

The Multiple Meanings of Scale: Implications for Researchers and Practitioners

Educational Researcher, 48(6), 369-377
First published online June 27, 2019

Richard Paquin Morel

Northwestern University, Evanston, IL

Cynthia Coburn

Northwestern University, Evanston, IL

Amy Koehler Catterson

Alder Graduate School of Education, Redwood City, CA

Jennifer Higgs

University of California, Davis, CA

The Multiple Meanings of Scale: Implications for Researchers and Practitioners

Richard Paquin Morel¹, Cynthia Coburn¹, Amy Koehler Catterson², and Jennifer Higgs³

Interest in the study of scale has grown over the past three decades, yet it still suffers from a lack of conceptual clarity. Despite attempts at conceptualizing scale, there is still wide diversity in how the term “scale” is used. These differences matter. They impact how scale is studied, the strategies used to achieve scale, and the lessons we can draw across studies of the scale of innovations. In this article, we argue that scale is a polysemic and dynamic phenomenon. There are multiple, legitimate definitions of scale, and such definitions can shift over time, depending on the goals and needs of reformers. Drawing upon an extensive review of the literature, we present a typology of scale comprising four predominant conceptualizations in the literature. We detail the conceptualizations and discuss the affordances and challenges of each. We conclude by offering implications of the polysemic, dynamic nature of scale for researchers and reformers. Presenting this typology, we aim to spark new conversations about scale and to help guide future scale research and practice.

Keywords: educational policy; educational reform; organization theory/change; policy; qualitative research

The challenge of bringing educational innovations¹ to scale has concerned researchers, policymakers, and reformers for over three decades. Innovations that succeed in small settings face challenges when expanded to more locations. Starting in the late 1990s with Elmore’s (1996) essay, scholars have increasingly turned their attention to the problem of scale. Initial efforts to study scale suffered from a lack of conceptual clarity (Coburn, 2003). However, a steady stream of research has improved the situation, offering clearer conceptualizations and empirical studies of scale (e.g., Clarke & Dede, 2009; Coburn, 2003; Peurach & Glazer, 2012; Schneider & McDonald, 2007a; 2007b; Stringfield & Datnow, 1998).

Still, a lack of clarity remains. Across studies, terms like “scale,” “scale up,” “scaling,” “spread,” and “at scale” are often used interchangeably but with quite different meanings. Such differences in definition are not trivial. A lack of shared understanding about the meaning of scale inhibits scholarly conversations on issues such as the identification of appropriate research designs and the development of new knowledge about relevant strategies for effectively fostering scale. The lack of shared understanding creates challenges for drawing lessons across studies to inform future research and efforts at achieving scale. With many resources and much effort dedicated to scaling educational

innovations, developing a clear conceptual foundation for scale is critical.

Drawing on an extensive review of the literature from a variety of fields, we argue that there are multiple, legitimate ways to conceptualize scale. While all conceptualizations share the requirement that innovations become widespread, we identified four distinct ideas among researchers and reformers about how people *should use* innovations for them to be considered “at scale.” We call these conceptualizations adoption, replication, adaptation, and reinvention. We show how each conception has distinct affordances and challenges for researchers and reformers, requiring different research designs and different strategies for fostering scale. We move beyond prior attempts to conceptualize scale that argue for one particular conceptualization (e.g., Coburn, 2003; McDonald, Keesler, Kauffman, & Schneider, 2006), to discuss when and under what conditions different conceptualizations might be appropriate. We also argue that scale is dynamic, rather than static, with the possibility that conceptualizations may shift over time. By acknowledging scale as a

¹Northwestern University, Evanston, IL

²Alder Graduate School of Education, Redwood City, CA

³University of California, Davis, CA

polysemic and dynamic phenomenon, we aim to stimulate more robust conversations about scale among researchers and reformers around a shared understanding of scale, opening new pathways for empirical investigations and a better understanding of strategies to foster scale.

Conceptualizing Scale: From Normative and Static to Polysemic and Dynamic

Early researchers often explicitly defined scale as the widespread use of a program (Adelman & Taylor, 1997; Fullan, 2000; Sloane, 2005; Stringfield & Datnow, 1998). This conceptualization focuses on quantitative indicators, such as the number of schools using an innovation, to determine when scale is achieved. Fullan (2000), for example, defines scale “as meeting two criteria: (1) the focus of reform is an entire system and/or (2) a minimum of 50 or so schools and some 20,000 or more students are involved” (p.8). Datnow, Hubbard, & Mehan (2002) define scale as “the transfer of an external reform model to multiple settings” (p. 2).² This classic definition of scale is common outside of education research, for example, in studies of diffusion of innovations (Rogers, 2003), community development (Myers, 1984), social movements (Briscoe, Gupta, & Anner, 2015; Strang & Soule, 1998), and organizational practices (Briscoe & Murphy, 2012; Fiss, Kennedy, & Davis, 2012).

Education scholars have criticized this conception of scale as incomplete, arguing that it leaves out core dimensions of interest to educators (Coburn, 2003; Elmore, 1996; McDonald et al., 2006). This critique led to at least two ways of conceptualizing scale. Coburn (2003) argued for a multi-dimensional conceptualization of scale, adding depth of implementation, shifts in norms and beliefs, and sustainability to widespread use. She argued that unidimensional conceptualizations focusing on numbers limit the scope of scale research, overlooking factors that lead to substantive changes important to reformers. McDonald and colleagues (2006) argued for a conceptualization that emphasizes student achievement in addition to widespread use. Challenging Coburn, they maintained that improving student achievement is the most critical outcome of scale and therefore fundamental to how scale is conceptualized.

Despite such efforts, there remain three critical problems in the literature on scale. First, multiple conceptualizations of scale have proliferated across studies, but this multiplicity is unacknowledged. It makes it difficult to assess the state of the field and draw insights across studies if the studies are talking about and measuring different things. Second, prior attempts to conceptualize scale, such as Coburn’s and McDonald and colleagues’, take a normative stance, arguing that one particular conceptualization is superior to others, without considering the possibility that there are multiple legitimate conceptualizations. This forestalls important conversations and empirical investigations about the comparative benefits and drawbacks of different approaches to scale.

Third, most conceptualizations of scale are static. They do not acknowledge that conceptualizations of scale may shift over time within the lifecycle of an innovation. Yet, empirical research on scale shows that it unfolds and transforms over time. Reformers may shift the meaning and goals of scale in response

to changing conditions in the environment and changing needs of reformers and users (Peurach, 2011). Even studies using a static conceptualization of scale document shifts in the meaning of scale as programs are implemented, such as Datnow and colleagues (2002). Acknowledging the dynamic nature of scale opens avenues for empirical studies of changing scale conceptualizations and the possibility of more guidance for reformers in developing their strategies to foster scale.

To address these problems, we develop a typology of scale that acknowledges the polysemic and dynamic conceptualizations in the literature. The typology emerges from research we conducted through a grant from the John D. and Catherine T. MacArthur Foundation. The purpose of the grant was to develop a conceptual framework on scale to guide their funding strategy related to Digital Media and Learning. We reviewed literature from multiple fields,³ analyzing how scholars conceptualized and analyzed scale. It soon became clear that scholars and reformers alike were using the term *scale* in a range of ways, that these different conceptualizations were useful analytically and practically, and that they could be flexible, shifting over the life course of the innovation.

A Typology of Scale

Any typology of scale needs to first bound the term. In the literature, researchers tend to conflate scale as a process and scale as an outcome. We define scale as the outcome or reformers’ desired end-state for how a large number of users engage with an innovation. How that end-state is conceptualized can vary. Scale as an outcome might look different depending upon whether reformers expect that people use the innovation in prescribed ways or encourage adaptation. Different conceptualizations likely require different strategies to achieve and different kinds of research designs to capture. Below, we describe four conceptualizations of scale that we derived from the literature: adoption, replication, adaptation, and reinvention. While the first three are common in education research and practice, we consider the fourth to be emergent. Stemming from research and practice in the digital media fields, it is starting to make in-roads in educational settings.

Adoption

Across fields, many researchers and reformers conceptualize scale as adoption. That is, they conceptualize scale as the widespread use of an innovation without explicitly conceptualizing the expected use of the innovation (Fullan, 2000; Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou, 2004; Jenkins, Ford, & Green, 2013; Myers, 1984; Sloane, 2005). For example, Fullan (2000), quoted above, defines scale quantitatively, as achieving a specific user-base. This conceptualization is common in the business literature, where scale is described in terms of market share (e.g., Halaburda & Oberholzer-Gee, 2014).

Despite criticisms of this conceptualization (e.g., Coburn 2003; McDonald et al., 2006), researchers have identified benefits to adoption. First, social movement scholars have documented how widespread adoption can build legitimacy for new or counter-normative ideas or practices, leading to substantive

changes in practices and beliefs (Strang & Soule, 1998). For example, Rao, Monin, & Durand (2003) show how French chefs adopted “nouvelle” styles of cooking. Even though instantiations of this new style varied widely, it ultimately changed the landscape of French cuisine. Briscoe and Safford (2008) document how the adoption of domestic partner benefits by *Fortune* 500 firms shifted the practice from controversial to accepted and expected.

Second, communication scholars have shown how adoption can foster “network effects” (DiMaggio & Garip, 2012; Easley & Kleinberg, 2010). Network effects occur when the benefits of adopting an innovation increase as more people adopt the innovation. This can produce cumulative advantages for innovations, weaving them into the fabric of everyday practice (DiMaggio & Garip, 2012). An example is email. When only a few individuals and organizations used email, the value of the network was relatively low. As email became ubiquitous, its value dramatically increased. Without widespread adoption, email has limited value.

There are also examples of the value of widespread adoption in education, despite variable and, at times, superficial implementation. The Carnegie Unit created a template for the division of instructional time that has become so widely used that it has defined the organization of time in schools (Tyack & Tobin, 1994). Kindergarten—began as a progressive reform in the 19th century—became a fundamental part of public schooling through widespread adoption (Cuban, 1993). Bloom’s Taxonomy moved from the periphery to mainstream through widespread adoption, becoming a standard for curriculum development and instruction in the K–12 system, even though schools use the taxonomy in a wide range of ways (Schneider, 2014). This suggests that adoption can have important consequences for educational practices, even when they are implemented superficially.

Replication

In educational settings, researchers and reformers often conceptualize scale as replication. For this conceptualization, an innovation is considered at scale if it is widespread, implemented with fidelity, and produces expected outcomes. Advocates of this conceptualization have argued that student achievement is the primary goal of education and therefore scale requires widespread impact (McDonald et al., 2006; Slavin & Madden, 2007). This approach to scale has become common in education research and offers a framework to researchers and reformers interested in student achievement as an outcome of scale for guiding their work.

Replication is conceptualized in two related ways. Some have conceptualized replication in terms of *outcomes*. McDonald and colleagues (2006), for example, argue that the first-order concern of scale is to “produce similarly positive effects in different settings and to help a greater number of students” (p. 16). This view is echoed by Baker (2007): “The goal of scaling up educational innovations is to produce robust, effective, replicable outcomes” (p. 37). In such cases, scholars often assume that well-designed innovations or interventions, implemented with fidelity, will reliably reproduce results in new settings.

Researchers and reformers have also conceptualized scale as the replicated *use* of an innovation to ensure the reproduction of outcomes. For example, Elias, Zins, Graczyk, and Weissbert

(2003) describe scale as “widely replicated procedures” (p. 303). Others describe replication as the reproduction of the capabilities required to engage in certain types of work (Peurach & Glazer, 2012; Winter & Szulanski, 2001). These capabilities include routines, practices, and skills that are needed to enact an innovation in a new location. Often implicit in this conceptualization is the assumption that replicated procedures lead to replicated outcomes.

Adaptation

Increasingly common in educational settings, some researchers and reformers conceptualize scale as adaptation. Scale is the widespread use of an innovation that is modified according to the needs of local users (Clarke & Dede, 2009; Fishman, 2005; Means & Penuel, 2005; Siskin, 2016; Wiske & Perkins, 2005). For example, Dede and Nelson (2005) define scale as “transferring and adapting [a] set of interrelated innovations to new contexts” (p. 111). Means and Penuel (2005) define scale as “adaptation by local actors to local contexts” (p. 177). Such modifications are bounded by the requirement to adhere to pre-defined “core principles” of the innovation.

This conceptualization has its roots in the concept of “mutual adaptation” described by Berman and McLaughlin (1976) in studies of the implementation of federal programs. They found that successful programs allowed local actors to make adaptations according to local needs and make adaptations to their local context to accommodate programs, a finding that has been confirmed in several studies (Berends, Bodilly, & Kirby, 2002; Bodilly, Keltner, Purnell, Reichardt, & Ikemoto, 1998; Cohen, Peurach, Glazer, Gates, & Goldin, 2014). Such modifications were often ad hoc and unplanned. Some researchers conceptualizing scale as adaptation have formalized this process by planning and designing for modifications based on local contexts (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003).

Like scale as replication, those who conceptualize scale as adaptation are often concerned with whether an innovation produces expected outcomes (Dede & Nelson, 2005). However, in contrast, they argue that local modifications enhance an innovation’s effectiveness (Cobb et al., 2003; McLaughlin, 1987; Means & Penuel, 2005). Local actors know their context and can use this knowledge to effectively adapt innovations. Local conditions cannot be “designed away,” but are key to successful outcomes. Indeed, advocates of “continuous improvement” research argue that it is through variations in design that we learn what works (Bryk, Gomez, & Grunow, 2011; Peurach, Glazer, & Winchell Lenhoff, 2016).

While supporters of adaptation reject notions of strict fidelity to program designs, many argue that innovations should maintain a core set of principles that bound local modifications. The maintenance of an innovation’s effectiveness can also be used to determine if local modifications are acceptable (e.g., Means & Penuel, 2005, p. 177). Indeed, some researchers speak of “lethal mutations” that violate core principles and potentially undermine effectiveness (Brown, 1991; Tatar et al., 2008). In this view, modifications are bound within an acceptable range to prevent the adaptations from deviating too far from the design’s core principles.

Reinvention

Of the scale conceptualizations presented here, reinvention is likely the least familiar in education. The concept derives in part from the field of digital media. The key idea is that innovations serve as a catalyst for further innovation (Bogers, Afuah, & Bastian, 2010; Morrison, Roberts, & von Hippel, 2000; von Hippel, 2007). Rather than reproducing or adapting an innovation, local actors build from it—“remix” it, in the language of digital media scholars—creating something new (Lessig, 2008). In contrast with scale as replication and adaptation, researchers and reformers conceptualizing scale as reinvention expect that innovations undergo radical transformations. What the innovation looks like, what it means, how it is used, what problems it solves, and what outcomes it produces all depend on the creative appropriation by local actors.

Scale as reinvention differs from the oft-documented phenomenon of educators transforming innovations as they implement them (Huberman, 1993; Little, 1990) and the “users-as-innovators” literature in organizational studies (von Hippel, 2007). These two literatures document how users recreate innovations to fit their needs in an ad hoc, idiosyncratic manner, viewing it as an unplanned consequence of implementation. By contrast, scale as reinvention involves intentional and systematic experimentation with an innovation.

The ethos of reinvention is salient in open-source production communities and the remix/meme culture associated with digital media (Benkler, 2006; Jenkins et al., 2013; Lessig, 2008; Santo, 2018; Weber, 2004). In open-source production platforms, like GitHub, people share innovations with the expectation that others will appropriate, modify, or totally recreate them. With memes, actors appropriate images and recreate them for comedic, social, or political purposes by altering text, context, or symbolism (Jenkins et al., 2013; Lessig, 2008; Varis & Blommaert, 2015).

Reinvention is not strictly a digital phenomenon. It is making in-roads in education through a variety of organizations seeking to catalyze local innovation. For example, the Institute of Play, an organization dedicated to promoting game-inspired learning in schools (Salen, Torres, Wolozin, Rufo-Tepper, & Shapiro, 2010), offers several “Design Packs” that allow educators to experiment with elements of their learning model. The packs present models of instructional practices, curriculum design, and school design, but encourage users to pick, choose, and re-envision those elements. The packs serve as a starting point for educators to engage in the local innovation reinvention requires. This, along with other projects that encourage educators to adopt a remixing perspective (Santo, 2018), show that reinvention, while still new, is a growing approach to scale in education.

This typology builds on existing conceptualizations of scale in three ways. First, while previous conceptualizations have taken a normative stance, we have argued, based on our reading of the literature, that scale comprises a range of legitimate conceptualizations. Both Coburn (2003) and McDonald and colleagues (2006) explicitly critique the conceptualization of scale as adoption. Neither acknowledges the potential benefits of adoption, such as building legitimacy or achieving network

effects. Second, it extends the potential range of outcomes of scale. For example, McDonald and colleagues’ conceptualization, which fits in the replication conceptualization of scale, focuses on a single, albeit quite important, outcome of scale: student achievement. However, reformers may seek additional outcomes as well, including shifting beliefs, catalyzing innovation, or co-producing approaches to teaching and learning that are context-bound to local educational settings. Third, it builds upon the dimensions of conceptualization articulated by Coburn. Depth of implementation, sustainability, and ownership will likely have continued relevance for conceptualizations like replication, adaptation, and reinvention. However, what these dimensions look like will likely vary for each conceptualization. In this article, we moved away from a normative discussion of scale, creating a typology that describes multiple possible conceptualizations, highlighting the potential affordances provided by each approach.

Scale Is Dynamic

Acknowledging the polysemic nature of scale opens the possibility that scale is also dynamic—that for the same innovation, reformers may pursue different outcomes at different points in its life cycle. Most researchers have studied scale through the lens of one conceptualization and assumed that this conceptualization remains fixed. Despite this, descriptive studies of scale have documented how innovations and the goals of reformers shift over time (Cohen et al., 2014; Datnow et al., 2002; Peurach, 2011). For example, Success for All developed an explicit strategy of expecting replication initially, before opening the program to local modifications (Datnow & Park, 2010; Peurach, 2011). However, existing studies do not make explicit how and why conceptualizations of scale may change, nor do they design research intentionally to capture scale as an outcome differently at different times in the life cycle.

Viewing scale as dynamic means that how reformers and researchers conceptualize scale can shift over time, depending on the needs of reformers. Reformers may desire widespread adoption to help build legitimacy for counter-normative innovations, but then shift to replication or adaptation to meet more specific desires for use and outcomes. In this approach, reformers attempt to build bottom-up acceptance of their ideas through widespread adoption, laying the groundwork for deeper instantiations.

Implications for Researchers

The acknowledgement of the polysemic and dynamic nature of scale has important implications for researchers. Each conceptualization of scale likely requires different research designs, sources of data, and indicators of success. Each conceptualization raises a unique set of empirical questions. Finally, each conceptualization poses somewhat different challenges for researchers.

Scale as *adoption* emphasizes getting innovations into the hands of many people or organizations. There is a long history of dissemination research, which is primarily concerned with documenting diffusion (Greenhalgh et al., 2004; Rogers, 2003). These studies often use archives of organizations to identify when practices were adopted and then apply network analysis,

spatial regression, or hazard models to capture the rate, extent, and factors that influence adoption (Briscoe & Murphy, 2012; Strang & Soule 1998; Wang & Soule, 2012).

Researchers may also attend to other potential outcomes of adoption, like network effects or legitimacy. To capture network effects, researchers can investigate how social networks influence adoption (Valente, 2005) or identify “threshold” points where the advantages of adopting an innovation outstrip the costs (DiMaggio & Garip, 2012). Researchers may consider indicators for when innovations become viewed as legitimate—that is, appropriate for use in educational settings (Colyvas & Jonsson, 2011). Studying these outcomes requires research designs that cover long time periods and frequently requires archival methods to document changes over time (Colyvas, 2007; Rao et al., 2003; Weber, Heinze, & DeSoucey, 2008). Colyvas (2007), for example, studied 14 years of university archives to document how the practice of patenting research findings went from marginal to mainstream.

In education, most discussions of research designs for studying scale have centered on *replication* (e.g., Schneider & McDonald, 2007a; 2007b). Some researchers studying scale as replication attend to *how* innovations are used, with a focus on the fidelity of implementation (Slavin & Madden, 2007). Researchers frequently use surveys or interviews to capture how faithfully educators adhere to prescriptions for use (Quint et al., 2013). Similarly, Peurach and Glazer (2012) advocate for studying the replication of organizational routines that reliably reproduce desired outcomes. Such studies require researchers to document the implementation of routines, which may require in-depth qualitative methods, such as participant observations or ethnographies.

Other replication research focuses primarily on outcomes, attempting to establish the effectiveness of an innovation and the generalizability of findings. These studies tend to use randomized control trials (RCTs) to ensure internal validity (McDonald et al., 2006) and may seek to strengthen the generalizability of results by conducting repeated RCTs in contrasting contexts (e.g., White, Kim, Kingston, & Foster, 2013). Some impact evaluations of innovations try to capture both effectiveness and fidelity, employing mixed methods designs (e.g., Quint et al., 2013).

While the scholarly discussion about how to effectively study *adaptation* is less well-developed, existing research offers insights. In contrast to replication, effectiveness in adaptation is not a generalizable causal relationship but a context-bound local relationship (Cobb et al., 2003). This requires attention to variations in implementation and their contribution to local effectiveness. Means and Penuel (2005) emphasize the importance of distinguishing “enactment variables,” which are important for local contexts, from “main effects,” which hold across contexts.

Continuous improvement research offers an alternative approach to studying adaptation. Advocates argue for partnerships between researchers and practitioners for engaging in iterative cycles to design, test, and modify innovations based on local needs and conditions (Bryk et al., 2011; Cobb et al., 2003). Practitioners and researchers collaboratively define local problems of practices, design or adapt interventions, and evaluate the effectiveness of the interventions. Findings from such inquiry-based studies are not expected to generalize to multiple settings.

Rather, the inquiry process may produce knowledge that others may adapt to their own local needs, through their own process of adaptation.

Due to a lack of studies on scale as reinvention in education, there are few examples to draw upon for guidance on matters of research design. A key criterion for scale as reinvention is widespread use of an innovation as a jumping-off point for further innovation. Since the meaning, purpose, and function of an innovation can change through reinvention, identifying and tracing its spread and scale are primary concerns. Researchers may also want to document the innovations that result from reinvention and the conditions that produced them. Santo (2018), for example, documented how educators repurposed an online curriculum promoting web literacy by participating in a learning network that promoted “remixing” materials. The initial curriculum served as a jumping-off point for further innovation by educators.

Researchers studying scale as reinvention may attend to factors that promote innovation within organizations by studying how educators engage with practices that are typically counter-normative in education settings. For example, the “working in the open” model adapts practices and values from the world of open-source digital production to education settings (Santos, 2018). Drawing from research exploring how social networks promoted innovation (Frank, Zhao, Penuel, Ellefson, & Porter, 2011; Moolenaar, Daly, & Slegers, 2011; Obstfeld, 2005; Reagans & McEvily, 2003; Ruef, 2002; Spillane, Kim, & Frank, 2012), scholars studying scale as reinvention may analyze the network structures that support reinvention within schools and evaluate strategies for generating such network structures (Santo, Ching, Pepler, & Hoadley, 2016).

Finally, acknowledging the polysemic and dynamic nature of scale makes possible new conversations and empirical questions about scale closed off by normative conceptions. Since scale is polysemic, researchers may investigate when and under what conditions conceptualizations of scale are appropriate and effective. Recent studies by Kim and colleagues (Kim et al., 2017; Quinn & Kim, 2017) are examples of such comparative work. Studying replication and adaptation models of scale for the same intervention, they found that adaptation was effective for experienced teachers, while replication was effective for novice teachers. Researchers might also make the dynamic nature of scale itself the subject of empirical investigation by studying processes that can lead to the shift in scale conceptions. How and why do conceptualizations of scale change? Under what conditions are shifts in conceptualizations effective? Which scale transitions are the most effective? Further, since scale is dynamic, appropriate research designs depend on the point of the life cycle of the innovation. For example, if a researcher were studying an innovation where reformers focused on scale as replication initially before transitioning to scale as adaptation, it might be appropriate to design a longitudinal study that charted that shift over time, focusing on different measures of implementation at different stages.

Implications for Reformers

For reformers, the polysemic nature of scale means that appropriate strategies for achieving scale depend how scale is

conceptualized. Each conceptualization makes distinct demands for how reformers design innovations, the supports needed for operating at scale, and expectations for the capacities of target users.

Reformers who want to encourage adoption are principally concerned with strategies to promote widespread diffusion. They may use branding and media campaigns to generate “pull,” or demand for innovation (Hagel, Brown, & Davison, 2012). To build legitimacy, reformers may consider targeting influential people in social networks who can facilitate the spread of innovations (Easley & Kleinberg, 2010; Obstfeld, 2005). They could also promote policies that support the adoption of innovations, as the Success for All Foundation did by lobbying for policies that promoted the use of evidence-based programs (Peurach, 2011).

Reformers interested in replication may expect substantial capacity on the part of reformers and users. They may need to create robust infrastructures to support and monitor implementation and to build the capacity of the reform organization itself (Peurach, 2011; Stringfield & Datnow, 1998). Teachers may have to learn new practices or detailed curricula; schools may need to be organized in different ways. In such cases, reformers likely need to develop strategies and support structures for building the capacity necessary to implement the innovation with fidelity. This may require elaborated designs that codify the roles, routines, and practices to guide implementation, along with trainers to provide professional development, mechanisms for quality control, and organizational designs to support that work (Cohen & Ball, 2007; Peurach, 2011; Slavin & Madden, 2007; Winter & Szulanski, 2001).

Reformers encouraging adaptation may also require substantial capacity from the users, as well as from themselves. Users are expected to make modifications to enhance an innovation’s effectiveness. This task likely requires a thorough understanding of the core principles of the innovation as well as local conditions so that users can make informed modifications. Reformers may need to help build the capacity of local users to engage meaningfully in adapting innovations and may also need to develop substantial understanding of local conditions themselves. Forms of adaptation, such as continuous improvement (Bryk et al., 2011) and design-based models (Cobb et al., 2003), also require infrastructure to support rapid data collection, analysis, and design iteration.

Adaptation also often involves substantial and ongoing collaboration between reformers and users, who may have different needs, expectations, and abilities. Such collaborations create unique challenges for reformers to manage. Scholarship on research-practice partnerships can offer insights on the strategies and structures that make collaboration effective (Coburn & Penuel, 2016; Coburn, Penuel, & Geil, 2013). For example, reformers and users may need to carefully consider their differences in language use, clarify roles and responsibilities, and understand the respective affordances and constraints of their organizational settings (Farrell, Harrison, and Coburn, 2019; Harrison, Davidson, & Farrell, 2017).

Designs for innovations that encourage reinvention depend heavily on the norms and capacities of potential users to engage

in systematic local innovation (Benkler, 2006; Jenkins et al., 2013). Reformers encouraging reinvention in educational contexts cannot rely on such preexisting norms and capacities to recreate innovations in disciplined and systematic ways. Prior research suggests a latent ability among teachers to transform innovations (Huberman, 1993; Little, 1990). Reformers may need to provide supports for teachers, so teachers can embrace that role move toward systematic iteration at the local level. This may require more than developing individual competencies; it may also require strategies to develop local conditions and norms that support innovation. Teachers and administrators may need to explicitly adopt new role identities consonant with a remix/reinvention ethos (Salen et al., 2010).

Conclusion

By acknowledging the polysemic and dynamic nature of scale, we aim to strengthen research on scale by encouraging greater attention to how scale is defined and the identification of appropriate research designs to capture it. We hope to open new areas of empirical exploration that have previously been overlooked, such as comparative explorations of different conceptualizations, the conditions which make one conceptualization more appropriate than others, and the dynamics of shifting conceptualizations of scale. We also aim to provide reformers interested in scaling innovations guidance on identifying appropriate and effective strategies for scale, as well as on explicitly planning for and addressing how and why their conceptualizations of scale shift depending on their goals or needs.

Acknowledging the polysemic and dynamic nature of scale is an important step toward greater conceptual clarity. If we do not identify and acknowledge these different conceptualizations, we risk research designs that are ill-equipped to capture the key elements of the process of scaling and the outcome of scale. We lose the opportunity for investigating, discussing, and debating effective strategies for promoting scaling and achieving scale given different goals. We need greater conceptual clarity if we are to move the field forward as both researchers and reformers. This article is meant as a step in this direction.

NOTES

This article was supported by funding from the John D. and Catherine T. MacArthur Foundation. We are grateful to Foundation staff and grantees of the Digital Media and Learning Initiative for providing feedback and pushing our thinking as we developed the analysis presented here. Richard Paquin Morel’s work on this project was supported by the Multidisciplinary Program in Education Sciences, Institute for Education Science, U.S. Department of Education (Grant # R305B140042 to Northwestern University).

¹Throughout this article, we use the term “reformers” to include anyone involved in design, development, and dissemination of an innovation. This can include designers, educational leaders, and researchers who develop and spread innovations. By “researchers” we refer to scholars studying innovations at scale. Sometimes, researchers both develop innovations and study their spread. We use the generic term “innovation” to capture educational ideas, practices, tools, interventions, and programs. We deliberately keep the term broad to including a wide range of possible approaches to fostering change. The type of innovation—whether

an idea, practice, or whole-school reform model—may influence the desired end-state of scale. While much of the early research on scale in education focused on whole-school reform models (e.g., Datnow et al., 2002; Elmore, 1996), this has expanded to a range of programs and interventions (e.g., Kim and colleagues' [2017] study of summer reading programs). It is possible to scale tools, ideas, and practices in addition to programs. The nature of the innovation, along with the typology of scale we present here, are two elements of a more comprehensive framework for the study and practice of scale that we will present in a future publication.

²It is important to note that some researchers who explicitly defined scale as adoption do study how innovations are implemented and used. For example, Stringfield and Datnow (1998), while explicitly defining scale in terms of adoption, documented how teachers implemented and adapted programs. Indeed, early studies of scale that attended to implementation provided evidence of the importance of implementation that informed Coburn's (2002) and McDonald et al.'s (2006) reconceptualizations of scale.

³We conducted an extensive review of literature on scale in education, organizational theory, social networks, social movements, and digital media studies. Using Google Scholar, we identified over 150 articles on spread, scale, and related concepts like diffusion. We distinguished between empirical and conceptual articles, using empirical articles as the core of our review. For each article, our team identified how the authors defined scale and conceptualized its process, and the factors that influenced the process. We then classified studies by the definition of scale. Synthesizing across studies, we derived four primary conceptualizations of scale, forming our typology.

REFERENCES

- Adelman, H. S., & Taylor, L. (1997). Toward a scale-up model for replicating new approaches to schooling. *Journal of Educational and Psychological Consultation*, 8(2), 197–230.
- Baker, E. L. (2007). Principles for scaling up: Choosing, measuring effects, and promoting the widespread use of educational innovation. In B. L. Schneider & S. K. McDonald (Eds.), *Scale-up in education volume 2: Issues in practice* (pp. 37–54). Lanham, MD: Rowman & Littlefield.
- Benkler, Y. (2006). Peer production and sharing. In *The wealth of networks: How social production transforms markets and freedom* (pp. 59–81). New Haven, CT: Yale University Press.
- Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American schools after a decade*. Retrieved from: http://www.rand.org/pubs/monograph_reports/MR1498.html.
- Berman, P., & McLaughlin, M. W. (1977). *Federal programs supporting educational change: Vol. II: Factors affecting implementation and continuation*. Santa Monica, CA: RAND Corporation.
- Bodilly, S. J., Keltner, B. R., Purnell, S. W., Reichardt, R., & Ikemoto, G. I. (1998). *Lessons from new American schools' scale-up phase: Prospects for bringing designs to multiple schools*. Retrieved from: http://www.rand.org/pubs/monograph_reports/MR942.html.
- Bogers, M., Afuah, A., & Bastian, B. (2010). Users as innovators: A review, critique, and future research directions. *Journal of Management*, 36(4), 857–875.
- Briscoe, F., Gupta, A., & Anner, M. S. (2015). Social activism and practice diffusion. *Administrative Science Quarterly*, 60(2), 300–332.
- Briscoe, F., & Murphy, C. (2012). Sleight of hand? Practice opacity, third-party responses, and the interorganizational diffusion of controversial practices. *Administrative Science Quarterly*, 57(4), 553–584.
- Briscoe, F., & Safford, S. (2008). The Nixon-in-China effect: Activism, imitation, and the institutionalization of contentious practices. *Administrative Science Quarterly*, 53(3), 460–491.
- Brown, A. (1991). Design experiments: Theoretical and methodological challenges in evaluating complex interventions in classroom settings. *Journal of the Learning Sciences*, 2, 141–178.
- Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. T. Hallinan (Ed.), *Frontiers in sociology of education* (Vol. 1, pp. 127–162). New York, NY: Springer.
- Clarke, J., & Dede, C. (2009). Design for scalability: A case study of the River City curriculum. *Journal of Science Education and Technology*, 18(4), 353–365.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9–13.
- Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3–12.
- Coburn, C. E., & Penuel, W. R. (2016). Research-Practice Partnerships in Education: Outcomes, Dynamics, and Open Questions. *Educational Researcher*, 1–7.
- Coburn, C. E., Penuel, W. R., & Geil, K. (2013). *Research-practice partnerships at the district level: A new strategy for leveraging research for educational improvement*. New York: William T. Grant Foundation.
- Cohen, D. K., & Ball, D. L. (2007). Educational innovation and the problem of scale. In B. Schneider & S.-K. McDonald (Eds.), *Scale-up education, volume 1: Ideas in principle* (pp. 19–36). Lanham, MD: Rowman & Littlefield.
- Cohen, D. K., Peurach, D. J., Glazer, J. L., Gates, K. E., & Goldin, S. (2014). *Improvement by design: The promise of better schools*. Chicago, IL: University of Chicago Press.
- Colyvas, J. A. (2007). From divergent meanings to common practices: The early institutionalization of technology transfer in the life sciences at Stanford University. *Research Policy*, 36(4), 456–476.
- Colyvas, J. A., & Jonsson, S. (2011). Ubiquity and legitimacy: Disentangling diffusion and institutionalization. *Sociological Theory*, 29(1), 27–53.
- Cuban, L. (1993). *How teachers taught* (2nd ed.). New York, NY: Teachers College Press.
- Datnow, A., & Park, V. (2010). Success for all: Using tools to transport research-based practices to the classroom. In C. E. Coburn & M. K. Stein (Eds.), *Research and practice in education: Building alliances, bridging the divide* (pp. 77–91). Lanham, MD: Rowman & Littlefield.
- Datnow, A., Hubbard, L., & Mehan, H. (2002). *Extending educational reform: From one school to many*. New York, NY: RoutledgeFalmer.
- Dede, C., & Nelson, R. (2005). Technology as proteus: Digital infrastructures that empower scaling up. In C. Dede, J. P. Honan, & L. C. Peters (Eds.), *Scaling up success: Lessons from technology-based educational improvement* (pp. 110–132). San Francisco: Jossey-Bass.
- DiMaggio, P., & Garip, F. (2012). Network effects and social inequality. *Annual Review of Sociology*, 38(1), 93–118.
- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge, UK: Cambridge University Press.
- Elias, M. J., Zins, J. E., Graczyk, P. A., & Weissberg, R. P. (2003). Implementation, sustainability, and scaling up of social-emotional and academic innovations in public schools. *School Psychology Review*, 32(3), 303–319.
- Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, 66(1), 1–27.

- Farrell, C. C., Harrison, C., & Coburn, C. E. (2019). "What the hell is this, and who the hell are you?" Role and identity negotiation in research-practice partnerships. *AERA Open*, 5(2), 1–13.
- Fishman, B. J. (2005). Adapting innovations to particular contexts of use: A collaborative framework. In C. Dede, J. P. Honan, & L. C. Peters (Eds.), *Scaling up success: Lessons from technology-based educational improvement* (pp. 48–66). San Francisco: Jossey-Bass.
- Fiss, P. C., Kennedy, M. T., & Davis, G. F. (2012). How golden parachutes unfolded: Diffusion and variation of a controversial practice. *Organization Science*, 23(4), 1077–1099.
- Frank, K. A., Zhao, Y., Penuel, W. R., Ellefson, N., & Porter, S. (2011). Focus, fiddle, and friends: Experiences that transform knowledge for the implementation of innovations. *Sociology of Education*, 84(2), 137–156.
- Fullan, M. (2000). The Return of Large-scale Reform. *Journal of Educational Change*, 1, 5–28.
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *Milbank Quarterly*, 82(4), 581–629.
- Hagel III, J., Brown, J. S., & Davison, L. (2012). *The power of pull: How small moves, smartly made, can set big things in motion*. Philadelphia, PA: Basic Books.
- Halaburda, H., & Oberholzer-Gee, F. (2014). The limits of scale. *Harvard Business Review*, 92(4), 94–99.
- Harrison, C., Davidson, K., & Farrell, C. (2017). Building productive relationships: District leaders' advice to researchers. *International Journal of Education Policy and Leadership*, 12(4).
- Huberman, A. M. (1993). *The lives of teachers*. New York, NY: Teachers College Press.
- Jenkins, H., Ford, S., & Green, J. (2013). *Spreadable media: Creating value and meaning in a networked culture*. New York, NY: New York University Press.
- Kim, J. S., Burkhauser, M. A., Quinn, D. M., Guryan, J., Kingston, H. C., & Aleman, K. (2017). Effectiveness of structured teacher adaptations to an evidence-based summer literacy program. *Reading Research Quarterly*, 52(4), 443–467.
- Lessig, L. (2008). *Remix: Making art and commerce thrive in a hybrid economy*. New York, NY: Penguin.
- Little, J. W. (1990). The mentor phenomenon and the social organization of teaching. *Review of Research in Education*, 16, 297–351.
- Means, B., & Penuel, W. R. (2005). Scaling up technology-based educational innovations. In C. Dede, J. P. Honan, & L. C. Peters (Eds.), *Scaling up success: Lessons from technology-based educational improvement* (pp. 176–197). San Francisco: Jossey-Bass.
- McDonald, S. K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, 35(3), 15–24.
- McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9(2), 171–178.
- Moolenaar, N. M., Daly, A. J., & Slegers, P. J. (2011). Ties with potential: Social network structure and innovative climate in Dutch schools. *Teachers College Record*, 113(9), 1983–2017.
- Morrison, P. D., Roberts, J. H., & von Hippel, E. (2000). Determinants of user innovation and innovation sharing in a local market. *Management Science*, 46(12), 1513–1527.
- Myers, R.G. (1984, October). *Going to scale*. Paper presented at the Second Inter-Agency Meeting on Community-based Child Development, New York, NY.
- Obstfeld, D. (2005). Social networks, the tertius iungens orientation, and involvement in innovation. *Administrative Science Quarterly*, 50(1), 100–130.
- Peurach, D. J. (2011). *Seeing complexity in public education: Problems, possibilities, and success for all*. New York, NY: Oxford University Press.
- Peurach, D. J., & Glazer, J. L. (2012). Reconsidering replication: New perspectives on large-scale school improvement. *Journal of Educational Change*, 13(2), 155–190.
- Peurach, D. J., Glazer, J. L., & Winchell Lenhoff, S. (2016). The developmental evaluation of school improvement networks. *Educational Policy*, 30(4), 606–648.
- Quinn, D. M., & Kim, J. S. (2017). Scaffolding fidelity and adaptation in educational program implementation: Experimental evidence from a literacy intervention. *American Educational Research Journal*, 54(6), 1187–1220.
- Quint, J. C., Balu, R., DeLaurentis, M., Rappaport, S., Smith, T. J., & Zhu, P. (2013). The success for all model of school reform: Early findings from the investing in innovation (i3) scale-up. Retrieved from <http://www.mdrc.org/publication/success-all-model-school-reform>.
- Rao, H., Monin, P., & Durand, R. (2003). Institutional change in toque ville: Nouvelle cuisine as an identity movement in French gastronomy. *American Journal of Sociology*, 108(4), 795–843.
- Reagans, R., & McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*, 48(2), 240.
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.
- Ruef, M. (2002). Strong ties, weak ties and islands: Structural and cultural predictors of organizational innovation. *Industrial and Corporate Change*, 11(3), 427–449.
- Salen, K., Torres, R., Wolozin, L., Rufo-Tepper, R., & Shapiro, A. (2010). *Quest to learn: developing the school for digital kids*. Cambridge, MA: The MIT Press.
- Santo, R. (2018). Open source culture as inspiration for designs of educator learning networks. In S. A. Yoon & K. J. Baker-Doyle (Eds.), *Networked by design: Interventions for teachers to develop social capital* (pp. 149–169). New York, NY: Routledge.
- Santo, R., Ching, D., Peppler, K. A., & Hoadley, C. M. (2016, April). *Working in the open: Lessons from open source software on building ecosystems that promote educational innovation*. Paper presented at the annual American Educational Research Association conference, Washington, DC.
- Schneider, B., & McDonald, S.K. (Eds.). (2007a). *Scale up in education, volume 1: Ideas in principle*. Lanham, MD: Rowman & Littlefield.
- Schneider, B. & McDonald, S.K. (Eds.). (2007b). *Scale up in education, volume 2: Issues in practice*. Lanham, MD: Rowman & Littlefield.
- Schneider, J. (2014). *From the ivory tower to the schoolhouse: How scholarship becomes common knowledge in education*. Cambridge, MA: Harvard Education Press.
- Sisken, L. S. (2016). Mutual adaptation in action. *Teachers College Record*, 118(13), 1–16.
- Slavin, R. E., & N. A. Madden (2007). Scaling up success for all: The first sixteen years. In B. Schneider & S.-K. McDonald (Eds.), *Scale-up in education, volume 2: Issues in practice*. Lanham, MD: Rowman & Littlefield.
- Sloane, F. C. (2005). The scaling of reading interventions: Building multilevel insight. *Reading Research Quarterly*, 40(3), 361–366.
- Spillane, J. P., Kim, C. M., & Frank, K. A. (2012). Instructional advice and information providing and receiving behavior in elementary schools: Exploring tie formation as a building block in social capital development. *American Educational Research Journal*, 49(6), 1112–1145.
- Strang, D., & Soule, S. A. (1998). Diffusion in organizations and social movements: From hybrid corn to poison pills. *Annual Review of Sociology*, 265–290.

- Stringfield, S., & Datnow, A. (1998). Introduction: Scaling up school restructuring designs in urban schools. *Education and Urban Society, 30*(3), 269–276.
- Tatar, D., Roschelle, J., Knudsen, J., Shechtman, N., Kaput, J., & Hopkins, B. (2008). Scaling up innovative technology-based mathematics. *The Journal of the Learning Sciences, 17*, 248–286.
- Tyack, D., & Tobin, W. (1994). The “grammar” of schooling: Why has it been so hard to change? *American Educational Research Journal, 31*(3), 453–479.
- Valente, T. W. (2005). Network models and methods for studying the diffusion of innovations. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and Methods in Social Network Analysis*. New York, NY: Cambridge University Press.
- Varis, P., & Blommaert, J. (2015). Conviviality and collectives on social media: Virality, memes and new social structures. *Multilingual Margins, 2*(1), 31–45.
- von Hippel, E. (2007). Horizontal innovation networks—by and for users. *Industrial and Corporate Change, 16*(2), 293–315.
- Wang, D. J., & Soule, S. A. (2012). Social movement organizational collaboration: Networks of learning and the diffusion of protest tactics, 1960–1995. *American Journal of Sociology, 117*(6), 1674–1722.
- Weber, K., Heinze, K. L., & DeSoucey, M. (2008). Forage for thought: Mobilizing codes in the movement for grass-fed meat and dairy products. *Administrative Science Quarterly, 53*(3), 529–567.
- Weber, S. (2004). *The success of open source*. Cambridge, MA: Harvard University Press.
- White, T. G., Kim, J. S., Kingston, H. C., & Foster, L. (2013). Replicating the effects of a teacher-scaffolded voluntary summer reading program: The role of poverty. *Reading Research Quarterly, 49*, 5–30.
- Winter, S. G., & Szulanski, G. (2001). Replication as strategy. *Organization Science, 12*(6), 730–743.
- Wiske, M. S., & Perkins, D. (2005). Dewey goes digital: Scaling up constructivist pedagogies and the promise of new technologies. In C. Dede, J. P. Honan, & L. C. Peters (Eds.), *Scaling up success:*

Lessons from technology-based educational improvement (pp. 27–47). San Francisco: Jossey-Bass.

AUTHORS

RICHARD PAQUIN MOREL, MA, is a doctoral candidate in human development and social policy at Northwestern University, 2120 Campus Drive, Evanston, IL 60208; richard.morel@u.northwestern.edu. His research focuses on how social networks and social movements shape the development and implementation of educational policies.

CYNTHIA COBURN, PhD, is a professor at Northwestern University, 2120 Campus Drive, Evanston, IL; cynthia.coburn@northwestern.edu. Her research focuses on policy making and policy implementation in urban schools.

AMY KOEHLER CATTERSON, PhD, is a researcher and faculty member specializing in literacy instruction at Alder Graduate School of Education, 2946 Broadway St., Suite B, Redwood City, CA 94602; akcatterson@aldergse.org. Her current research focuses on the co-teaching practices of mentor teachers and their apprentice teacher candidates, or residents, and the relationship between co-teaching and K–12 student learning in teacher residency programs.

JENNIFER HIGGS, PhD, is an assistant professor at the University of California, Davis, 1 Shields Avenue, Davis, CA 95616; jmhiggs@ucdavis.edu. Her research focuses on uses of digital tools that support equitable learning and teaching in and across school and non-school contexts.

Manuscript received October 27, 2016

Revisions received September 2, 2017;

December 3, 2018; April 2, 2019

Accepted May 27, 2019