CA: RAND. Retrieved from http://www.rand.org/pubs/research_briefs/RB8019/index1.html
- Berends, M., Kirby, S. N., Naftel, S., & McKelvey, C. (2001). *Implementation and Performance in New American Schools Three Years into Scale-Up*.
- Bradach, J. (2003). *Going to Scale: The Challenge of Replicating Social Programs* (Stanford Social Innovation Review). Stanford, CA: Stanford University. Retrieved from https://ssir.org/images/articles/2003SP_feature_bradach.pdf
- Bryk, A. S., Gomez, L., Grunow, A., & Hallinan, M. T. (2011). Getting Ideas into Action: Building Networked Improvement Communities in Education. In *Frontiers in Sociology of Education*. Springer Publishing.
- 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.

- Cannata, M., Cohen-Vogel, L., & Sorum, M. (this issue). Partnering for Improvement: Communities of Practice and Their Role in Scale-Up. *Peabody Journal of Education*.
- Chambers, D. A., Glasgow, R. E., & Stange, K. C. (2013). The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. *Implementation Science*, 8, 117. <https://doi.org/10.1186/1748-5908-8-117>
- Clarke, J., & Dede, C. (2009). Design for Scalability: A Case Study of the River City Curriculum. *Journal of Science Education and Technology*, 18(4), 353–365.
- Coburn, C. E. (2003). Rethinking Scale: Moving Beyond Numbers to Deep and Lasting Change. *Educational Researcher*, 32(6), 3–12. <https://doi.org/10.3102/0013189X032006003>
- Coburn, C. E., & Penuel, W. R. (2016). Research–Practice Partnerships in Education Outcomes, Dynamics, and Open Questions. *Educational Researcher*, 45(1), 48–54. <https://doi.org/10.3102/0013189X16631750>
- Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, 12(3), 327–345.
- Cohen, D. K., Peurach, D. J., Glazer, J. L., Gates, K. E., & Goldin, S. (2013). *Improvement by Design: The Promise of Better Schools*. Chicago ; London: University Of Chicago Press.
- Cohen-Vogel, L., Cannata, M., Rutledge, S., & Socol, A. R. (2016). A Model of Continuous Improvement in High Schools: A Process for Research, Innovation Design, Implementation, and Scale. *Teachers College Record*, 116(13), 1–26.
- 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*, 0895904814560886.

- Datnow, A., & Park, V. (2009). Conceptualizing policy implementation: Large-scale reform in an era of complexity. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook of Education Policy Research* (1 edition, pp. 348–361). New York : Washington, D.C.: Routledge.
- Davison, M. L., Young, S. S., Davenport, E. C., Butterbaugh, D., & Davison, L. J. (2004). When do children fall behind? What can be done? *Phi Delta Kappan*, 85(10), 752–761.
- Dede, C. (2006). Scaling up: Evolving innovations beyond ideal settings to challenging contexts of practice. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 551–566). Cambridge, UK: Cambridge University Press.
- Desimone, L. (2002). How Can Comprehensive School Reform Models Be Successfully Implemented? *Review of Educational Research*, 72(3), 433–479.
<https://doi.org/10.3102/00346543072003433>
- Durlak, J. A., & Dupre, E. P. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *Am J Community Psychol*, 41, 327–350.
- Elmore, R. F. (2016). “Getting to scale...” it seemed like a good idea at the time. *Journal of Educational Change*, 17(4), 529–537. <https://doi.org/10.1007/s10833-016-9290-8>
- Fishman, B. J., Penuel, W. R., Allen, A.-R., & Cheng, B. H. (2013). *Design-based implementation research: theories, methods, and exemplars*. New York: National Society for the Study of Education.
- Fullan, M. (2016). The elusive nature of whole system improvement in education. *Journal of Educational Change*, 17(4), 539–544. <https://doi.org/10.1007/s10833-016-9289-1>

- Gallagher, M. J., Malloy, J., & Ryerson, R. (2016). Achieving excellence: Bringing effective literacy pedagogy to scale in Ontario's publicly-funded education system. *Journal of Educational Change*, 17(4), 477–504. <https://doi.org/10.1007/s10833-016-9284-6>
- 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. <https://doi.org/10.1177/0022487115602126>
- Hartmann, A., & Linn, J. (2008). *Scaling Up: A Framework and Lessons for Development Effectiveness From Literature and Practice* (No. Working paper 5). Washington, D.C.: Wolfensohn Center for Development at Brookings. Retrieved from https://www.brookings.edu/wp-content/uploads/2016/06/10_scaling_up_aid_linn.pdf
- Hopkins, M., & Woulfin, S. L. (2015). School system (re)design: Developing educational infrastructures to support school leadership and teaching practice. *Journal of Educational Change*, 16(4), 371–377. <https://doi.org/10.1007/s10833-015-9260-6>
- Lee, J. (2004). Multiple facets of inequity in racial and ethnic achievement gaps. *Peabody Journal of Education*, 79(2), 51–73.
- LeMahieu, P. G., Grunow, A., Baker, L., Nordstrum, L. E., & Gomez, L. M. (2017). Networked Improvement Communities: the discipline of improvement science meets the power of networks. *Quality Assurance in Education*. <https://doi.org/10.1108/QAE-12-2016-0084>
- Mazzeo, C., Fleischman, S., Heppen, J., & Jahangir, T. (2016). Improving High School Success: Searching for Evidence of Promise. *Teachers College Record*, 116(13), 1–32.
- McDonald, S. K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, 35(3), 15–24.

- McLaughlin, M. W. (1990). The Rand Change Agent Study Revisited: Macro Perspectives and Micro Realities. *Educational Researcher*, 19, 11–16.
- Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology*, 50(3–4), 462–480.
- Meyers, D. C., Katz, J., Chien, V., Wandersman, A., Scaccia, J. P., & Wright, A. (2012). Practical Implementation Science: Developing and Piloting the Quality Implementation Tool. *American Journal of Community Psychology*, 50(3–4), 481–496.
<https://doi.org/10.1007/s10464-012-9521-y>
- Newman, D., Zacamy, J., Lazarev, V., & Lin, L. (this issue). School Processes that Can Drive Scaling-up of an Innovation or Contribute to Its Abandonment. *Peabody Journal of Education*.
- 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. <https://doi.org/10.3102/0013189X11421826>
- Peurach, D. J. (2011). *Seeing Complexity in Public Education: Problems, Possibilities, and Success for All* (1st ed.). Oxford University Press, USA.
- Peurach, D. J., Lenhoff, S. W., & Glazer, J. L. (2016). Large-Scale High School Reform through School Improvement Networks: Exploring Possibilities for “Developmental Evaluation.” *Teachers College Record*, 118(13), 1–28.
- Peurach, D. J., & Neumerski, C. M. (2015). Mixing metaphors: Building infrastructure for large scale school turnaround. *Journal of Educational Change*, 16(4), 379–420.
<https://doi.org/10.1007/s10833-015-9259-z>

- Sabelli, N., & Dede, C. (2013). Empowering design-based implementation research: The need for infrastructure. In B. J. Fishman, W. R. Penuel, A.-R. Allen, & B. H. Cheng (Eds.), *Design-based implementation research: theories, methods, and exemplars* (pp. 464–480). New York: Teachers College, Columbia University.
- 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. <https://doi.org/10.3102/00346543072003387>
- Thompson, M., & Wiliam, D. (2008). Tight but loose: A conceptual framework for scaling up reforms. In E. C. Wylie (Ed.), *Tight but loose: Scaling up teacher professional development in diverse contexts* (pp. 1–44). Princeton, NJ: ETS.
- World Bank,. (2003). *Scaling-Up the Impact of Good Practices in Rural Development: A working paper to support implementation of the World Bank's Rural Development Strategy* (No. 26031). Washington, D.C.: The International Bank for Reconstruction and Development. Agriculture and Rural Development Department. Retrieved from http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2004/01/30/000160016_20040130163125/Rendered/PDF/260310White0co1e1up1final1formatted.pdf
- Murphy, J. & Torre, D. (2014). *Creating productive cultures in schools: For students, teachers, and Parents*. Thousand Oakes, CA: Corwin Press.
- Newmann, F. M., Smith, B., Allensworth, E., & Bryk, A. S. (2001). Instructional program coherence: What it is and why it should guide school improvement policy. *Educational Evaluation and Policy Analysis*, 23(4), 297–321.
- Nunnery, J. A. (1998). Reform ideology and the locus of development problem in educational restructuring enduring lessons from studies of educational innovation. *Education and Urban Society*, 30(3), 277–295.

- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, 78(1), 33–84.
- 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.
- Peurach, D. J., & Glazer, J. L. (2012). Reconsidering replication: New perspectives on large-scale school improvement. *Journal of Educational Change*, 13(2), 155–190.
- Rowan, B.P., Correnti, R.J., Miller, R.J., & Camburn, E.M. (2009) School improvement by design: Lessons from the study of comprehensive school reform Programs. In G. Sykes, B. Schneider, & D.N. Plank (Eds.), *Handbook of Education Policy Research* (pp. 637-651). New York, NY: Routledge.
- Spillane, J.P. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition and Instruction*, 18(2), 141–179.
- Spillane, J.P., & Callahan, K.A. (2000). Implementing state standards for science education: What district policy makers make of the hoopla. *Journal of Research in Science Teaching*, 37(5), 401–425.
- Spillane, J.P., & Jennings, N.E. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. *Teachers College Record*, 98(3), 439–481.
- Spillane, J.P., Reiser, B.J., & Gomez, L.M. (2006). Policy implementation and cognition: The role of human, social, and distributed cognition in framing policy implementation in Meredith I. Honig (ed.) *New Directions in Education Policy Implementation: Confronting Complexity*. Albany, NY: State University of New York Press.
- Stringfield, S., & Datnow, A. (1998). Scaling up school restructuring designs in urban schools. *Education and Urban Society*, 30(3), 269–276.
- Supovitz, J. A., & Weinbaum, E. H. (2008). *Implementation gap: Understanding reform in high schools*. New York, NY: Teachers College Press.
- Vernez, G., Karam, R., Mariano, L. & DeMartini, C. (2006) *Evaluating comprehensive school reform models at scale: A focus on implementation*. Santa Monica, CA: RAND.
- Wandersman, A., Duffy, J., Flaspohler, P., Noonan, R., Lubell, K., Stillman, L., Blachman, M., Dunville, R., & Saul, J. (2008). Bridging the gap between prevention research and practice: The interactive systems framework for dissemination and implementation. *American Journal of Community Psychology*, 41(3-4), 171–181.

Weiss, M. J., Bloom, H. S., & Brock, T. (2014). A Conceptual Framework for Studying the Sources of Variation in Program Effects. *Journal of Policy Analysis and Management*, 33(3), 778–808.

Figure 1. Implementation for Scale Up Framework: Logic Model

