

# BROOKINGS

## THE BROWN CENTER CHALKBOARD

« Previous

Thomas J. Kane | January 29, 2014 11:00am

# An Urgent Need for Short-Cycle Clinical Trials in Education



*“If you said to me, are we making progress on [U.S. education reform] or not, I could talk for a long time, but I wouldn’t be able to give you a number.”* --Bill Gates with David Rubenstein, Sanders Theater, September 21, 2013

In the three decades since the release of the *Nation at Risk* report, the U.S. education reform effort has failed to achieve lift-off. *Why is that so?* Regardless of the reform strategy—whether new standards, or accountability, or small schools, or parental choice, or teacher effectiveness—there is an underlying weakness in the U.S. education system which has hampered every effort up to now: most consequential decisions are made by district and state leaders, yet these leaders lack the infrastructure to learn quickly what’s working and what’s not. They launch new initiatives with no detailed analysis of their effects. At best, they track aggregate measures such as overall proficiency and graduation rates, which can hide the consequences for the specific schools, or grades or subjects actually affected by their initiatives. And, when there is turnover, new leaders are forced to re-invent the wheel, blind to the mistakes and successes of their predecessors. For their part, philanthropists fund new initiatives in their local schools, and never know whether their funds have made a difference for children.

We are not lacking innovation in U.S. education. We lack the ability to learn from our innovations.

Over the last decade, the U.S. Department of Education has built a robust infrastructure for evaluating programs at the federal level. The National Center for Education Evaluation has funded more than 34 large-scale impact evaluations since 2002. These cost an average of \$12 million and have taken 5.6 years to complete. Such a system might be sufficient in countries where education policy decisions are made at the federal level and where there is greater continuity of leadership. Unfortunately, state and local decision-makers—who play the critical role in the U.S. system—too rarely make the connection between the lessons learned in federally-funded evaluations and their own policy decisions. We need a new model to supplement these federal efforts, which is faster, less expensive and more closely tied to the decisions being made at the state and local level.

The Common Core standards and new teacher evaluation policies are a good example. Although the federal government helped create the policy framework, its ultimate impact will be determined by thousands of implementation and policy decisions at the state and local level. So, how could we ensure that state and local leaders get the evidence they will need to find the best solutions?

Here's an outline of one approach, which a group of states and large districts could undertake collectively:

1. Suppose a group of states were to invite panels of teachers to assemble packages of materials targeted at the most demanding new standards in each grade and subject. Each package should contain a training component and a feedback component. For instance, in addition to receiving training and curriculum materials, teachers might be given cameras to submit videos of their lessons teaching the new standards, for comment from peers, principals and content experts. (Implementing the new standards will require massive adult behavior change. And any adult behavior change requires feedback. Postponing the implementation of teacher evaluations would be like launching a Weight Watchers program with no bathroom scales or mirrors for the participants.)
2. Teams of teachers by grade level and subject in schools would be invited to participate in the trials, from all the participating states.
3. The states would find a partner to organize the trials to test the packages. From the volunteer schools, a subset would be chosen by lottery to receive the treatments in specific grade levels and subjects. (A school might win the lottery in one grade and subject, but not in others. They would serve as control group schools in the grades and subjects where they were not chosen for treatment.)

Randomly assigning treatments at the school/grade level would eliminate the need to analyze student-level data. Assembling and cleaning student-level data accounts for much of the cost and delay in traditional evaluations.

4. The research would piggy-back on federal data reporting requirements (using school-level and subgroup means by grade and subject rather than student-level data). That way, the tables could be prepared beforehand and impact estimates could be produced within days of state reporting during the summer following each school year.

Initially, the effort would need to be launched with federal or philanthropic support. If such a program were to start immediately, the group could be randomly assigning materials this summer. By next summer (the summer of 2015), the group could be pointing to successful models for combining teacher training and teacher evaluation, backed up by strong evidence.

The urgent need for the crisp evidence provided by randomized trials cannot be underestimated. Strangely, education leaders speak about “effective professional development” in the same way that foreign policy experts speak of Middle East peace: everyone says it’s essential, but no one seems to believe it can work! That’s a problem, because it’s impossible to rally a groundswell of support unless you believe (and can demonstrate) your strategy will work. State and district leaders have a brief window of time—perhaps two or three years at the most—to prove it. (Otherwise, the next cycle of reformers will reinvent the wheel again.)

This would just be the first step. Not every intervention would lend itself to random assignment. Universities will need to train state and local agency staff to use their data to evaluate policies and programs. Software providers will need to automate the statistical algorithms used by PhD-level analysts in the more expensive customized evaluations. Once they’ve seen the payoff, state and district leaders will need to find the resources needed to hire them.

---

Thomas J. Kane

Nonresident Senior Fellow, Governance Studies, Brown Center on Education Policy

Thomas Kane is professor of education and economics at the Harvard Graduate School of Education and faculty director of the Center for Education Policy Research, a university-wide research center that partners with states and districts to evaluate innovative policies. He was deputy director in the education group at the Bill & Melinda Gates Foundation, where he directed the Measures of

Effective Teaching project.

[More Posts from Thomas >](#) | [View Expert Page >](#)

## 0 comments



Start the discussion...

Best ▾

Brown Center Chalkboard

Share

Login ▾

Be the first to comment.

[Subscribe](#)

[Add Disqus to your site](#)