Support a Science of Performance Improvement

BY ANTHONY S. BRYK

Given the many challenges facing education, it is natural to look to our nation’s research and development capacity to point the way. Even at its best, however, extant education research today only weakly informs the complex processes involved in improving teaching and learning. We simply lack the necessary infrastructure to guide the transformations we need.

This isn’t due to an absence of activity. An industry of education researchers is identifying problems and developing new concepts. Individual educators regularly experiment with new practices in their classrooms, and school districts are constantly introducing new programs. Individual consultants and commercial firms market new services, as state and federal policies incentivize new activity, and foundations promote bold new solutions seeking to reinvent the institutional enterprise. In short, an impressive array of individual projects focuses on many different issues.

Taken as a whole, however, this enterprise isn’t leading to solutions for major educational problems. We must reengineer both how we carry out education R&D and the institutional environments in which this work occurs if we want to achieve more productive ends. Something new and quite different must emerge if we are to confront educational challenges successfully.

Build a Capacity to Inform Improvement

Education needs a Design, Educational Engineering, and Development (D-EE-D) infrastructure. This

ANTHONY S. BRYK is president of the Carnegie Foundation for the Advancement of Teaching, Stanford, California.
activity should be organized around the core problems of practice embedded in the day-to-day work of improving teaching and learning and in the institutions where teaching and learning take place. Making progress in addressing educational problems requires a commitment to a rapid prototyping process by which researchers and practitioners co-develop innovations, try them in schools and other learning contexts, and then refine and try them again. This new infrastructure demands an engineering orientation in which adaptability to local contexts is a direct object of study. In this regard, knowing that a program can work is not good enough; we need to know how to make it work reliably over many diverse contexts and situations. This means accumulating a rigorous knowledge base on practice improvement where the real test of adequacy is its capacity to advance demonstrable, broad-based improvements in teaching and learning.

Undergirding this proposal is a core set of ideas about a science of performance improvement. First and most important, a D-EE-D agenda puts the day-to-day work of educators at the center of the inquiry and focuses attention on solving problems of practice that have genuine consequences for people’s lives. Most education research today isn’t organized in this fashion, and the field doesn’t hold itself accountable for actually making a difference in this regard. While a policy concern (e.g., teacher quality) may stimulate a creative mix of projects, nothing in the current system forces these diverse projects to converge on achieving common goals.

Second, D-EE-D needs to bring scientific discipline to bear on this activity. You can’t improve practice at scale unless you measure it and unless the evidence base is organized around some working theory about how various instructional processes, organizing routines, and cultural norms interact to affect desired outcomes. Educational interventions typically posit complex causal chains between some reform and the desired outcomes. D-EE-D must be disciplined by an explicit cause-and-effect logic about the multiple subprocesses that must cumulate for improved learning to occur (Cohen, Raudenbush, and Ball 2003). Moreover, we must evaluate the cause-and-effect logic itself by examining evidence that emerges each time the innovation is used in practice.

Third, D-EE-D places a premium on achieving reliability in the performance of work activity. In other words, we want innovations to achieve positive results as they are used in practice. In any human resource-intensive enterprise, such as schooling, variations in performance are the natural state of things. We have ample testimony to this from decades of educational innovations. That a practice, program, or service can work is of little value unless we discern how to make it work at scale in the hands of many different individuals working under diverse circumstances. For example, consider the introduction of school-based coaches to improve teaching and learning. The efficacy of this innovation depends on a host of factors, including the selection and professional preparation of the individuals who take on this new work, support of the local school contexts (e.g., principal leadership, prevailing faculty norms about opening practice to outside scrutiny), how coaches actually enact the role (e.g., decisions about which teachers to coach and on what topics), and the technical core of the coaching program itself (e.g., materials, coaching procedures, and routines provided to guide this work). Each of these elements has potential to change the program effects. Understanding the sources of this variability is key to making coaching work at scale.

Fourth, this activity entails a hybrid methodological stance that combines the strengths of action research and systematic large-scale inquiry. The strength of action research is its context specificity; but this is also its weakness in that it typically lacks a strategy for accumulating evidence over different contexts and times. In contrast, randomized control trials can yield precise information about the average effect of a program aggregated over some sample of sites, but they offer little guidance about how the innovation might actually operate (and what effects might actually accrue) as it moves into many different contexts. A D-EE-D strategy entails a melding of systematic inquiry about effects with a deep understanding of the mechanisms operating to produce these effects.

Aim for Usable Knowledge

Think about these principles in the context of professional development initiatives. Consider the “full causal cascade” that leads from the initial design of a professional development initiative all the way to the intended outcomes — improvements for students. For example, a university creates a professional development program for local public school teachers. Immediately we should ask, “What do we want teachers to learn and be able to do? And, how do we know whether teachers can do this based on the professional development that we designed?”

Next, assuming that teachers have acquired some knowledge and skill, we must determine whether
participation in the professional development program translated into observable changes in instruction. Can we track what happens day to day in classrooms? And then, assuming we’re able to document changes in practices, the next big step is to determine if these practices are related to measurable improvements over time in children’s learning gains in those classrooms.

That a practice, program, or service can work is of little value unless we discern how to make it work at scale in the hands of many different individuals working under diverse circumstances.

The goal, simply put, is to create usable knowledge about practice. This means being very specific about what it is, testing it, refining it, and representing it in various tools, artifacts, social routines, and common rituals (Hiebert, Gallimore, and Stigler 2002). It is not enough to say, “This works. Use it.” We need to know if people can learn to use this knowledge with some reliability. This is the “learnability test.”

Atul Gawande offers a poignant example in writing about innovations in medical practice early in the 20th century (2007, pp. 177-180). At that time, there was a great deal of concern about infant and maternal mortality during childbirth. Physicians were beginning to use forceps to assist in childbirth. However, teaching doctors how to use them properly was difficult because this was basically a matter of feel, and feel is very hard to teach. Moreover, the consequences of getting it wrong tended to be catastrophic. Ultimately, this innovation failed the learnability test. While forceps are a superb tool in the hands of some doctors, others failed badly.

Organize for Teacher Learning

The work of teaching and its improvement has been plagued by a classic dualism. On one side, we have scripted instruction. The basic idea here is that learning will improve if teachers follow an externally defined protocol for each lesson and follow the pacing guide for how lessons sequence over time. At the opposite end is a view that every teacher, child, and classroom setting is unique. This orientation calls on each individual teacher, each day, to invent instruction anew.

Scripted instruction has difficulty responding to variations in the background, interests, and prior knowledge of students and the varied ways in which they learn. It also tends to break down when the instructional tasks become more ambitious and open-ended. The second, more craft-oriented approach to teaching has resulted in some wonderful individual classrooms, but those have not been replicated at scale.

A more productive middle ground conceives of teaching as an organized complexity (Simon 1996). Skillful teaching involves a dynamic interplay of understanding around students’ background knowledge, motivations, dispositions, skills, and interests; the immediate goals for instruction; the command of a set of pedagogical tools and resources and automaticity in their use; and a capacity to continuously adjust on the fly.

This capacity to recognize critical patterns in classroom activity and react quickly and appropriately distinguishes novices from experts across many domains. Consequently, it is sensible to conceptualize teacher learning as a problem of expertise development. Individuals develop expertise by having many opportunities to engage in guided practice with others who are more expert. While this observation shouldn’t appear controversial, little teaching improvement activity is actually organized around this perspective. Instead, from the scripted-instruction tradition, we have tended toward a language about implementing practices or programs with fidelity. Embedded in this language is a concept that teaching is routine work that can be defined, delivered, observed, and monitored. This leaves little room for recognizing the dynamic and complex character of teacher decision making described above.

In contrast, the concept of a professional learning community appears more consistent with the dynamic and context-bound nature of instructional decision making. However, these initiatives also suffer from their own distinctive zone of wishful thinking. The question of how these communities might actually organize and function to improve instruction typically remains unaddressed. In operation, these efforts seem predicated on the idea that if we just gather teachers together to talk about practice, something good will happen. This too strikes us as a weak working assumption for change at scale.

I propose a modest amendment. The social organization for improvement is a professional learning community organized around a specific instructional system.
Montessori primary education provides an example of this. The instructional system details learning objectives for children, an organization of the classroom, specific roles for teachers and students, instructional routines, materials, and a classroom culture. Specific training programs have been developed to prepare novices to teach in the complex ways demanded by this approach (Cossentino 2005). Finally, the community focuses on improving instruction and preparing individuals to teach this way. In short, Montessori education is a professional learning community organized around a specific instructional guidance system.

Reading Recovery offers another exemplar. Entry into this professional community begins with an intensive year-long training program. Novices are introduced immediately to a systematic reading tutorial process that uses a common set of pedagogical practices and materials integrated around a working theory of how students learn to read. Regular collection of common data, such as taking running records, informs all instruction. Individuals who lead the program have demonstrated skill as Reading Recovery teachers and have been trained to teach other adults. Reading Recovery collects and analyzes data on student progress in order to continually improve instruction. Finally, a participant community is formed around all of this so that participants learn from each other.

Taken together, these two examples illustrate the organizing power of a specific instructional system on the activity of a professional learning community. The instructional system details what teachers need to know about students in terms of background knowledge, skills, and interests. It involves some very specific pedagogical practices and social routines and expects automaticity in their use. Educators have a shared language about goals for students and understand how these align over time around some larger conception of student learning. Teachers also share a common evidence base about what constitutes learning. This allows them to analyze and refine the cause-and-effect logic that organizes their shared work. Finally, tying this all together is an explicit process for socializing new members into the community and for organizing ongoing social learning among all participants.

Aim for Ambitious Goals for Students

The stakes for improving education in the United States have never been higher and the demand for productive R&D never more important. In a global economy in which increasing numbers of students around the world are achieving basic academic skills, the United States must do more if we are to maintain preeminence as a first-world economy and if we are to sustain our national belief in equal opportunity for all. Success in a knowledge economy requires students who can flexibly apply conceptual knowledge and basic skills in the analysis of complex problems. It demands more sophisticated communication skills and the ability to work effectively with diverse others. It entails efficacy in the use of a variety of new multimedia tools to support analysis and representation of ideas that enhance communication and social learning. It demands ease in the pursuit of complex, ambiguous tasks in which missteps are common. And it entails forming those habits of mind and heart that manifest in the tenacious pursuit of worthy goals. Especially in this new conceptual age, Thomas Edison’s observation remains salient, “Genius is one percent inspiration and ninety-nine percent perspiration.”

In truth, no one really knows exactly what all of this means for the future of education in the United States. However, two concluding points do seem clear. First, our education systems must do much more for many more than they have ever done before. Second, a vital R&D infrastructure must aim high because the long-term cost of mediocrity in our education system is simply unacceptable in societal terms. The goal for R&D is simple: Dramatic improvements in student learning are the coin of the realm.

REFERENCES


