Scaling Up: Evolving Innovations beyond Ideal Settings

to Challenging Contexts of Practice

Chris Dede, Harvard University

Dede, C. (in press). Scaling Up: Evolving Innovations beyond Ideal Settings to Challenging Contexts of Practice. In R.K. Sawyer (Ed.), Cambridge Handbook of the Learning Sciences. Cambridge, England: Cambridge University Press.

"Scaling up" involves adapting an innovation successful in some local setting to effective usage in a wide range of contexts. In contrast to experiences in other sectors of society, scaling up successful programs has proved very difficult in education (Dede, Honan, and Peters, in press). Insights from changing operations at one fast-food location may easily transfer to every store in that franchise and perhaps to any comparable type of restaurant. However, a new type of teaching strategy that is successful with one practitioner often is difficult to generalize even to other instructors in the same school, let alone to a broad range of practitioners. In general, the more complex the innovation and the wider the range of contexts, the more likely a new practice is to fail the attempt to cross the "chasm" from its original setting to other sites where its implementation could potentially prove valuable (Moore, 1999). In other words, scalable designs for educational transformation must avoid what Wiske and Perkins (2005) term the "replica trap": the erroneous strategy of trying to repeat everywhere what worked locally, without taking account of local variations in needs and environments. This involves resolving problems of magnitude (fostering the necessary conditions for change in large numbers of settings with average resources at

considerable distances from one another) and variation (diverse and often unfavorable conditions across settings).

In the context of innovations in teaching/curriculum, Coburn (2003) defines scale as encompassing four interrelated dimensions: depth, sustainability, spread, and shift in reform ownership. "Depth" refers to deep and consequential change in classroom practice, altering teachers' beliefs, norms of social interaction, and pedagogical principles as enacted in the curriculum. "Sustainability" involves maintaining these consequential changes over substantial periods of time, and "spread" is based on the diffusion of the innovation to large numbers of classrooms and schools. "Shift" requires districts, schools, and teachers to assume ownership of the innovation, deepening, sustaining, and spreading its impacts. A fifth possible dimension to extend Coburn's framework is "evolution," in which the innovation as revised by its adapters is influential in reshaping the thinking of its designers, creating a community of practice that evolves the innovation.

This chapter defines "scaling up" individual educational innovations as different from "systemic reform." An extensive scholarly literature documents the challenges of systemic reform, which involves implementing suites of innovations that collectively transform the mission, goals, processes, personnel, and products of an educational institution (Berends, Bodilly, and Kirby, 2002). Scaling up a suite of innovations through systemic reforms that assure their incorporation and effectiveness by changing the entire context of their implementation to embed their conditions for success is less frequent and more challenging than the typical

situation of scaling up a single innovation through implementation in a potentially resistant educational setting that lacks some conditions related to the innovation's prior successes. In this chapter, both types of scaling up are discussed.

An Analogical Conceptual Framework for Scaling Up in Educational Settings

Adaptation of an organism, innovation, or organization to local conditions is a fundamental phenomenon in both natural and human settings. Analogies related to various types of adaptation are helpful for understanding the special case of scaling up in educational settings. This section draws on insights about adaptation from studies in both the biological and social sciences to construct an analogical conceptual framework for the adaptation of educational innovations.

The adaptation of biological species is distinct from human adaptation in both its rate and its degree of intentionality. Over long periods of time, shifts in an ecosystem can induce the adaptation of species that inhabit that specific context (Moya and Font, 2004). As one illustration, changes in a natural environment may mean that particular forms of coloration for a type of insect provide better camouflage, enhancing the chances of its survival. As a result, through natural selection over time that coloration becomes more prevalent in that insect species. In contrast, the shifts in their context were evidently too great for random mutations to generate the necessary adaptive changes in time to save many types of dinosaurs. In ecosystems, adaptation is a non-purposive, slow process driven by natural selection; the "innovation" of a change in a species is emergent rather than deliberately designed and implemented.

In human settings, a somewhat analogous process of adaptation to contextual conditions occurs. For example, in the rapidly shifting context of modern economies, many businesses must frequently alter their products and processes to stay competitive or even viable. In turn, employees of these businesses find that various types of skills or knowledge wax or wane in importance. Some workers intentionally adjust their mix of capabilities to adapt to this shift; others unwilling or unable to change may lose their jobs. In contrast to the slow and non-purposive process of natural selection in ecosystems, the human progression of contextual change, organizational response, and individual adaptation is sometimes deliberate, sometimes inadvertent, and has accelerated markedly over the past few decades (Levy and Murnane, 2004).

For both natural and human settings, the context itself is shaped by successes and failures of adaptation. Through systemic processes, the interactions among various flora and fauna in an ecosystem are altered by the prevalence of each species. For example, pine trees lose needles that, on the forest floor, inhibit some other types of trees from growing. Similarly, the characteristics of a society are affected by the types of businesses that prosper or fail. As an illustration, the success of a factory creates regional conditions (e.g., economic prosperity, a skilled workforce, rail transport) conducive to opening other factories. These reciprocal interactions between the context and its organisms or organizations create complex longitudinal processes of mutual adaptation (Midgely, 2003).

Altering suites of internal institutional policies and practices in response to a change in organizational context may require complex forms of adaptation. If the shifts in setting are small, many institutions can readily make relatively minor

alterations in subsets of their standard procedures and human resource capabilities. For example, to reduce costs, a manufacturing plant might shift from keeping large inventories of parts to using a just-in-time logistical system based on sophisticated tracking mechanisms and constant communication with suppliers. The products the plant builds remain the same, as do the roles of many employees outside of the operational sector that has altered. However, larger changes in context may demand purposive, transformational adaptations in objectives, products, policies, practices, and personnel throughout an organization. As an illustration, the manufacturing enterprise may decide to alter its entire product line in response to changing market conditions, affecting almost every role within the company. Such transformational shifts are quite difficult to accomplish for any enterprise; the challenges of rapid, discontinuous institutional changes are well documented in the scholarly literature on organizational innovation (Agyris, 2004).

Despite their differences, all these types of individual and organizational adaptation, biological and social, have fundamental similarities. Systems models of organisms, people's cognitive and affective processes, organizations, and contexts are all based on nested, interrelated, longitudinal dynamic interactions that reinforce change (positive feedback loops) or resist change (negative feedback loops). Both types of feedback are important, but change-resisting mechanisms tend to predominate, to preserve the entity's integrity and to enhance survival. Such conservative responses to change are particularly evident in educational settings (Senge, 2000). For example, local communities in the United States historically have exerted a high degree of control over public schools

through politically volatile vehicles such as school boards. As a result, attaining the coherent, sustained will and resources to achieve a nationwide transformational shift has been extremely difficult (Tyack and Cuban, 1996), even though innovation within individual districts is easier than in countries dominated by national education policies. Other countries with educational systems historically more strongly shaped by national policies have experienced fewer barriers to transformational change.

While this preponderance of preservational processes is valuable for many reasons, resistance to change is often a liability at times of rapid shifts in individual and organizational context. For example, economists and high-tech businesses believe that the emergence of a global, knowledge-based economy demands that education provide its graduates with different skills and knowledge than were optimal for industrial civilization (Partnership for 21st Century Skills, 2003; Levy and Murnane, 2004). Yet current educational reform initiatives in many countries are regressively emphasizing basic skills and broad, shallow content, rather than stressing higher order skills based on deep knowledge of a few core principles (Dede, 2003). Transcending change-resisting processes to enable the evolution of new types of behaviors is challenging for individuals, but even more difficult for organizations, since their institutional policies and practices as well as the employees who carry them out must alter. Not surprisingly, given the rapid rate of societal change in modern times, organizations have a higher rate of extinction than individuals do (Klein, 2000).

In response to the threats to identity and viability that shifts in context pose, both individuals and organizations seek to gain the power to shape their

settings. Kings and dictators impose national policies and induce cultural beliefs that increase their personal power over their countries. Businesses may seek monopoly status for their products and services, may try to alter the governmental policies that govern their activities, or may attempt to influence the culture of societies in which they function in ways favorable to their interests.

In turn, human settings also shape which types of individuals and organizations prosper within societies through cultural and economic mechanisms. For example, the characteristics and capabilities of successful people and thriving organizations differ within capitalist and socialist economies. Also, a country may seek to shape its larger global context to reduce threats to its identity and survival. Thus, whether a particular entity is seen a context to which other entities adapt (e.g., a nation to which its organizations adapt) or an entity adapting to its context (for example, a nation shaped by its role in the larger setting of global civilization) depends on the analytic perspective utilized.

In summary, this analogical conceptual framework is based on the assumption that coupled, cyclic, hierarchically nested, longitudinal, change-reinforcing and change-resisting processes (hereafter referred to as "systems dynamics models") can model the behaviors of individuals, innovations, and organizations in ways that lead to insights about their interrelationships and their ability to evolve (Morecraft and Sterman, 2000). In applying this modeling strategy to adaptation in a variety of natural and social systems, the following high-level observations emerge:

- Contexts shape the entities that inhabit them (organisms, people, organizations) through rewarding or inhibiting various types of

behaviors. As an illustration, knowledge based economies provide financial incentives for individuals to attain 21st century skills, which in turn influences the mission of schools (Partnership for 21st Century Skills, 2003).

- Entities influence their setting by changing its characteristics in ways that alter the behaviors that context reinforces or suppresses. For example, parents who have attained high levels of education tend to value quality schools for their children and to act in ways that encourage their society to invest in education, a self-reinforcing feedback loop.
- In social settings, the rate of innovation has accelerated in recent decades through mutually reinforcing feedback loops augmenting shifts in context, organizations, and individuals (e.g., the rapid evolution of knowledge economies relative to the comparable development of industrial or agricultural economies). Tyack and Cuban (1996) delineate the many change-resisting processes characteristic of educational institutions, and Cuban (2001) further documents the particular challenges educational innovations based on information technology entail. Because of these factors, the general acceleration of innovations over recent decades has affected schooling less than other sectors of society.
- People can respond to environmental feedback by deliberately altering their individual knowledge and skills. However, for people to rapidly modify these in a major way is difficult because

of various change-resisting mechanisms, such as cognitive and affective limits on how quickly and thoroughly teachers can unlearn behaviors that were successful in their schools before an innovation changed their context (Spillane, 2002).

- Institutions can respond to environmental feedback by modifying their practices and practices and by shifting capacity-building strategies for their employees. However, for organizations to rapidly modify these processes in a major way is difficult because of various change-resisting mechanisms. For example, many school districts have multiple layers of review and approval that make changing operational practices difficult, as well as affiliated institutions (e.g., teacher unions) that can retard changes in human resources, policies, and roles.

How does this analogical conceptual framework inform studies of scaling up innovations in education? The challenges involved in adapting an intervention successful in some local educational setting to effective usage in wide range of other contexts draw on aspects of all the observations delineated above. Insights about adaptation from studies in both the biological and social sciences may suggest new approaches to the process of educational innovation, based on the adaptation of strategies for scaling up proven successful in these other sectors and settings.

The remainder of this chapter applies this high level-framework in describing and contrasting the analytic strategies used by scholars to study various educational cases in which success was achieved in scaling up. The cases

considered are drawn from presentations at an invitational research conference on scaling up technology-based educational innovations held at Harvard in 2003 (http://www.gse.harvard.edu/scalingup/) and papers in the book that resulted from that meeting (Dede, Honan, and Peters, in press).

Scaling Up a Set of Exemplary Instructional Practices and Curricula

As part of a NSF grant to create a Center for Learning Technologies in Urban Schools (http://www.letus.org), in 1997 the University of Michigan and Northwestern University partnered with the Detroit Public Schools and the Chicago Public Schools to improve urban science education. This LeTUS initiative developed hands-on, project-based, technology-intensive curricula (Krajcik, Blumenfeld, Marx, and Soloway, 2000), then worked to scale up the implementation of those curricula beyond those teachers who participated in designing this innovation into classrooms throughout these districts, and potentially to other districts as well. In Detroit, the scaling up strategy involved both broad-based professional development and close collaboration with school and district administrators, including a high-level champion in the district's central office.

In particular, the professional development activities included extended summer workshops, monthly Saturday work sessions, in-classroom consultations, and online professional development environments (Fishman, Marx, Best, and Tal, 2003). In addition, the curriculum itself was designed to be educative for teachers (Schneider and Krajcik, 2000) and to meet community needs (Moje, Collazo, Carillo, and Marx, 2001). The LeTUS team also worked to incorporate

this professional development initiative into the larger context of a systemic reform movement in the Detroit Public Schools.

Through these combined efforts, approximately 65 teachers, representing 26 percent of all middle-grade science teachers in Detroit, are now using these innovative curricula and pedagogies (Eighty-five teachers in Detroit have worked with LeTUS over the years, but due to promotion and attrition, roughly twenty no longer directly teach science; there are approximately 250 middle-grade science teachers in Detroit.) This is certainly a success in scaling up relative to many other curricular/pedagogical innovations, but it also indicates the difficulties of persuading the majority of an educational innovation's potential users to adapt it.

The LeTUS team developed a framework for evaluating the fit between innovations and intended contexts of use (Blumenfeld, Fishman, Krajcik, Marx, and Soloway, 2000). The underlying model for this framework is "usability," the extent to which people can use tools or innovations to accomplish work (Nielsen, 1993). Three dimensions of usability (capability, policy and management, and school culture) are arrayed as three axes originating from a common point (the origin, which represents the current capacity of the district to use the innovation) to form a three-dimensional space. The innovation is mapped into this space, and its distance from the origin represents a gap between the capacity required to successfully use the innovation and the current capacity of the school district.

Conceptualized in this manner, scaling up involves closing gaps that exist between the innovation's demands and an organization's capacity. Closing a gap on the culture axis (the extent to which an innovation adheres to or diverges from the existing norms, beliefs, values, and expectations for practice at different levels

of the system) may entail providing opportunities for teachers and administrators to gain new visions of practice and policy consistent with features of the innovation. Closing a gap on the capability dimension may involve providing professional development for teachers or modifying activities and redesigning technology to reduce new knowledge and skills required for effective implementation. Closing a gap on the policy and management dimension may require changing school and district policies and procedures, as well as adapting the innovation to improve its fit with current practices. In close collaboration with the Detroit Public Schools, the university's design team made all these types of adaptations to close the usability gap with that particular context (Blumenfeld, Fishman, Krajcik, Marx, and Soloway, 2000).

Applying the analogical conceptual framework sketched earlier to this case study and similar initiatives of scaling up curricular/pedagogical innovations requires the use of a structural category system that maps the parts of the innovation's organizational context. Structural issues are not important in non-purposive change (such as biological evolution or inadvertent institutional shifts), but are central to deliberate organizational innovation. As an example, Russell, Bebell, and O'Dwyer (2003) studied a variety of factors thought to influence the conditions for success of the implementation of instructional technology in school districts (Table 1).

==INSERT TABLE 1 HERE==

In any given situation, various factors might assume greater or lesser importance.

Reasons of space preclude mapping all the interventions involved in the scaling up strategy used by the University of Michigan and the Detroit Public Schools onto this hierarchically nested structure of factors. Such a mapping can provide insights as to which types of interventions from the full range of possibilities this group's adaptation strategies emphasized. In this case study, the partnership particularly focused on influencing teachers' beliefs about pedagogy, providing teachers with technology support services and professional development about technology usage, and adapting the innovation in various ways to reduce the size of the gap teachers and administrators faced along the three usability dimensions (Fishman, in press).

Such a strategy emphasizes the key role of "alignment" (coherence and mutual reinforcement among shifts in policies and practices) in scaling up innovations. Cohen and Hill's study of effective state educational reform approaches (2001) shows the importance of coherence among curriculum, professional development, and student assessment. Striking a chord of mutually reinforcing innovations is important in enabling each to have the "conditions for success" required for scaling up (Dede & Nelson, in press). The effective use of antibiotics illustrates this concept: Antibiotics are a powerful "design," but worshiping the vial that holds them or rubbing the ground-up pills all over one's body or taking all the pills at once are ineffective strategies for usage – only administering pills at specified intervals works as an implementation strategy. A huge challenge educators face, and one of the reasons this field makes slower progress than venues like medicine, is the complexity of conditions for success—and the sophistication of the processes necessary to achieve these conditions—

required in effective interventions. Nothing powerful in facilitating learning is as simple or as easily administered as an inoculation in medicine. Fostering coherence and alignment among a suite of innovations that is less comprehensive than a full-scale systemic reform effort, but simultaneously is implemented to provide each other's conditions for success, is an effective approach to scaling up.

Beyond delineating the nested, hierarchical interrelationships among various conditions for success, analysis using the structural category system above helps to show how adapting innovations is shaped by both change-resisting and change-reinforcing processes (as described earlier in the high-level observations on the analogical framework). For example, providing technology support services involved overcoming various types of concerns from the Detroit Public School's management information systems department, which is responsible for its technology infrastructure. Linking the innovation to the larger context of the systemic reform initiative underway in the district provided a change-reinforcing mechanism to help resolve these concerns.

Part of the goal of LeTUS is to influence participating districts not only via direct effects of implementing technology-intensive, inquiry-based curricular units, but also through shaping schools' policies and culture to build capacity for usability, enabling the facile adaptation of future, similar innovations. As the high-level observations on the analogical conceptual framework suggest, two substantial factors in this evolution are teachers' ability to reconceptualize their professional roles and the district's capability to alter its policies for both practice and human resources management. The emphasis in the usability framework on policies and culture reflects the appropriate importance the University of

Michigan / Detroit Public Schools partnership places on altering underlying systemic dynamics, rather than focusing primarily on superficial changes in policy and practice targeted just to this specific innovation. Such a strategy for adapting innovations is a stepping-stone towards the "systemic reform" type of scaling up, discussed next.

Comparing Scaling Up Innovations via Systemic Reform Initiatives

As discussed at the start of the chapter, this analysis views scaling up a single innovation into a school or district as different than conducting a full-scale systemic reform of that educational organization. Contrasting the endpoints of the continuum between, on the one hand, adaptation of an isolated intervention into a single setting lacking some of its conditions for success and, on the other hand, transforming the entire context of a school district through a suite of innovations that embed their conditions for success highlights similarities and differences in strategy between these two types of educational improvement. Union City, New Jersey is a well studied instance of systemic reform based on adapting suites of innovations.

As Carrigg, Honey, and Thorpe (in press) describe:

> In 1989, the Union City school district was the second-worst-performing district in New Jersey. It had failed forty-four of fifty-two indicators that the state uses to determine the efficacy of school systems; in fact, the state had threatened to take over governance unless radical and successful restructuring was implemented within five years… The transformation in academic achievement that the district experienced during the 1990s and has sustained into the current decade constitutes a surprising success story.

> By 1995, Union City's average scores on the state's eighth-grade readiness test surpassed those of its urban counterparts by as much as 20 percentage points… By 2002, Union City's test scores ranked highest among New Jersey cities with populations of 50,000 or more.

What suites of innovations led to these dramatic advances, and what systemic reform processes were used to scale these across the district?

The reform efforts initially emphasized literacy, seeing this as a pre-requisite for many other forms of learning. Particularly in the early grades, district leadership implemented a variety of pedagogical and curricular innovations in teaching reading. Also, teachers began to infuse language and reading into all areas of the curriculum, using a long-range strategy for gradual implementation of innovations; a decade was required for full implementation of shifts in literacy instruction from grades K-12. Details on the specifics of the literacy innovations are available on the Union City district website (http://www.union-city.k12.nj.us/curr/k12curr/escurr/1-4humanities/index.html).

In part because improvements in student outcomes came rapidly and were well publicized, the district benefited from stable political leadership and community support over this entire time period. To promote this, school leaders regularly solicited community feedback during this process and encouraged parent buy-in and collaboration. To aid teachers as they worked toward proficiency in the innovations, the district shifted to a five stage model of professional development: awareness, practice, sharing, peer coaching, and mentoring.

Basing systemic reforms on strong, stable leadership; teacher ownership; and community support is a pattern commonly seen in successful systemic reform. Snipes, Doolittle, and Herlihy conducted a study, *Foundations for success: Case studies of how urban school districts improve student achievement* (2002), sponsored by the Council of the Great City Schools and MDRC (a research organization). This study focused on determining which large urban districts have improved on a system-wide scale and what common factors across these district initiatives seem responsible for successful reforms. Their major findings on district characteristics that promote success in systemic reform are:

- Urban school districts that have improved performance on a broad scale share certain preconditions for reform, such as political and organizational stability over a prolonged period and agreement among school board members, the superintendent, and community leaders that student achievement is the top priority.
- District leadership can play a key role in scaling up improvements through strategies such as setting district-wide goals, holding district- and building-level administrators personally accountable for results, adopting uniform curriculum and instructional approaches that apply to every school, and redefining the main role of the central office as one of guiding, supporting, and improving instruction at the school building level.
- Faster-improving urban school districts provide principals and teachers with early and ongoing assessment data, along with

> training and support to help them use these data to improve teaching and learning.

While Union City is smaller than many of the urban districts participating in this study, its case study affirms these conditions for success in adapting innovations.

Beyond the factors already discussed, to monitor students' growth and achievement in literacy over time, Union City utilized both formal and informal assessment methods (Carrigg et al, in press). The objective was to focus on students as individuals, evaluating their progress based on their abilities and learning styles. The district implemented a mix of diagnostic, formative, and summative assessments. In part because the diagnostic and formative measures were used to improve and individualize instruction, students' test scores improved substantially without basing the curriculum around high-stakes testing.

These and other aspects of the Union City case study resonate with a key design principle Goldman (in press) describes as underlying successful systemic reforms: Conducting a process of a continuous inquiry—based on access to information, analysis of information, and actions aimed at supporting what is going well and improving what is not—is important in making sustainable progress towards educational improvement. In that respect, scaling up is like biological evolution--but more purposive: One watches to see what emergent innovations occur (mutations), then allocates resources for those that are successful. Means and Penuel (in press) support this principle of data-based decision making, indicating that the research base needed for scaling up goes

beyond the question, “What works?” to a more complex question, “What works when, and how?” They state

> Rather than an average effect size, local decision makers need research findings that shed light on the expected effects under different circumstances and on the contextual and implementation factors that are likely to influence success. An emphasis on average effects can be counterproductive if it results in inattention to these critical factors in efforts to move interventions to new settings and to scale. (p. 216)

Findings that link gains in student achievement to detailed, practical strategies for implementation also help practitioners and policymakers make the case for further innovation.

How does the Union City case study illustrate differences in strategy between scaling up a single innovation and conducting a systemic reform? Without mapping all the interventions involved in the scaling up strategy used by Union City onto the structural categories delineated earlier and comparing this with the strategy used by the University of Michigan and the Detroit Public Schools, a few generalizations are apparent. Systemic reforms tend to utilize an entire suite of mutually reinforcing innovations, a broader range of adaptations to provide the conditions for success for those innovations, and a longer time-frame to enable the full institutionalization of transformation change. Systemic reforms also use district-wide, data-based decision making to develop strategies that increase investment in promising innovations, decrease support for innovations not successfully adapted to this context, and identify change-reinforcing processes to aid in institutional transformation. But what about scaling up in settings that

not only are unwilling to undertake full-scale systemic reform, but also are largely uninterested in implementing even isolated innovations?

Scaling Up Innovations Without Partnering with Local Contexts

Another type of scaling up is designing educational innovations to function effectively across a range of relatively inhospitable settings (Dede, 2004). This is in contrast to the models presented thus far for effective transfer of an innovation to another context, which involve partnering with a particular school or district to make that setting a conducive site for adapting a particular design. Scalability into typical school sites that are not partners in innovation requires developing interventions that retain substantial effectiveness in relatively barren contexts, such as urban schools, in which an innovation's conditions for success (e.g., supportive administration, qualified and enthusiastic teachers, a well maintained technology infrastructure, a student population consistently present) may be absent or attenuated. Under these circumstances, major intended aspects of an innovation's design may not be enacted as intended by its developers, who can anticipate that parts of their design will be "defenestrated" (thrown out the window).

Evolving a design for scalability even into contexts in which its conditions for success are attenuated or lacking requires enhancing the robustness of its effectiveness when parts of its intended enactment are defenestrated. Such "design-for-defenestration" is exemplified in studies the author and his colleagues are conducting. With National Science Foundation funding, we are creating and studying graphical multi-user virtual environments (MUVEs) that enhance middle

school students' motivation and learning about science and society (Dede, Nelson, Ketelhut, Clarke, and Bowman, 2004).

Our “River City” MUVE is centered on higher order inquiry skills such as hypothesis formation and experimental design, as well as on content related to national standards and assessments in biology and ecology (http://muve.gse.harvard.edu/muvees2003/). Through design-based research (Dede, in press), we are documenting how students can gain this knowledge through immersive simulations, interaction with digitized museum artifacts, and "participatory" historical situations. Students learn to behave as scientists by collaboratively identifying problems through observation and inference, forming and testing hypotheses, and deducing evidence-based conclusions about underlying causes. The goal is to promote learning for all students, particularly those who are unengaged or low performing.

Design-for-defenestration involves identifying conditions for success likely to be attenuated in many contexts, then evolving the design to retain substantial effectiveness under those circumstances. For example, in some implementations of the MUVE, a few teachers ignored all or most of the professional development made available online. These teachers then typically encountered problems in implementation, such as not understanding the purpose and process of the curricular intervention, lacking knowledge about the higher order inquiry skills and standards-based scientific content the intervention helps students to learn, and missing skills in leading the small group and whole class interpretive discussions important for students’ understanding of both their MUVE experiences and the data collected. Although this list sounds quite grim,

in practice the curricular intervention worked fairly well in these situations. The MUVE is designed for scalability, creating curricular interventions so compelling for students and with sufficient internal guidance so that they have a fulfilling, self-directed learning experience—albeit with reduced educational outcomes—even with a confused teacher.

In response to attenuation of the teacher-preparation condition for success, we evolved the professional development portion of the design to increase its scalability. For example, we produced a just-in-time, "light" version of the online professional development that an overwhelmed teacher can skim for ten minutes per day during the unit, providing essential information needed to guide students for that stage of the learning experience. We are also designing variants of the MUVE that simplify the teacher's role without substantially compromising the effectiveness of the innovation.

However, some aspects of any educational innovation are difficult to resolve through robust designs. For example, for MUVE implementations in urban sites, student attendance rates for class averaged about 50% (although this improved during the implementation of the learning experience, an encouraging measure of its effectiveness). Also, in the shadow of high stakes testing and accountability measures mandated by the federal No Child Left Behind legislation, persuading schools to make available two weeks of curricular time is difficult for any design that does not use traditional pedagogy to inculcate students with basic skills and factual content.

These pose challenges difficult to overcome by even the best robust designs. However, innovators can still attempt to get leverage on these factors.

For example, the MUVE curriculum is very engaging for students and teachers, uses standards-based content and skills linked to the high stakes tests, and shows strong outcomes with sub-populations of concern to many schools worried about making adequate yearly progress across all their types of students.

Reasons of space preclude mapping design-for-defenestration strategies onto the analogical conceptual framework described earlier. However, the pattern of investments in scalability that such an analysis generates is different than for strategies that involve partnering with implementation sites to enhance their capabilities. In robust designs, more resources are invested in developing variants of the innovation adapted to special circumstances, less in building capacity at implementation sites. Overall, design-for-defenestration may represent a more effective strategy for moving to very large scale across many sites, but is likely less effective in developing high levels of usage at a particular site than the alternative scaling up strategy of working with that particular context to build its capacity.

Thus far, this discussion of scalability has centered on curricular and pedagogical innovations in pre-college schooling. How do effective scaling up strategies differ with various types of innovations and different audiences served? The case study that follows exemplifies the challenges of scaling up innovations in teacher professional development not linked to any particular local context, but disseminated worldwide.

Scaling Up a Context-Independent Strategy for Teacher Professional Development

What issues are involved in scaling up when adaptation to a single setting is not a concern because the innovation is a generic online service for teacher professional development? Operating through the Internet, WIDE (Wide-scale Interactive Development of Educators) World focuses on professional development of constructivist teaching practices for schools and other settings (Wiske and Perkins, in press). Participants include practicing teachers, professors, teacher developers, administrators, and others actively engaged in education. WIDE World courses emphasize active experimentation with various pedagogical frameworks in one's professional context (http://wideworld.pz.harvard.edu/). A part of WIDE World's research program is studying how this initiative encounters and resolves challenges to using the World Wide Web for scaling up educational improvement.

This innovation was designed in response to the realization that even well constructed education-oriented websites do not provide the kind of sustained guidance and support most teachers need in order to make sustained, significant changes in their practice. This shortfall is not surprising in a passive, largely presentational medium like websites. In contrast, WIDE World offers semester-long professional development courses in which participants learn about research-based pedagogies, apply these principles in designing and enacting new approaches with their own students, receive frequent support and feedback from a coach, and engage in regular reflective exchanges with fellow participants in the course. The courses focus on the development of new practices, in contrast to on-

site or on-line university courses that foreground academic learning. In the fall of 2004, WIDE courses involved approximately 773 participants representing 40 countries throughout the world. Since WIDE World's inception, WIDE has worked with more than 3,000 participants representing 82 countries.

WIDE World is designed so that its professional development takes into account the many dimensions of context that affect attempts to implement organizational innovations. Consistent with other typologies summarized in this chapter, Bolman and Deal (1997) developed a taxonomy of four dimensions that need attention and coordination when fostering change in organizations: *human resources* (knowledge, skills, and beliefs of people), *structural* (roles, relationships, schedules, and other forms of organizational structures), *cultural-symbolic* (norms, values, symbols, rituals, and rewards that affect perceptions of meaning and well-being), and *political* (the allocation of authority and responsibility, and commitment from stakeholders). To this, Wiske and Perkins (op. cit.) add a fifth dimension, *technical* (tools, technologies, materials, and other tangible resources).

Given this framework, the design of WIDE World is based on two central assumptions (Wiske and Perkins, in press.). First, improving education requires bridging the knowledge-action gap: the gulf between current understandings of best practice and actual practice (Perkins, 2003; Pfeffer and Sutton, 2000). One strategy is to cultivate research-based pedagogical craft while adapting to or adjusting the context to support wide-scale change. This requires investments in the human resources and technical dimensions of the Bolman-Deal taxonomy as augmented by Wiske and Perkins (in press), as well as shifts in the structural,

cultural-symbolic, and political dimensions. Second, designs should be based on explicit scaling models: causal theories about how the designs address some, if not all, aspects of craft and context as they are complicated at scale by problems of magnitude and variation

The scalability model for WIDE World that emerges from these assumptions relies on human interaction rather than prepared materials to build teachers' capacity for innovation. WIDE World utilizes expert coaches to provide tailored support and suggestions and to promote interaction among peers as teachers change their practice, to manage mutual adaptation of WIDE innovations and local context, and to augment the support provided by materials. Wiske and Perkins (in press) describe the pedagogical framework thus (page 61):

Learning occurs through presentations of ideas by the instructor, short assignments of reading or examination of on-line models, activities in which participants try out new practices and post designs or reflections on-line, feedback from coach and peers, and participation in reflective on-line discussions. WIDE World courses guide participants in designing, applying, critiquing, and revising new practices, through multiple cycles of exchange with peers and coaches, with an emphasis on changing participants' practice as well as their minds.

Challenges to this scalability model include:

- Financial issues (will a sufficient number of participants pay for WIDE World services to sustain the enterprise?)
- The availability of coaches (will a sufficient fraction of WIDE World participants choose to serve as coaches for future participants?), and

- Effectiveness issues (to what extent can the learning model described above resolve across distance the problems of adaptation to context discussed throughout this chapter?).

Similar challenges of scalability are intrinsic to other initiatives that provide online teacher professional development services, such PBS's TeacherLine (http://teacherline.pbs.org/teacherline/), EDC's Ed Tech Leaders Online (http://www.edtechleaders.org/), and TERC's collaboration with Lesley University, Science Online (http://scienceonline.terc.edu).

Like the design-for-defenestration approach, WIDE World does not focus on a particular site like Detroit or Union City. Such decontextualized strategies for scaling up have the strength of potentially influencing a much broader audience, but the challenge of generating a substantial impact on practice when all the support mechanisms rely on mediated communication, rather than face-to-face interactions. In these models for scaling up, more responsibility falls on the remote participant to accomplish successful adaptation of both the innovation and that setting. This involves both selecting which aspects of the Bolman-Deal structural category system described earlier are most important to modify and devising implementation strategies that use change-reinforcing processes to overcome change-resisting mechanisms.

This chapter has described several alternative approaches to scaling up: curricular/pedagogical innovations, systemic reforms, design-for defenestration, context-independent online teacher professional development. Depicting their comparative strengths and limits provides policymakers and practitioners with a sense of which strategy might best suit their particular situation. However,

conducting research on innovation processes this complex is quite challenging. Beyond standard methodological approaches, what new types of research models are needed?

Developing New Analytic Methods for Studying Adaptation and Scalability

Developing new methods of analysis for studying adaptability and scalability, then examining their validity and value is an important frontier for research. For example, researchers, policymakers, and practitioners would all benefit from the creation of a generalizable metric for assessing the scalability of an educational intervention or design (Dede, 2004). Such an index would measure the degree to which the educational effectiveness of the design is robust despite attenuation of its conditions for success. Through identifying factors within the intervention's context that represent important conditions for success and summarizing the extent to which the effect of the intervention is sensitive to variation in each, this index could provide prospective adopters of the innovation a better sense of what its likely effectiveness would be in their own particular circumstances.

A limited taxonomy of important contextual factors that can serve as viable conditions for success across many types of educational interventions is essential for such an index. The work of Russell and his colleagues (2003) on a structural category system for factors that influence the effectiveness of educational technology, discussed earlier, is one example of such a taxonomy. The conditions for success exist at several nested levels of the educational hierarchy. This nesting not only complicates the creation of a scalability index,

but also increases the challenge of estimating the precision of this measure at each level.

Fortunately, for many types of innovations, a relatively small set of contextual factors are often very influential in determining effectiveness. Potential influential factors in this subset may include teachers' knowledge of content and pedagogy, students' socioeconomic and linguistic backgrounds, students' mobility and absenteeism, and (for technology-based innovations) the extent and reliability of the computer/networking infrastructure. Examining scalability in the context of this subset of powerful conditions for success may still yield a workable index, but only investigating its feasibility through using real data can determine the potential validity and value of such a measure.

At its core, the evaluation of the sensitivity of an intervention's impact to select contextual conditions is a question of statistical interactions. In evaluating the sensitivity to the conditions for success, one asks: Is the effect of the intervention dependent on the selected contextual conditions? For example, is the intervention more effective for children of lower SES, or higher? Does the impact of the intervention depend on specific teacher capabilities? On features of the classroom and school infrastructure?

An accurate scalability index must ensure that such interactions are included in the statistical models that underpin the data-analyses conducted to assess the implementation of educational interventions. If the interactions have a statistically significant effect, then we know that the effect of the treatment is sensitive to the conditions that participated in the interaction. Estimating the various effect sizes anticipated for the intervention under each of the interacting

conditions may enable pooling these into a global index of scalability that captures the extent to which the intervention's effect size is sensitive to variation in the conditions for success.

Several important technical challenges to implementing this approach in practice make the validity of this measure uncertain. First, how one should pool the several effect sizes, representing variation in the intervention's impact across levels of a particular contextual factor, into a single index of sensitivity or scalability is uncertain. Second, as conditions for success are drawn from higher levels of the organizational hierarchy (classrooms, schools, districts), mustering the statistical power necessary to detect interactions between these conditions and the intervention being studied is increasingly difficult. As a result, for conditions of success lying at higher levels of the organizational hierarchy, it may not be possible to estimate the sensitivity of the treatment effect to these conditions in a single study. However, a synthesis of findings across many studies, in the manner of meta-analysis, might suffice. Overall, examining the feasibility of new methods of analysis, such as a scalability index, is an important frontier for research on scaling up and may generate strategies for quantifying various aspects of the analogical conceptual framework for scaling up described earlier.

Summary

In their summary of participant discussions at the Scaling Up Success conference referenced earlier, Dede and Honan (in press) identify four key themes in adapting an educational innovation successful in some local setting to effective usage in wide range of contexts:

- *Coping with change:* context, leadership, and funding

- *Promoting ownership:* building constituent support; institutionalizing innovations
- *Building human capacity:* working with collaborators and partners; providing professional development
- *Effective decision making:* interpreting data; creating and applying usable knowledge

This chapter describes how these themes are articulated in four types of scaling up strategies.

Developing mechanisms to provide funding and build capacity in research, practice, and policy for studying these and related issues is an important next step. At present, U.S. resources for scholarship are focused on the creation of "scientifically based knowledge" through clinical intervention studies involving random assignment, a model for evolving educational effectiveness that captures only part of the research needed to effectively enable the scaling up approaches discussed above. Clinical trials can aid in determining what to scale, but not how to effectively adapt that innovation to various local situations. Hopefully, this synthesis and related work will inspire public and private sources who realize the crucial nature of adaptation-oriented research (as discussed earlier, "What works when and how") to provide the substantial, sustained support required to undertake these types of sophisticated studies.

References

Agyris, C. (2004). Reasons and rationalizations: The limits to organizational knowledge. Oxford: Oxford University Press.

Berends, M., Bodilly, S.J., & Kirby, S. (2002). Facing the challenges of whole-school reform: New American Schools after a decade. Santa Monica, CA: RAND Corporation.

Blumenfeld, P., Fishman, B.J., Krajcik, J.S., Marx, R.W., & Soloway, E. (2000). Creating usable innovations in systemic reform: Scaling up technology-embedded project-based science in urban schools. *Educational Psychologist,* 35(3), 149–164.

Bolman, L.G., & Deal, T.E. (1997). Reframing organizations: Artistry, choice, and leadership (2nd ed.). San Francisco: Jossey-Bass.

Carrigg, F., Honey, M., and Thorpe, R. (in press). Moving from successful local practice to effective state policy. In C. Dede, J. Honan, & L. Peters, Eds, Scaling up success: Lessons learned from technology-based educational innovation. New York: Jossey-Bass.

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

Cohen, D.K., & Hill, H.C. (2001). Learning policy: When state education reform works. New Haven, CT: Yale University Press.

Cuban, L. (2001). Oversold and underused: Computers in classrooms. Cambridge, MA: Harvard University Press.

Dede, C. (in press). Why design-based research is both important and difficult. *Educational Technology*

Dede, C. (2004) *Design for Defenestration: A Strategy for Scaling Up Promising Research-based Innovations.* Chicago, IL: NORC.

Dede, C. (2003). No Cliché Left Behind: Why Education Policy is not like the Movies. *Educational Technology* **43**, 2 (March-April), 5-10.

Dede, C., Honan, J., & Peters, L., Eds. (in press). Scaling up success: Lessons learned from technology-based educational innovation. New York: Jossey-Bass.

Dede, C., & Nelson, R. (in press). Technology as Proteus: Digital Infrastructures that Empower Scaling Up. In C. Dede, J. Honan, & L. Peters (Eds.), Scaling Up Success: Lessons Learned from Technology-Based Educational Innovation. New York: Jossey-Bass.

Dede. C., & Honan, J. (in press). Synthesis of Research Themes on Scaling Up. In C. Dede, J. Honan, & L. Peters (Eds.), Scaling Up Success: Lessons Learned from Technology-Based Educational Innovation. New York: Jossey-Bass.

Dede, C., Nelson, B., Ketelhut, D., Clarke, J., & Bowman, C. (2004). Design-based research strategies for studying situated learning in a multi-user virtual environment. *Proceedings of the 2004 International Conference on Learning Sciences*, pp. 158-165. Mahweh, NJ: Lawrence Erlbaum.

Fishman, B. (in press). Adapting innovations to particular contexts of use: A collaborative framework. In C. Dede, J. Honan, & L. Peters, Eds, Scaling up success: Lessons learned from technology-based educational innovation. New York: Jossey-Bass.

Fishman, B., Marx, R., Best, S., & Tal, R. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, 19(6), 643–658.

Goldman, S.B. (in press). Designing for scalable educational improvement: Processes of inquiry in practice. In C. Dede, J. Honan, & L. Peters, Eds, Scaling up success: Lessons learned from technology-based educational innovation. New York: Jossey-Bass.

Klein, J. (2000). Corporate failure by design: Why organizations are built to fail. New York: Quorum Books.

Krajcik, J. S., Blumenfeld, P., Marx, R. W., & Soloway, E. (2000). Instructional, curricular, and technological supports for inquiry in science classrooms. In J. Minstrell & E.H.V. Zee (Eds.), Inquiring into inquiry learning and teaching in science (pp. 283–315). Washington, DC: American Association for the Advancement of Science.

Levy, F., & Murnane, R.J. (2004). The new division of labor: How computers are creating the next job market. Princeton, NJ: Princeton University Press.

Means, B, & Penuel, W.R. (in press). Scaling up technology-based educational innovations. In C. Dede, J. Honan, & L. Peters, Eds, Scaling up success: Lessons learned from technology-based educational innovation. New York: Jossey-Bass.

Midgley, G., Ed. (2003). Systems theories and modeling (Volume II of Systems thinking). New York: Sage.

Moje, E.B., Collazo, T., Carrillo, R., & Marx, R.W. (2001). "Maestro, what is 'quality'?": Language, literacy, and discourse in project-based science. *Journal of Research in Science Teaching*, *38*(4), 469–498.

Moore, G.A. (1999). Crossing the chasm (Rev. ed.). New York: Harper Perennial.

Morecraft, J.D.W., & Sterman, J.D., Eds. (2000). Modeling for learning organizations. New York: Productivity Press.

Moya, A., & Font, E., Eds. (2004). Evolution: From molecules to ecosystems. Oxford, England: Oxford University Press.

Nielsen, J. (1993). Usability engineering. San Francisco: Morgan Kaufmann.

Partnership for 21st Century Skills. (2003). Learning for the 21st Century. Washington, DC: Partnership for 21st Century Skills. Retrieved November 20, 2004 from http://www.21stcenturyskills.org/downloads/P21_Report.pdf

Russell, M., Bebell, D., & O'Dwyer, L. (2003) Use, support, and effect of instructional technology study: An overview of the USEIT study and the participating districts. Boston, MA: Technology and Assessment Study Collaborative. Retrieved August 29, 2004 from http://www.bc.edu/research/intasc/studies/USEIT/description.shtml

Schneider, R.M., & Krajcik, J.S. (2000, April). *The role of educative curriculum materials in reforming science education*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Senge, P.M., McCabe, N.H.C., Lucas, T., Kleiner, A., Dutton, J., & Smith, B. (2000). Schools that learn: A fifth discipline fieldbook for educators, parents, and everyone who cares about education. New York: Currency Press.

Snipes, J., Doolittle, F., & Herlihy, C. (2002). Foundations for success: Case studies of how urban school districts improve student achievement. Washington, DC: Council of the Great City Schools. Retrieved from http://www.cgcs.org/reports/Foundations.html on November 27, 2004.

Spillane, J. P. (2002). Local theories of teacher change: The pedagogy of district policies and programs. *Teachers College Record,* 104(3), 377–420.

Tyack, D., & Cuban, L. (1996). Tinkering toward utopia: A century of public school reform. Cambridge, MA: Harvard University Press.

Wiske, M.S., & Perkins, D. (in press). Dewey goes digital: Scaling up constructivist pedagogies and the promise of new technologies. In C. Dede, J. Honan, & L. Peters, Eds, Scaling up success: Lessons learned from technology-based educational innovation. New York: Jossey-Bass.

Table caption

Table 1: A Taxonomy of Potential Factors Influencing the Scalability of Effective Instructional Usage of Technology (from Russell, Bebell, and O'Dwyer, 2003)