J Educ Change (2016) 17:529–537
DOI 10.1007/s10833-016-9290-8

"Getting to scale..." it seemed like a good idea at the time

Richard F. Elmore[1]

Published online: 31 October 2016

It all seemed quite straightforward back then. American schools were in the early stages of responding to state and national policies designed to improve measured student performance, which was the working definition of learning at the time. Educators seemed cautiously positive, at least about the national attention they were receiving, if not the content of the policies themselves. American academics interested in something called "policy implementation," a relatively well-developed research specialty, were interested in education reform policy as a special case of the broader issue of how public institutions respond to large-scale policy initiatives. And local, state, and federal officials were uneasy enough about the uncertain consequences of the changes they were living with to be curious about what research could tell them. With this as context, in 1995, I wrote an article for the *Harvard Educational Review* suggesting that the current period of educational reform might be an opportune time to revisit the perennial question of how promising educational practices might be developed and adopted "at scale."

Just writing that last sentence, at this moment in time, fills me with chagrin at my unthinkable presumptuousness and naïveté. The article was written by someone who, at that time, was a strong, if slightly fuzzy, believer in "policy-driven reform"—no more. It was written 5 years before I embarked on research and clinical practice that would take me into close observation of instructional practice in some 4000 classrooms, in nearly five hundred schools, in five countries. It was written 15 years before I embarked on a second clinical practice designed to try to assist some 100-plus low-performing American schools to become simpler, more focused, more coherent, more instructionally flexible learning environments. It was written, in other words, by someone—a person who now seems very unfamiliar to

✉ Richard F. Elmore
richard_elmore@gse.harvard.edu

[1] Harvard Graduate School of Education, Harvard University, Cambridge, MA, USA

me—who believed in what seems to me now an irresponsibly simplistic and schematic view of human learning and development.

I like to warn the people I work with in schools that I am a recovering political scientist. I did an entire program of doctoral coursework in public law and political science before I switched to the Harvard doctoral program in education policy. I studied political economy and public policy, in addition to education, at Harvard. I taught in a public policy and management school for 11 years before I moved to an education school. As a result of this background I am hypersensitive to the nuances of power and influence in the settings in which I work. I carry the residue of that training in my intellectual subconscious. Among the many voices I carry with me from that period of my life are three: A grizzled Massachusetts politician, Fred Salvucci, was famous for saying, "Culture eats policy for breakfast." Thomas Schelling, a brilliant economic theorist and Nobel Prize winner with whom I studied, once said, "there is a vast difference between knowing the right thing to do, and actually doing the right thing." And a treasured consulting colleague and teacher, Donald Schön, once said, "when you're the only one in the room who knows the right answer, you should stop talking and listen." I wish I had paid more attention to these voices before I wrote that article in the *Harvard Educational Review*.

I currently live in a world in which I routinely watch well-intentioned, highly-motivated educators—teachers and administrators—talk obsessively about "best practice" as if it were a kind of super-hero jumpsuit you could slip on before you step into your formal role as "change agent." I am routinely asked by well-meaning system-level leaders to talk to groups of more than 100 teachers and administrators about deeply complex practices of leadership and instructional practice that can only be learned through deep, daily immersion in guided practice. When I demur, they say things like, "we only have three professional development days this year, and we would really like to make the most of them with challenging content." I despair. I routinely observe classrooms in which the artifacts of "high-level, challenging content" are prominently displayed on the walls and in the materials distributed to students, but it is clear to me that the teachers in those classrooms have never experienced themselves the kind of learning they are asking students to engage in. Can you "teach" people how to learn in ways you yourself have never experienced? I think not. I have come to call this casual reference to "high level content" "knocking the corners off the grand piano to get it through the classroom door." I routinely work with schools that are asked to operate in systems that have completely, blatantly dysfunctional administrative structures, clotted with multiple levels and cross-functional relationships that, on their face, do not, cannot, and never will have a positive impact on learning. The main function of people who work in these organizations seems to be telling other people what to do—usually for things they themselves do not know how to do. When I ask about these dysfunctions, I am greeted with a kind of standard shoulder-shrugging response that can only be read as, "tell me something I don't know." Learned helplessness in the face of this dysfunction is the disease of a dying institution. Yet there seems to be endless optimism, at least among the committed reformers, and professional policy experts, that somehow we can make things better by "implementing" something

called "best practices" "at scale." Each of these terms, in my view embodies deep and profound misconceptions about how human beings learn, develop, adapt, and change.

First, "implementation" is something you do when you already know what to do; "learning" is something you do when you don't yet know what to do. The casual way policy-focused people use the term obscures this critical distinction. The knowledge of what to do has to reside not in the mind of some distant policy wonk or academic, but in the deep muscle-memory of the actual doer. When we are asking teachers and school leaders to do things they don't (yet) know how to do, we are not asking them to "implement" something, we are asking them to learn, think, and form their identities in different ways. We are, in short, asking them to be different people. When we expect them to "implement" when they don't know what they are doing, we are asking to pretend to be people they are not, a skill that regrettably is highly valued in hierarchical organizations. This is mostly what I see in classrooms undergoing "reform." Learning is a profoundly developmental practice; implementation is a technical practice. There is a growing science of learning (of which most educators seem blissfully unaware) that tells us how complex and multi-layered development is. Judging the success of reform policies by focusing on implementation is like issuing Michelin stars to restaurants based on how well they follow the recipe on the outside of the box.

Second, the idea of "best practices" is antithetical to developmental models of learning. Professions do indeed have practices that they expect their members to use and they evaluate themselves according to how well those practices work. But in real professions, of which education is not one, those practices come at the end of a long causal chain of learning and cultural socialization that creates foundational knowledge, dispositions toward the acquisition of new knowledge, and formal and informal institutions that stand apart from the workplace and that reinforce the culture that produces the practice. Real professions select their members based on mastery of knowledge, not based on bureaucratic and institutional procedures. Educators, because of the weakness of professional culture and autonomy, tend to treat "best practices" as tips and tricks that can be readily assimilated by reading the right books, or hiring the right consultants, rather than by investing seriously in developing the cultural and institutional infrastructures of professional practice. Trying to graft a professional culture onto an essentially bureaucratic, heavily institutionalized culture is a thankless task. The things that are easily "scaleable" in this environment are the things that require the least depth of preparation and practice.

Third, if we have learned anything from 25 to 30 years of attempts to "reform" education it is that every effort at reform is heavily influenced by the contexts, micro and macro, in which it exists. In the face of this understanding, the idea of "scale" seems either very superficial or downright wrong. In the era of policy-driven reforms we have struggled with the fact of how to adjust our thinking to accommodate the kinds of social, economic, cultural, and institutional differences that influence how learning occurs. Policy-driven reform has made uniformity the rule, and diversity a suspect and problematic exception. Advocacy groups speak of solutions as entitlements that must be present in every school and classroom in order

for "all students" to receive "equal" treatment. Policy makers speak and act as if variability in practice and outcomes is the result of subversive, self-interested motives, or down-right bone headedness. Major consulting firms and armies of education professional development providers attach themselves to simple, context-free formulas that can be easily painted over the top of major contextual differences. What sticks are the formulas, not the ideas or understanding that formed them, if they ever existed in the first place. Policy experts treat international metrics, and the measurements and the constructs behind them, as if they represent some universal set of cultural and social attributes that have equal value in every society and culture. The drive for "scale," in other words, is a drive for a kind of uniformity that makes the world more intelligible to people who are uncomfortable with complexity. Scale is an artifact of policy thinking. It's not a bad construct if you know that it is just that—a construct that suppresses complexity and relieves that sense of queasiness that afflicts the orderly mind in a complex world.

We are, in other words, pushing the limits of our ability to influence human improvement through the vehicle of education by using the constructs associated with policy-driven reform. Implementation and scale are two such constructs. It is time to move on. But to what?

The articles in this issue have much to contribute to our understanding of what the next stage of work might look like. First, there is an underlying theme about the instability and fickleness of policy as a basis for nurturing the growth on learning in society. One of the first things I used to teach my students when I still taught policy courses was the importance of the "issue-attention cycle" in political systems. Policy happens when there is a convergence of a problem articulated by societal interests, with a set of "solutions" in the form of policy proposals, with a window of opportunity in which policymakers see the pairing of the problem with the solution will produce political credit. Often the fit between the problem and the solution is imperfect. Often the solution has little or nothing to do with the "carrying capacity" of the institutions to which it is addressed. Usually the political credit accrues to those who initiate policy, not to those who have to deal with its downstream effects. The fact that policy makers live in a world in which there are multiple opportunities to connect solutions to problems for credit means that their attention shifts from one problem to another based on what generates political credit, not based on the "success" of the last policy. No one ever lost an election based on some bureaucrat's inability to implement an idea that a policy maker thought made sense at the time.

So the relationship between policy and practice can be, and usually is, fractious and unreliable. By my count only two of the cases in this volume succeeded in creating a stable and supportive policy environment that allowed for sustained and enlightened learning over an extended period of time—Ontario and Long Beach. We don't know much from the articles about how this happened, but the circumstantial evidence suggests there was enough depth and stability in the professional culture of these two systems to correct for and adjust to shifts in the political winds. Also, it appears that both systems were clever enough to offer enough new content in the actual reform agenda to continue to generate political credit for policy makers.

Beyond these two cases, there are a number of examples in which the relationship between policy and practice was predictably much choppier. In the Mexico case, the Learning Communities Program adjusted to instability in the policy environment by creating a powerful alternative model of social organization, based on networked relationships among communities and schools, occasionally supported by sympathetic allies in the state and national government. In other words, it has survived by creating its own parallel culture, more powerful in its own way than the ambient culture of the state. The Columbian case ends with the waning of governmental support for *Escuela Nueva*, which apparently has little to do with the movement's successes and more to do with shifts in political interests. The Indian and South African literacy cases are, at least from the U.S. perspective, much more predictable—gradual erosion of public and private attention, having less to do with the value of the programs than with lack of institutional capacity to sustain the work over time.

One logical place to go from these findings, if you believe in policy-driven reform, is to ask, "what conditions can policy makers create that sustain and expand interesting and powerful practices of whatever type over time?" I will take a contrarian view. I believe that policy makers are rational actors, that they respond to the incentives that face them, not to idealistic visions of the possible, and that relying on policy makers to think in long-term, generational terms is a lost cause. It is much more plausible to suggest that policy makers will be supportive of things *that already exist and are demonstrably successful in ways that generate political credit*. Acknowledge that political incentives are inherently unstable, fickle, and often counter-productive to sustained development of strong ideas in practice, and create a culture of innovation that capitalizes on support when it exists and avoids opposition and neglect when it occurs. If you hear Machiavelli in this last sentence you are absolutely correct.

This realist view of the role of policy in sustained development of learning leaves unanswered what we should use as a basis for action, if not policy. Let me be bold here: People interested in pushing against the frontiers of knowledge and policy around human development and learning should learn to rely less on governmental institutions and more on the development of strong theories and their enactment. "Scale" for its own sake is less important than demonstrating that powerful ideas can work in diverse environments and creating powerful networks that are capable of operating with or without the cover of public authority.

This change requires a shift in mindset, which has some of the following characteristics:

- From basing new learning designs on received ideas that are "feasible" in existing institutions, to basing learning designs on the theory and science of how humans learn
- From focusing on "universal" prescriptions for organization and practice to the processes required to adjust powerful ideas to diverse contexts
- From universal prescriptions for learning to multiple, diverse, and promising adaptations to diverse populations

- From complex, technical expressions of expert knowledge to simple, transparent expressions, accessible to adults and children alike
- From a focus on producing predictable effects across many settings, to a focus on expecting to be surprised by effects and patterns that emerge from divergent thinking and practices.

These are the characteristic markers of what might be called a "learning system" rather than an "education system." They are much more aligned to a future in which learning will migrate out of established educational institutions and into the broader society. They break the lock of institutionalized thinking about the capabilities of learners, and they introduce the possibility that the way we currently think about learning is, in powerful ways, scripted by the institutional interests of the education sector, when they should be influenced by emerging knowledge from the sciences of learning. They encourage divergent thinking. They think about "scale" and "spread" in terms of the development of theories and practices of learning through human interaction and the creation of culture, rather than adapting ideas and evidence to the "realities" of existing institutional cultures.

I have been spending a good deal of time over the past couple of years in "outlier" powerful learning environments.[1] I have noticed in my interactions with adults and young people in these settings a few patterns that they have in common, despite their profound differences in goals and practice. The first of these I would call *the principle of transparency*. This principle states that every act or set of ideas that is central to the practice of a learning environment should be open and accessible to everyone in that environment. Stated differently, our theories and practices should be known and understood by adults and young people alike, they should be formulated in such a way that everyone can articulate them and explain their meaning in the context of practice, and they should be subject to continuous scrutiny and revision based on evidence from practice. You can see this principle in the design of the Learning Communities Project in Mexico, where the practice of *tutoria* is known, understood, and practiced without distinction between adults and young people. The main distinctions are among levels of mastery of the practice—expertise—rather than formal role—teacher versus student. People can operate at different levels of understanding and development in the practice, but the practice requires that everyone be developing their own expertise, and the expertise required for the practice is transparent and accessible to all. One sees a different demonstration of this principle in *Escuela Nueva*, where the practice is more complex pedagogically, but also accessible to young people and adults on an equal footing. We can talk about powerful pedagogies that seem to spread with relative ease in highly challenged settings, but we tend understate the ability to create, propagate, and spread practices that are, at once, powerful and transparent. Overdeveloped, highly institutionalized learning practices require armies of experts, consultants, and administrative overseers and the preconditions for their success become codified in the research literature on "scale," while there are existence

[1] My working definition of powerful learning: The capability of human beings to consciously modify understandings, attitudes, and beliefs over time in the presence of evidence, experience, and knowledge.

proofs that demonstrate that the power lies in the ideas and their transparency rather than in their inaccessibility to non-experts.

Another pattern I observe among successful outliers is active *pursuit of divergent thinking within a well-defined practice*. One such setting, for example, is a design studio, called NuVu,[2] for 14–18-year-olds, created by a group of architects from MIT, in which young people routinely develop, design, and produce prototypes for solutions to pressing social problems—adolescent health, refugee survival and welfare, mobility for the physically impaired in third world countries. The young people in this learning environment represent the full range of social, racial, and economic backgrounds. The way the studio works is that the adults introduce young people to the studio model of architectural practice through a series of simple design problems; they set to work on a broadly defined topic and form groups to work on a range of solutions; they present and critique prototypes through various stages; and they ultimately present their designs to experts in the field. A central working principle of learning practice at NuVu is that a "good" problem in one to which *neither the adults nor the young people know the answer*. The process requires divergent thinking in order to produce a large enough range of possible solutions to create an interesting and productive design space. But the process also requires that young people master a discipline of subjecting their practice to development and critique, bringing them into a constant state of challenge to learn something they previously did not know, and to find the source where it can be learned.

A third principle I observe in divergent learning environments might be called *the pursuit of intentional surprise*. The rise of standards based, test driven reform has unfortunately been accompanied by a similar rise in what Thomas Kuhn once called "normal science." Kuhn used this term to describe the stylized, textbook version of the scientific method used to explain science to non-scientists. Real science, Kuhn argued, does not take this form, and rather is focused on a much messier process of fits and starts driven by an, often vague, pursuit of discovery. The normal science version of inquiry that has been grafted onto education reform has led inexorably to randomized field trials, more and more rigorous and less informative measures of output, references to often small, decontextualized differences among schools, systems, states, etc. The stress in normal science is the hard-edged pursuit of generalizable, replicable "results" in a field where the science has yet to solve the puzzle of why strong interventions that work in one place fail, at the same level of fidelity, in another. What you see when you watch how divergent learning environments work is something radically different: You see adults and young people engaged in a common pursuit of interesting questions, disciplined by a set of agreements for membership in a community, with the expectation that the world is capable of delivering gratifying surprises if one stays engaged with it. I am spending a great deal of my life at the moment reading in the literature on the neuroscience of learning. One of the fundamentals of this field is that human beings are learning organisms for which the pursuit of learning is a survival mechanism, biologically and genetically programmed by eons of development. In this view, evidence of the absence of learning is not evidence of a human

[2] https://cambridge.nuvustudio.com.

predisposition against learning, but evidence of systemic, institutional, and cultural impediments to learning. The pursuit of surprise, a powerful motivator of learning, is threatening to institutionalized definitions of learning.

A fourth principle I observe in divergent learning environments might be called *the pursuit of intentional design*.[3] Among the wonderful learning opportunities available to me at this point in my life is that I get to work in a consulting capacity with an architectural firm that specializes in the design of learning environments. It is an unusual firm in the sense that in order to work with them as a client you have to be willing to state the theory of learning from which you expect the physical design to evolve. There are no a priori prescriptions about the end state of the physical design, only that you, as a client, work with the firm to develop the learning theory on which the design will be based. I have written about the connection between this kind of *design thinking* and the emerging neuroscience of learning in another context. One of the things I have noticed in this architectural practice and in my observation of divergent learning environments is how much educators' preconceptions of how learning occurs are determined (not just "influenced," but *determined*), by the institutionalized definitions of learning embodied in the physical environments in which they work. Many educators, for example, quite literally cannot imagine learning taking place in an environment where there isn't someone called a "teacher" present, accompanied by all the physical affordances by which that person exercises control over other people called "students." I confess that I often, in workshop settings play a perverse trick, in which I ask individual educators to recall and record the last time they "learned" something powerful enough to affect their lives. I then ask them where, and under what conditions, this learning occurred. Only the smallest minority of these adults locate the learning in a formal schooling environment, or a setting in which they were "taught" something and expected to "remember" it.

What is happening outside the formal schooling environment is that people who are relatively uninterested in replicating "school" are *reversing the established institutional definition of the relationship between form and function in learning*. Rather than asking what kinds of powerful learning can occur in this (literally) concrete, heavily institutionalized environment, they are asking, "if we set out to design a learning environment that was the best approximation of our current theory of learning, what would it look like?" Notice the terms "approximation" and "current," which suggest that learning organizations change and deepen their understanding of what they are doing over time and that all designs are provisional representations of their best judgment, subject of modification with additional learning.

One can see in the articles in this issue that educators are beginning the process of testing the limits of established learning organizations to address the outer limits of young peoples' roles in directing and developing their personal capacities as learners. Both Long Beach and Ontario have, predictably, pushed beyond the more conventional definitions of instructionally-focused reform, accompanied by

[3] I have tried to develop some of these ideas of intentional design around explicit theories of learning in my on-line course at HarvardX: https://www.edx.org/course/leaders-learning-harvardx-gse2x-0.

ambitious strategies of teacher involvement and development, into projects and programs that attempt to connect the world of the school to the world outside. These connections will inevitably lead, if they haven't already, to confronting different modalities of learning, different definitions of competence, and different challenges to learners' dispositions toward learning than young people have experienced in school. The question for the future is whether the institution of schooling is flexible enough to accommodate these divergent patterns of thinking and learning, or should divergent thinking be the province of other types of learning environments. In other words, can the social activity called "learning" grow fast enough and wide enough in the institution called school?

I am currently working in two school districts each of which is literally within less than an hour's drive from the Long Beach Unified School District, demonstrably one of the few beacons of "improvement at scale" in the U.S. The reason I am working in these two school districts is that they are chronically underperforming by the metrics of standards based reform. Each of these districts contain a number of schools, which, if the dictates of reform policies enacted at the state and federal levels in the early 1990's were actually enforced, would have been closed and reconstituted 10 years ago. The schools still exist. They are still chronically underperforming. The adults who work in these schools still go about their daily work feeling the overhang of an oppressive and hugely dysfunctional state and local bureaucracy. The social distance between the schools in Long Beach and those in these districts is as large as the distance between Long Beach and a failed state in Africa or Asia; the physical distance is a short drive. The demographics are virtually the same. In what world does the idea of "improvement at scale" have any meaning when conditions like this exist side-by-side in the same metropolitan area? After 25 years of policy-driven reform shouldn't we be thinking about different ways of understanding learning as a social activity, not just reinforcing the preserve of a privileged set of institutions to do the same things they have been doing for 25 years? If children in Mexico and Colombia can benefit from divergent models of learning, why can't children within a stone's throw of Long Beach benefit in the same way? Our conceptions of what is possible are too often constrained by our beliefs about what is feasible and practical in an archaic set of institutions.

In the final analysis, I read these interesting and engaging articles as a kind of valedictory to a period of education reform that will soon (hopefully) become history. I worry that we will not have learned how limited policy is as a mechanism for transforming learning in society. I worry about the excessive attention to, and preoccupation with, fidelity to practices some of which, at best, find their roots in an obsolete industrial, colonialist society. I worry about the growing attenuation of the gap between "education" as it occurs in schools, and "learning" as it is expanding exponentially in the broader society. I also worry that an entire generation of young people who would be astonishingly good at designing new learning environments for our children, will probably migrate away from this project because they don't want to work in schools. Against these worries, the concern about "scale" as a measure of value seems to pale. Which takes me back to the beginning of this essay. It seemed like a good idea at the time. Maybe we should move on.