Routledge
Taylor & Francis Group

Check for updates

Getting the Maximum From the Minimum: an Adaptive System for Scaling School Reform

Nell Scharff Panero

Hunter College, City University of New York, New York, NY, USA

ABSTRACT

This article presents findings from a study of the minimum conditions needed to scale a previously successful model of high school reform in New York City. Most studies of scale work backward to identify conditions that were in place when scale was successful, rather than identifying those that were minimally required, and thus more likely to be achievable across settings. Findings from this study suggest that only three minimum conditions are needed for scale – adequate knowledge of the model; an appropriate environment; and a working feedback loop – when these conditions interact in a functioning, self-adjusting ecosystem.

Improving high schools entails challenges that are "far easier to catalogue than to surmount" (Mazzeo, Fleishman, Heppen, & Jahangir, 2016, p. 2). High schools have been the most difficult to reform because the challenges teachers face are arguably the hardest: skill gaps present since elementary school have widened; the demand to graduate all students college and career ready has grown; and the structural and cultural constraints to change – siloed departments and norms of privacy, teacher autonomy, and the view of the teacher as content expert rather than learner, all of which impede continuous improvement – are most intense on the highest rungs of the K-12 ladder as well (McLaughlin & Talbert, 2001; Seashore Louis & Lee, 2016; Siskin & Little, 1995). Where individual high schools or pockets of schools have beat the odds, scaling up success poses formidable challenges (Coburn, Russell, Kaufman, & Stein, 2012; Elmore, 1996; Fullan, 2000; Stringfield & Datnow, 1998). Scaling up in the research has taken on multiple meanings. It includes the most traditional meaning of spreading a reform outward to a larger and larger number of schools as well as the critical notion of depth – that the reform is embedded in the culture to the extent that it infiltrates the "grammar" of a building or district (how things are done around here), thus allowing the reform to become self-generative and more likely to be sustained (Coburn, 2003; McLaughlin & Mitra, 2001).

The particular reform studied here is called strategic inquiry (SI), which was found to be successful in turning around struggling high schools in New York City (Panero & Talbert, 2013; Wohlstetter, Kim, & Flack, 2018). Its success was particularly striking in large high schools, typically most resistant to reform, where schools that committed to the model for 3 or more years brought incoming struggling students on track to college and graduation so significantly that being a student in one of these schools "increased the likelihood of being on track for graduation and college readiness approximately as much as having [entered with] relatively high 8th grade reading skills" (Talbert, Cor, Chen, Kless, & McLaughlin, 2012, p. 14).

This study focuses in on a key challenge to scale, as defined by spread, depth, and sustainability – that of being able to deliver a reform with fewer resources than originally available, in this case, via a train-the-facilitator approach (see Stringfield, 2002, positing that scale, in essence, is about making what "once was

---

available only to the rich … available to the broad expanse of a culture" [p. 63] with fewer resources). As is often the case in early stages of development and implementation of a new reform, the first iteration of SI was heavily resourced by foundation funding. This allowed a number of ideal conditions to be in place, including unusual amounts of time for program facilitators to be trained (one full day per week) and to be on site (one dedicated program facilitator at each large school 3 full days per week).

In the model studied here, the experiment was one of scale in that it involved a partnership with a district to implement a train-the-facilitator approach aimed to improve a larger number of schools with fewer resources. The SI facilitators (now called "consultants," to distinguish them from the district- and school-based facilitators they would be developing) worked at a higher level of remove from the schools they were aiming to improve. In agreeing upon the model's design in conversation with the district, SI leaders drew upon what prior research had identified as core design elements (Panero & Talbert, 2013) and their experience to articulate non-negotiables, what they believed must be in place in order for the scaled model to work. Non-negotiables included, for instance, a minimum number of days for SI consultants to deliver training to district coaches and a specific design for coach allocations to the schools.

As is probably typical in large-scale, high-stakes implementation of district reform, many of these non-negotiables, despite being agreed to up front, were not in fact met: time for training was reduced; coaches were not allocated as intended; and most notably, perhaps, what was designed as a three-year initiative was cut short at the end of two. As a result, there were many schools where the program did not take hold as intended. Despite these limitations, however, an independent evaluation found that the program worked. Specifically, students in SI schools were found to be almost two and a half times more likely to be on track to graduate than students in non-SI schools. The researchers found, in essence, that despite limitations of implementation, the model was effective when implemented at scale (Wohlstetter et al., 2018).

This led the researcher to wonder which of the conditions previously deemed essential were, in fact, necessary rather than ideal. In the context of scale, where there is always the need to accomplish more with less and where political churn and shifting priorities are less the exception than the rule, it is critical to understand the minimum conditions that account for success. This study seeks to do just that and in doing so to reveal something new about what it truly takes to improve high schools at scale within real-world contexts.

Literature review: what do we know about the minimum conditions needed to improve high schools at scale?

There are two specific limitations in existing research that this study seeks to address. First, there is almost nothing in the literature that addresses minimum conditions" head-on. Most studies of scale take successful cases as a starting point and extrapolate backward. For example, the 2002 MDRC study of districts that "beat the odds" identified a sampling of urban systems that improved student achievement and the commonalities in organizational capacity and approach that likely enabled change. Commonalities included political stability; a unifying vision focused on improvement for the most struggling learners; sustained resources aligned to the vision for improvement; etc. (Snipes, Doolittle, & Herlihy, 2002). The study's findings make sense; they reflect conditions to which most urban districts aspire and that likely facilitated these districts' success. They do not reveal, however, which conditions were minimally required in that success could not have happened without them. This seems critical knowledge since few districts come to the reform table with all or even many of these conditions met.

Second, because scaled success is rare at the high school level (Fullan, Bertani, & Quinn, 2004), lessons about scale are gleaned almost entirely from studies that do not include them. The 2002 MDRC study described above, for example, qualifies findings by stating that evidence of success in the featured districts was limited to elementary and middle schools (see McLaughlin & Mitra, 2001 & Datnow, 2002 for additional examples of theory-building about scale drawn from an evidence base

not inclusive of high schools). Researchers hope that lessons learned will apply broadly but acknowledge that generalization is difficult, given the organizational and cultural complexities of high school change (Firestone & Herriott, 1982; McLaughlin & Talbert, 2001). That said, a broad, evidence base, drawn mostly from studies of elementary and middle schools, points to conditions at the school, district, and program levels generally needed to support improvement at scale.

There are a number of school-based factors that, according to research, help embed a reform deeply and sustainably in a building. Datnow (2002) summarizes these findings as follows: schools need "a genuine interest in change; supportive teachers and administrators; a critical mass of teachers involved in implementation; sustained professional development; and a practical plan for implementation and monitoring of the change effort" (p. 225) along with stable staff, stable leadership, and a consistent reform focus over time. In addition, research points to the critical importance of the principal's role in buying-in, supporting, shepherding and distributing leadership of the change process (as summarized in Fullan, 2002), as well as of high-quality social networks/structured professional collaborations (Coburn et al., 2012; Gallimore, Ermeling, Saunders, & Goldenberg, 2009; Panero & Talbert, 2013) in shifting teachers' thinking and practice in ways that sustain.

The research focused on scripted or pre-packaged reforms (curriculum and/or material adoption) has tended to emphasize up-front buy-in, existing expertise, and fidelity of implementation as critical school-level factors. Studies of theory- or concept-driven models, on the other hand, emphasize the need for school-level actors' deep learning of a program's driving principles through supported engagement with those principles over time. Deep knowledge of "first principles," rather than adherence to a script, allows flexible adaptation and application to shifting local contexts (McLaughlin & Mitra, 2001; Peurach & Glazer, 2012) and is a prerequisite for the shifting of reform ownership needed for efforts to sustain (Coburn, 2003).

Districts must hold a clear vision for improvement grounded in a belief that all children can learn and marshal all available resources to get it done (Chenowith, 2017). They must select and commit to a reform strategy aligned to their vision for improvement, refraining from promoting contradictory theories of change simultaneously (Talbert, 2011); commit sufficient, aligned resources; and stay the course (Stringfield & Datnow, 1998). Districts also need "absorptive capacity," "the ability to recognize the value of new information, assimilate it, and apply it in novel ways as part of organizational routines, policies and practices" (Cohen & Levinthal, 1990; Zahra & George, 2002, as cited in Farrell & Coburn, 2017). This is challenging not only because districts typically underestimate the amount of resources they will need (financial, human, time for partnering, etc.) and pull them too soon (Farrell & Coburn, 2017; Stringfield, 2002) but also because of the tendency of district personnel, as with all learners, to assimilate new information into preexisting frames, failing to accurately see that which is new as new (Spillane, 2000).

Programs must do more than have a strong model that has worked in a finite number of settings. To succeed at scale, programs must embed a plan for reciprocal learning, mutual adaptation, and the eventual shifting of reform ownership within the original design (Coburn, 2003; Farrell & Coburn, 2017; Peurach & Glazer, 2012). They must be designed, in other words, with an understanding of complexity theory, knowledge that "complexity, dynamism and unpredictability are not merely things that get in the way. They are normal!" (Fullan, 1993, p. 20). In addition, they need an infrastructure matched to the scope of scale since "growing from 1 school to 10, or from 10 schools to 100, requires a core group of highly committed generalists … [whereas] to successfully grow from 100 to 1000 requires the building of a highly professional, specialized infrastructure" (Stringfield, 2002, p. 68). Programs must strike a difficult balance, offering something of value without appearing too different, since reforms have succeeded at scale largely because of their consistency with, rather than challenge to, existing norms (Cohen & Mehta, 2017). The issue in part may be one of messaging, how to signal alignment while enabling transformation (Spillane, 2000). Finally, scale requires an expanded number of expert facilitators of specified models for change. Research to date says little about what it takes to develop such facilitators (Talbert et al., 2012, p. 33).

Context for the study: strategic inquiry for high school renewal in New York City

The context for this study was a high-profile effort to "renew" rather than close 35 of New York City's most struggling high schools between November 2014 and November 2016.[1] Due to its strong track record turning around high schools in New York City, SI's strategy for collaborative inquiry-based improvement was selected for the renewal high schools along with a focus on closing literacy-based skill gaps that had surfaced in prior iterations of SI but not yet been fully tested (see Tyre, 2012 & Panero & Talbert, 2013 & Panero, 2016 for detailed descriptions of the literacy gaps surfaced through SI and the approach to closing them that was developed in response). The combination, it was hoped, would accelerate improvement in the struggling schools.

Because of the large number of schools involved and the limited resources available to support them (no foundation funding, a small number of SI-trained consultants), a train-the-facilitator model was selected as the best approach. The model, in a nutshell, was to develop "facilitator pairs," with each pair comprised of a district coach and a school-based teacher leader, to lead and spread leadership of SI in each school. A team of SI consultants would train cohorts of facilitator pairs at monthly full-day training sessions for three years. The expected outcomes for the pairs' learning and the spread of improvement at each school were as follows:

- In year 1, pairs would demonstrate their ability to lead one SI team to improve outcomes for targeted 9th grade struggling students. The teachers on this team would demonstrate shifts in thinking and practice aligned with SI's theory of change, along with improved outcomes for targeted students.
- In year 2, pairs would demonstrate their ability to train other teacher leaders as "facilitators" capable of producing similar outcomes for an expanded number of SI teams (a critical mass of 9th grade teams and at least one 10th grade team). These teams would produce similar shifts in thinking and practice for their teachers and improved outcomes for an expanded number of students (the entire 9th grade and targeted 10th grade struggling students).
- In year 3, pairs would continue to strategically extend SI's reach (via support for leaders of teams now involving a critical mass of 10th grade teachers and at least one 11th grade team and demonstrating outcomes for 9th and 10th grade students as a whole and targeted 11th grade struggling students) and by establishing systems and structures to sustain continuous improvement.

SI's strategy for high school renewal was ambitious; it aimed, in effect, to secure the same outcomes as in the prior iteration but from a much greater level of remove. Unfortunately, at the start of year 3 (in November 2016), for political reasons counter to the wishes of the Directors of the Renewal program, the partnership with SI was suddenly stopped. A second cohort of 19 high schools had entered the Renewal Program in year 2, having worked with SI for only one year when the partnership ended. In these schools, the work was not sufficiently embedded to sustain without the partner's support. Among the first cohort of schools, however, were a number in which SI had become sufficiently "structured into the daily lives of the school community" (Datnow, 2002, p. 24) that it was likely to sustain (see Schwartz, 2017 for a description of the work at one successful renewal high school in what would have been year 3). Success across the initiative and sustainability at some schools in spite of less than ideal conditions – even with the sudden and unexplained withdrawal of support – raised the question of minimum conditions upon which this study is based.

Method

The participants in this study were drawn from the group who knew the work the best, the SI consultants. In the renewal project, the SI consultants trained facilitator pairs as a cohort; coached pairs and other leaders at school sites; led intervisitations across sites; and met regularly as a team to

identify evidence-based patterns about what facilitator pairs needed next and to design training. In other words, they had an up-close and a birds' eye view of the work along with deep knowledge of its theoretical underpinnings. Of the eight consultants working with the renewal project, six became participants in this study. Four had worked with SI and received extensive training previously, one full day per week for three years; one had been trained in and supported the model in another city; and one was new to SI. All were highly experienced educators and professional developers, three with a Ph.D. in education and the other three having served, among other roles, as school principals.

The data were collected via a focus group interview (Dilshad & Latif, 2013). It was typical for the consultants to debrief at the end of each year. In June 2017 (at the end of what would have been the third year if the work had not been canceled), the consultants gathered to reflect upon the program overall with a focus on what had been a persistent question at a number of their meetings, the minimum conditions needed to enable success.

In order to get at this question, a focus group interview protocol was designed in which the consultants were asked first to come to an agreement about what constituted success (namely, schools meeting the goals articulated in the prior section for each year). Then, each consultant was asked to discuss specific schools in which the work was NOT successful and what at a minimum would have allowed for success at those sites. Open conversation was encouraged, to foster as much collective thinking as possible on the topic. Consultants were reminded to be specific and – when they began to talk about ideal conditions, those they wished had been in place – to answer the question of what *minimally*, in their opinion, would have allowed for success.

The conversation was audiotaped and later transcribed and analyzed using a grounded theory approach (Charmaz, 2000; Strauss & Corbin, 1998). Participants' statements that addressed the question of minimum conditions were coded and then sorted into "buckets" (Ely, Anzul, Friedman, Garner, & Steinmetz, 1991) or categories that emerged from the data as follows: minimum conditions relating to the *district*, the *trainees*, the *schools*, and the *program*. Then, codes were further sorted and placed into a smaller number of categories (larger buckets) according to their function, which resulted in three overarching categories as follows: *knowledge of the model; appropriate environment; and feedback loop*. Where statements were contradictory or inconsistent, the researcher sought to identify what made them both true, to lift to a level of abstraction where they were consistent in support of a particular condition. Codes appearing prominently across categories pointed to emergent themes, such as *time* and *prioritizing the work*. The fact that almost all of the codes fit to some extent in all three overarching categories pointed to the interactions of the conditions as an important finding.

Findings: very few minimum conditions operating in a self-adjusting ecosystem

The primary overarching finding is that because the studied program is designed to adapt to shifting contexts, there are very few minimum conditions required to scale it, far fewer than typically assumed. In short, *adequate knowledge of the model, an appropriate environment*, and *a working feedback loop* are the minimally required necessary conditions. These three overarching conditions form an interconnected ecosystem that "self-adjusts" to maximize efficiency and effectiveness of the model in the face of unexpected, new or ongoing challenges. Working in concert, they minimize the number of conditions needed for success. Knowledge of the model, for example, allows for some latitude in the conditions of the environment and minimizes some of the need for feedback. Appropriate environment, in turn, accelerates understanding and adoption of the model and maintains productive feedback. A working feedback loop identifies and provides components of missing knowledge, as well as identifying and suggesting adjustments to the environment. Together, these conditions create an adaptive system that allows the key principles of the model to be preserved in any number of conditions, including those that are less than ideal.

The sections below define each of the necessary conditions; illustrate how each operates within the ecosystem; and surface what is different from what program designers originally thought would be required and/or what is typically assumed to be required for success at scale.

Adequate knowledge of the model

The most critical knowledge that facilitators must gain is the knowledge that the model is grounded in principles rather than rules. They must learn to apply these principles flexibly in varied contexts, and to do so knowing that it is the principles themselves, rather than any rule book, that guides. Adherence to rules, rather than principles, is a typical characteristic of the novice. Fully "cooked" facilitators, as SI consultants sometimes refer to those deemed ready to lead SI independently, have demonstrated their ability to NOT be rule-bound, but rather to hold to underlying goals and purposes in the face of varied challenges and to draw upon SI principles flexibly and creatively to get there.

Learning core principles and to trust them as the drivers of the action take time; it involves cycles of implementation through which one comes to understand how the principles operate in a self-repairing system, meaning that the principles themselves dictate next steps. At first, a trainee implements according to their best understanding of the principles at that time. If goals are met, they keep going. Meeting a challenge, however, brings them up against the limits of current knowledge and to conversation with a more expert colleague or trainer who provides new ideas or insights to be tested in practice. When progress stalls – with no evidence of movement to desired outcome goals – the facilitator has likely been adhering to rules rather than to principles. Acting in closer accord with principles becomes a test of the model itself; when doing so moves the work forward, facilitators come to understand and to believe in the principles as guides for further action.

For example, in SI, there is a tool called a "Tennis Chart" (Panero & Talbert, 2013, pp. 159–160) that a facilitator learns to adapt for maximum impact. Most importantly, the facilitator must come to understand that the tool *must* be adapted. The purpose of the chart is to help teams understand what, in specific cases, students need to learn next to get unstuck. It helps teams articulate and test a working hypothesis, that if students knew a, b, and c, they could master d (as in, if a tennis player mastered ball toss, foot position, and swing, she would be able to "serve"). It aims to develop as well a number of skills and habits of mind for teachers working together on a team, including shared accountability and the capacity to examine and adjust practice as dictated by evidence.

The tennis chart below (see Figure 1) was created by SI consultants for use in a renewal training. Using an actual student's work (Kareem's, a pseudonym) collected by one facilitator pair, the consultants developed an activity to model use of such a chart in support of the desired learning goals. SI consultants hoped the pairs would understand how such a chart can work and then create similar but modified charts for use in their own contexts.

This particular team was comprised of 9th grade cross-content teachers that had been led to select a common skill focus (writing) and had zoomed in to identify shared students' struggle with the conjunctions *because, but*, and *so*. Each teacher had agreed to teach these conjunctions embedded in their content and to collectively evaluate the results.

In reviewing Kareem's work to prepare the activity, the SI consultants saw that Kareem appeared able to use the conjunctions effectively in some content areas but not others. They also saw that the incorrect usage occurred in response to confusing teacher tasks and designed the chart with the specific intent of surfacing this discrepancy for teachers. The chart was designed strategically, in other words, not only to surface a next step based on evidence, but also to push new thinking and practice, in this case with regard to a shared skill gap that was also intended to forge a cross-content connection.

At the subsequent training, however, Ricardo (a facilitator in training, a pseudonym) shared with the group that his own work implementing the tennis chart had failed. When asked to elaborate, he explained that his team was a group of social studies teachers who felt strongly that students were

Tennis Chart for Kareem - Does he get the skill (*Because, but, so*)?

	Because			But			So		
	Skill	Content	Mechanics	Skill	Content	Mechanics	Skill	Content	Mechanics
In Social Studies									
In English									
In Living Environment									
In Algebra									
Total count (# of Y/N out of total):	✕			✕			✕		

Figure 1. Tennis chart for Kareem.

Tennis chart for Kareem – Does he get the skill (*Because, but, so*)?

struggling most with content and were irritated that the content box on the bottom row of the tennis chart was crossed out. "Why can't we aggregate the missing content," they had asked, to which Ricardo had replied: "Because that's just not what we do in SI; we focus on skills, not content."

Based on this experience and others like it, since this was one of the many times when facilitator pairs followed the letter rather than the spirit of the law in designing tennis charts, the SI consultants created an activity focused on strategic design. While SI consultants thought they had been clear up front that Kareem's chart was an example, designed for a specific purpose and context, that message had to be nailed home and learned more deeply through experience. For the second activity, consultants wrote scenarios to capture varied inquiry team contexts (different team composition and goals for adult learning) all charged with reviewing the same student work. The pairs generated multiple tennis charts each to suit a specific context and analyzed similarities, differences, and the reasons behind each design.

In this setting, the feedback provided by the participant (that he did what he thought he was supposed to do and it did not work) provided information to the SI consultants that allowed them to understand the particular aspects of missing knowledge that most needed to be addressed. When Ricardo designed a new, more context-specific tennis chart and succeeded this time in moving the adults on his team toward his goal, the loop was closed.

Appropriate environment

An appropriate environment is one in which adequate knowledge of the model can be gained. This environment allows cycles of implementation to occur such that the process described above of testing and coming to understand how principles operate and drive the work can be learned. It is an environment that allows for experimentation over time and meaningful interaction with more knowledgeable others (those with expertise in a given model) at critical junctures when progress is stalled to infuse new ideas to be tested in practice. It also involves whatever structures, attitudes, and habits of mind are needed to nurture the process of learning core principles and their role through experience. However, not as many of these conditions are necessary up front as is typically assumed, because where ideal structures, habits of mind, and knowledge are not in place, existing conditions interact to create those that are minimally required for an appropriate environment.

In the studied program, for example, certain structural conditions were minimally required for the work to take hold. Namely, there needed to be regular time for inquiry to take place at school

sites. Principals needed to protect this time or at least to refrain from usurping it on a regular basis. What is typically assumed is that principals need to buy into and understand the work up front, in part because they are then more likely to protect time for it. In the studied model, some principals understood, believed in, and protected inquiry from the start. Where this was not the case, however, the program drew on whatever leverage it had in specific settings to ensure that the work would proceed anyway, with the assumption that principal commitment would develop over time.

The first cohort of renewal schools, for example, had the benefit of district coaches allocated to each school. Where mandated structures (including at least 90 minutes per week for inquiry teams to meet) were not initially put into place or protected, coaches used their formal authority as arms of the Executive Superintendent of Renewal to enforce the mandate. In many cases, what was began as a mandate, against the wishes of principals who wished to use team time differently, became a genuine principal priority as evidence of shifting team culture and improved student outcomes began to emerge.

The second cohort of 19 schools, however, had neither the benefit of renewal coaches nor the clear lines of formal authority that operated in cohort 1. Schools in the second cohort reported to varying superintendents who did not themselves report to the Executive Superintendent of Renewal. Ensuring time and basic structures for inquiry – even the consistent sending of teachers to SI training – was more challenging in this environment and required use of the carrot rather than the stick.

Because there was an open conversation (a feedback loop) between SI consultants and district leaders, the partners collaborated to develop and run a new training aimed specifically at winning over the Directors of School Renewal (DSRs). The DSRs functioned in a role similar to the coaches in cohort 1; however, they reported to local superintendents. They were not directly accountable to anyone charged with running the renewal initiative, but they served a critical role in supporting or hindering its success.

DSRs were invited to a series of specially tailored sessions in which they were led through a modified inquiry process focused on the teachers and schools they were charged with supporting. The hope was to engage them through work that was immediate and real while developing a deep understanding and thus their capacity for leading (and advocating for) the SI work. This investment (a reallocation of resources and a shift in strategy to focus on the DSRs) was effective in a number of cases. Some DSRs became local champions of the SI process and new arms (through carrots rather than sticks) of the renewal initiative.

A working feedback loop

A working feedback loop is an open line of back and forth communication between various elements of a program, most importantly between program leaders (in SI, the consultants) and those implementing it on the ground (the facilitator pairs). The feedback loop must provide information about the extent to which other conditions are in place (adequate knowledge and appropriate environment) such that elements can be adjusted. It must also provide evidence of progress in a program's ultimate goals. In SI, for instance, program designers needed continuous feedback about the state of knowledge and the environment as well as about movement in the student outcomes these conditions were intended to effect. The feedback loop's reaching all the way to the end goal (improved student performance) ensures a check on the program design itself so that it too can be adjusted.

As with the other conditions, a working feedback loop allows some latitude for missing knowledge or environment by providing a mechanism for compensation. In the studied program, for instance, a number of conditions were not in place as expected, but the program was able to respond. For example, for most of the first year, planned on-site coaching by consultants was canceled due to lack of funding. Consultants obtained needed information, however, by collecting and analyzing verbatim "low-inference transcripts" of inquiry meetings (see Panero & Talbert, 2013), which provided granular,

"low-inference" and thus reliable evidence of facilitators' leadership skills and team challenges, as well as via analysis of teacher tasks and target students' work that pairs were regularly required to submit. These products compensated for consultants' inability to see the work first-hand.

Similarly, from the start of year 1, district coaches were not allocated as intended, resulting in a series of conversations designed to shift district leaders' minds and ultimately to a series of adjustments. The decision to train pairs rather than individuals was the result of a joint desire to develop district buy-in and capacity. In addition, the SI consultants knew that trained district personnel provided important redundancy – that their learning the model deeply in year 1 meant they would be available to help spread it in subsequent years. Trained district personnel would be especially critical in schools where trained site-based facilitators might have left. To learn the model well enough to function in this way, coaches needed to be present for (to co-lead and learn from doing so) all inquiry team meetings in at least one school for the first year.

While district leaders agreed to allocating coaches accordingly and likely intended to do so, in fact, coaches had been hired according to a fundamentally different theory of change. The predominant district theory was that content-based improvement (content coaching) would most effectively drive change. Coaches were hired for their content expertise and, in spite of up-front verbal and written agreements to the contrary, were assigned to a different school every day to maximize the number of teachers in their content area that they could coach. When push came to shove, the underlying belief systems held by those who hired and allocated the coaches drove placement.

Research points to conflicting theories of change as a factor that undermines scale. A more nuanced understanding of this idea, however, suggests that contradictory theories of change should be expected and that programs must be designed to bring them into alignment. In the studied program, consultants brought their concerns to renewal directors at different times and in different ways over the course of the program, trying to discuss the benefits and limitations of the existing design. Two months into the first year, one of the three supervisors shifted her allocation of coaches to align with SI; the other two did not. SI consultants tracked carefully which coaches now had the opportunity to learn SI deeply and which did not.

By the start of the second year, renewal directors had come to understand why coaches would have trouble supporting SI without having implemented it themselves along with the need for an expanded number of trained facilitators to enable spread. Thus, they collaborated with the consultants to assign coaches based on what knowledge resided where – which schools were sending returning, well-trained facilitators and which were not; which coaches were sufficiently trained and which were not – to ensure that needed knowledge resided to the extent possible at every site. Consultants prioritized those sites most in need for their visits.

A number of factors are often assumed or put forth in the literature as necessary for the success of a model at scale that while ideal should not be expected and can be compensated for when lacking. Given the complexity of scaling any model of reform, programs and districts should anticipate lack of up-front knowledge and buy-in; the existence of contradictory theories of change; and the likelihood for those with formal authority (especially district personnel, coaches, and administrators) to have trouble recognizing that which is new in the reform as just that. The authority bestowed by their position makes it especially hard for them to position themselves as learners, and the typical lack of sufficient resources pushes them to lead implementation without sufficient opportunities for learning to do so. All of this should be expected and can be accounted for by a set of conditions that interact in an adaptive system.

Discussion and implications for policy and research

The findings of this study, namely that there are relatively few minimum conditions needed for scale when these conditions operate in an adaptive system, represent a shift from typical thinking and research about scale, as captured in the chart below (see Figure 2). Typically, the dialog and associated research focus on the various factors found to be in place when scale was successful,

drawing across situations to make in essence an accumulated list of what is deemed necessary. There is no doubt that the conditions surfaced in this manner are important and at times essential and that they are in fact ideal. However, findings in this study build on earlier research about the need for adaptive approaches (Coburn, 2003; Farrell & Coburn, 2017; Fullan, 1993; Peurach & Glazer, 2012) to make the case that a self-repairing ecosystem mitigates the need for the ideal.

This study suggests that adequate knowledge of a reform model is necessary but that there are multiple ways to ensure needed knowledge takes root in spite of environmental challenges, including high teacher turnover, as is most likely to occur in those schools and districts most in need of reform, and the presence of contradictory theories of change, also likely, given that reforms seeking to change business as usual generally counter existing theories rather than fit neatly within them. It suggests as well that aligned approaches and incentives up and down levels of a system (federal, state, district, and school) with a reform effort should not be expected as a prerequisite for launching a reform. Rather, working feedback loops, particularly those that allow information to flow between the bottom and top of a reform system (linking evidence of student improvement and teacher knowledge directly to program consultants, in the studied model) allow for adaptability that continuously checks and redirects itself according to evidence of progress in the driving goals. Evidence of progress, in turn, generates buy-in and alignment (supportive environmental conditions) over time.

The most important implication for policy may be that reform at scale, even at the high school level, is more possible than generally imagined – that the lack of ideal conditions does not doom an effort from the start. Those seeking to implement or promote reform at scale should select models that operate as adaptive ecosystems. Such systems will be most able to repair themselves over time and thus to survive in the face of new, ongoing or unexpected challenges. Adaptability may be the most important requirement for scale.

Future research should test the extent to which this proves to be the case in varied settings. Specifically, research should test whether the key conditions identified in this study (adequate knowledge, appropriate environment, and a working feedback loop) are the minimum conditions required for scale of varying reform models that operate as adaptive systems. Research should illustrate what each minimum condition looks like and comprises in these different contexts and

Category of condition	From	To
Knowledge	Up-front understanding and buy-in by key stakeholders, especially principals and district leaders Consistent people sent to be trained over time (minimal turnover)	Buy-in and understanding will be developed, even among key stakeholders, over time Turnover can be accommodated through redundancy and shifts in allocation of resources over time
Environment	Common theories of change across the environment (policy, state, district) that align with reform model	Varied theories of change will be brought into closer alignment with reform model's over time
Feedback	Nested, aligned incentives and approach to change from the top to the bottom (policy, state, district, reform model)	Working feedback loop between top and bottom of initiative (goal at school level and reform program) to bring about alignment up, down and across the system over time

Figure 2. Shifts in thinking about the conditions required for scale within an ecosystem.

paint nuanced pictures of adaptive systems operating in practice. This would provide a more robust research base to support a vision of scale that is attainable within real-world constraints.

Note

1. The Renewal School Program was a signature initiative of New York City's Mayor, Bill de Blasio, who was elected in part on his promise to support rather than close struggling schools, as his predecessor had done. Although SI support for High Schools was found to be effective (Wohlstetter et al., 2018), research of the initiative overall indicates mixed results. The almost 800 million dollar initiative ran from September 2014 until February 2019. Work relevant to SI began in 2014 with one cohort of 14 high schools. A second cohort of 19 additional schools joined in 2015. The number of high schools involved in the initiative shifted over time as new schools were added, and, in some cases, schools were merged or closed. See https://www.chalkbeat.org/tag/renewal-schools/ for an archive of articles about NYC's renewal initiative and https://www.schools.nyc.gov/about-us/initiatives/renewal-and-rise-schools for the latest information about the program published by the NYC Department of Education (DOE).

Disclosure statement

No potential conflict of interest was reported by the author.

References

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 509–535). London, UK: Sage Publications.

Chenowith, K. (2017). *Schools that succeed: How educators marshal the power of systems for improvement*. Cambridge, MA: Harvard Education Press.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3–12. doi:10.3102/0013189X032006003

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. doi:10.1086/667699

Cohen, D. K., & Mehta, J. D. (2017). Why reform sometimes succeeds: Understanding the conditions that produce reforms that last. *American Educational Research Journal*, *54*(4), 644–690. doi:10.3102/0002831217700078

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, *35*(1), 128–152. doi:10.2307/2393553

Datnow, A. (2002). Can we transplant educational reform, and does it last? *Journal of Educational Change*, *3*(3–4), 215–239. doi:10.1023/A:1021221627854

Dilshad, R. M., & Latif, I. L. (2013). Focus group interview as a tool for qualitative research: An analysis. *Pakistan Journal of Social Sciences*, *33*(1), 191–198.

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, *99*(1), 1–26. doi:10.17763/haer.66.1.g73266758j348t33

Ely, M., Anzul, M., Friedman, T., Garner, D., & Steinmetz, A. M. (1991). *Doing qualitative research: Circles within circles*. London, UK: Falmer Press.

Farrell, C. C., & Coburn, C. E. (2017). Absorptive capacity: A conceptual framework for understanding district central office learning. *Journal of Educational Change*, *18*(2), 135–159. doi:10.1007/s10833-016-9291-7

Firestone, W. A., & Herriott, R. E. (1982). Prescriptions for effective elementary schools don't fit secondary schools. *Educational Leadership*, *40*(3), 51–53.

Fullan, M. (1993). *Change forces: Probing the depths of educational reform*. London, UK: The Falmer Press.

Fullan, M. (2000). The three stories of education reform. *Phi Delta Kappan*, *81*(8), 581–584.

Fullan, M. (2002). Leadership and sustainability. *Principal Leadership (Middle School Ed.)*, *3*(4), 14–17.

Fullan, M., Bertani, A., & Quinn, J. (2004). New lessons for districtwide reform. *Educational Leadership*, *61*(7), 42–46.

Gallimore, R., Ermeling, B. A., Saunders, W. M., & Goldenberg, C. (2009). Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams. *The Elementary School Journal*, *109*(5), 537–553. doi:10.1086/597001

Mazzeo, C., Fleishman, S., Heppen, J., & Jahangir, T. (2016). Improving high school success: Searching for evidence of promise. *Teachers College Record*, *118*(13), 1–31.

McLaughlin, M. W., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change*, *2*(4), 301–323. doi:10.1023/A:1014616908334

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

Panero, N. S. (2016). Progressive mastery through deliberate practice: A promising approach for improving writing. *Improving Schools*, 3(19), 229–245. doi:10.1177/1365480216643209

Panero, N. S., & Talbert, J. E. (2013). *Strategic inquiry: Starting small for big results in education*. Cambridge, MA: Harvard University Press.

Peurach, D. J., & Glazer, J. L. (2012). Reconsidering replication: New perspectives on large-scale improvement. *Journal of Educational Change*, 13(2), 155–190. doi:10.1007/s10833-011-9177-7

Schwartz, K. (2017, February). Is it time to go back to basics with writing instruction? *MindShift*. Retrieved from https://www.kqed.org/mindshift/47069/is-it-time-to-go-back-to-basics-with-writing-instruction

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

Siskin, L. S., & Little, J. W. (Eds.). (1995). *The subjects in question: Departmental organization and the high school. The series on school reform*. New York, NY: Teachers College Press.

Snipes, J., Doolittle, F., & Herlihy, C. (2002). *Foundations for success: Case studies of how urban school systems improve student achievement*. Washington, DC: Council of Great City Schools.

Spillane, J. P. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition and Instruction*, 18(2), 141–179. doi:10.1207/S1532690XCI1802_01

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). London, UK: Sage Publications.

Stringfield, S. (2002). Big change questions: Is large-scale educational reform possible? *Journal of Educational Change*, 3(1), 63–73. doi:10.1023/A:1016539814319

Stringfield, S., & Datnow, A. (1998). Scaling up school restructuring designs in urban schools. *Education and Urban Society*, 30(3), 269–276. doi:10.1177/0013124598030003001

Talbert, J. E. (2011). Collaborative inquiry to expand achievement in New York City schools. In J. O'Day, C. Bitter, & L. Gomez (Eds.), *Education reform in New York City: Ambitious change in the nation's most complex school system* (pp. 131–155). Cambridge, MA: Harvard Education Press.

Talbert, J. E., Cor, M. K., Chen, P., Kless, L. M., & McLaughlin, M. (2012). *Inquiry based school reform: Lessons from Sam in NYC*. Stanford, CA: Stanford University, Center for Research on the Context of Teaching.

Tyre, P. (2012, October). The writing revolution. *The Atlantic Magazine*. Retrieved from https://www.theatlantic.com/magazine/archive/2012/10/the-writing-revolution/309090/

Wohlstetter, P., Kim, B., & Flack, C. B. (2018). *Strategic inquiry and New York City's renewal high schools*. New York, NY: Columbia University, Teachers College. Retrieved from https://www.tc.columbia.edu/media/news/images/2018/december/Wohlstetter_Strategic-Inquiry-Final-Report-1.0-1.pdf

Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review*, 27(2), 185–203. doi:10.5465/amr.2002.6587995