

Using research to improve education under the Every Student Succeeds Act

Mark Dynarski

Executive Summary

The Every Student Succeeds Act, the new reauthorization of the federal program designed to support the education of disadvantaged students, requires that states and districts use evidence-based interventions to support school improvement. Researchers have studied the effectiveness of education programs for decades and that effort is now producing substantial gains in knowledge of what works and what doesn't. But educators note that this kind of research is not as useful as it could be for them because it is conducted in settings that differ from theirs. They are interested in research that fits their contexts.

Recently, another kind of research paradigm has emerged in which researchers work directly with educators to identify and implement paths for improvement within particular settings. This new kind of research—which has come to be known as improvement science—operates in local contexts of districts and schools. But it faces a capacity problem because there are relatively few researchers participating or able to participate in these efforts compared to the number of districts and schools that could benefit from more evidence-based programs and practices.

The two approaches need to be coordinated. In the first stage, research would identify effective programs and practices writ large. In the second stage, districts or schools not meeting targets or objectives would work with improvement-science teams to adapt those research-proven programs to local contexts.

The 'Every Student Succeeds Act' also creates a new program to support research on innovations in education. Using the existing infrastructure of the regional lab network can help identify priorities for this new research on innovations. The priorities should fill gaps in knowledge and proven programs that states and districts have identified as important to them.

Background

Scaling up effective education policies or programs is in everyone's interests. Who argues that education should not improve? And findings from research often underscore where improvements are possible and where education can be more effective. But scaling up findings from research—having the findings lead to actions on a larger scale—is a challenge.

The issue is partly the size and dispersion of authority of the public-education enterprise, with its 15,000 districts, 65,000 schools, 4 million teachers, and 55 million students. For an improvement to find its way into even a fraction of this enterprise counts as progress and might be seen as miraculous.

How does the process of adoption appear to be working? Do improvements identified by research find their way into the enterprise at all? Brian Jacob's recent [note](#) here on Evidence Speaks commented on learning from research 'failures,' which arise when evidence emerges that a promising idea did not improve education outcomes when tested rigorously. A related question is whether the enterprise learns from research successes, when evidence emerges that a promising idea works. Do these successes become education practices?

The answer is hard to know because the extent to which research finds its way into schools and classrooms has not been measured. When educators are asked about research, however, they point to their perception of research having a "local perspective" as one reason for caution about using the research for programmatic decisions. They give more credence to findings that arise in contexts similar to their districts or schools compared to findings emerging elsewhere.ⁱ

This 'localism,' for lack of a better word, combined with limited avenues for research dissemination, has led to new forms of research in which researchers work directly with educators to develop local practices and programs. In the words of one of its foremost practitioners, this research takes a 'design-engineering-development' perspective, working from the ground up to tackle educators' problems.ⁱⁱ This approach is now known as 'improvement science.'ⁱⁱⁱ

This local approach sounds like an ideal way to move research closer to what educators value, program development and evidence about outcomes that occurs in their particular schools or districts and that results from a researcher-educator collaboration. Maybe there will come a day in which most schools or school districts are sufficiently resourced to have their own program development and evaluation teams. But even then local improvement science would need to be a complement to research efforts to

identify effective practices, not a substitute for them. Working with educators to promote more effective reading or math instruction, for example, needs to begin with sound research on reading and math instruction. Improvement science then can focus on encouraging and promoting the take-up of that sound research, working with educators to adapt what are believed to be key ingredients of the approaches. We'll return to that two-stage approach below.

Identifying effective practices is not the same as implementing them

The dominant approach to studying a question of whether a practice or program improves an education outcome is known as 'effectiveness' research. Often experiments are used to answer the question—Are teachers who attend these workshops more effective? Does this dropout-prevention program keep students in school? Does using this software lead to stronger math skills? There are different kinds of methodologies that address effectiveness, including randomized controlled trials, regression-discontinuity designs, well-designed quasi-experimental studies, and single-case designs, but it is convenient to use 'experiments' as a term for all of these ways of measuring the effects of programs, practices, or policies.

Experiments were relatively uncommon in education, certainly compared to their use in medicine, until 2002, when Congress created the Institute of Education Sciences. Since 2002, IES has funded hundreds of experiments and disseminated their results, mostly through the web and through workshops, webinars, and social media channels. It also disseminates syntheses of research and appraisals of individual effectiveness studies through its 'What Works Clearinghouse,' by way of the web and through the other channels. The information reaches a large audience. Practice guides produced by the What Works Clearinghouse are downloaded about 22,000 times a month. One of the Clearinghouse's most popular practice guides was downloaded nearly 90,000 times in its first *month* of release.^{iv}

For disseminating research findings cheaply, it's hard to top this model. In principle, every educator can hear about every finding of relevance to them for the cost of looking up a web page or watching a recorded webinar on YouTube. But whether the findings actually change educator practices is not known. Educators could be ignoring the findings and continuing to do what they are doing. The Government Accountability Office expressed concerns about this possible disconnect in its [recent review](#) of IES.^v

Findings from experiments are information, but changing practices to do something with the findings is implementation. As Pfeffer and Sutton (2000) have written, knowing is a long

way from doing.^{vi} ‘Improvement science’ strives to close the gap between knowing and doing. At a risk of oversimplifying, improvement science poses a model—such as the ‘design- engineering-development’ one mentioned above—in which researchers work directly with educators in districts and schools. The focus is on using rapid tests of change and ‘Plan-Do-Study-Act’ cycles to learn by doing, and connecting participants (teachers, principals, administrators) through networks to expedite their learning.

For example, the Carnegie Foundation for the Advancement of Teaching is collaborating with community colleges to promote success in math, with urban school districts to improve skills of their new teachers, and with school districts and organizations to design classroom experiences that promote ‘academic mindsets’ and support students to develop their own learning strategies.^{vii} Vanderbilt University is collaborating with school districts to enhance middle- school math curricula and create new kinds of professional development and teacher networks to improve math teaching.^{viii}

By its nature, improvement science focuses intensively within districts. That focus is a plus because researchers and educators are at the table, but also a minus because there are not enough researchers to be at the many tables that need them. There are about four times more school districts than higher-education institutions, and many higher-education institutions have no one with the time or, in many cases, interest or expertise, to anchor the improvement science effort at a local school district. Unless improvement science can generate knowledge of how schools and districts can improve without researchers being involved in thousands of school districts, the limited number of researchers essentially precludes scaling up. And if the knowledge improvement science generates in a few districts has to be disseminated to many others, improvement science ends up being in the same place as effectiveness research: educators might not hear about the findings or might view them as too distant to be useful in their local areas.

So findings from experiments that are intended to produce generalizable results can be inexpensively disseminated but might not be used by educators, and improvement science may yield more local knowledge but cannot operate widely because of capacity and generalizability issues. The key is for effectiveness research to be more ‘localized’—more applicable to educators in their schools and districts—without it necessarily having to be produced locally.

Experiments can be more useful to educators

Depending on the intervention or policy being considered, a district or school needs to first learn of an improvement—

for example, a relevant research finding that has been published, written about in the media, or passed along to local educators through word of mouth. Then, the educators’ questions become local. How much did staff in the study differ from staff in their schools? Did characteristics of students have a role in the experiment’s findings? Are the powers-that-be in the local education system, including its teachers, open to the type of program that generated the positive research findings? Does the state or district have the authority to implement the program that the research studied? Can school or district afford the out-of-pocket expense for licenses, fees, or materials, and the staff time to learn how to do the program?

These are a lot of judgments, and a significant gap opens up between how a researcher may view the findings—‘the study shows that the approach worked’—and how an educator might view the findings—‘it worked for somebody but I don’t know if it can work for me.’ What a researcher views as evidence becomes what an educator views as one variable in a risk equation when the risk itself is avoidable by sticking to the tried and true.

Move the approaches closer together

Both effectiveness research and improvement science can add to knowledge and both approaches can be useful. Both need to measure effects, provide information to prospective adopters about how to implement the program or approach, and be explicit about how much it will cost.

Scaling up should be the starting point in thinking about how to design experiments and improvement science efforts. If studies were designed from the view of scaling up, they would focus on developing information prospective adopters need: how large are effects, how can the program be implemented, and how much is it likely to cost? Approaches to effectiveness studies vary somewhat in how they measure effects, but they vary much more in how they study costs and implementation. Costs are rarely analyzed, and while some experiments report several hundred pages of detailed information about implementation, others describe implementation in a report chapter and many published papers simply do not mention it.^{ix} The risks educators face in implementing programs shown by research to be effective would be mitigated if research on implementation focused on creating a manual on how to carry out the program. Researchers developed a process for ‘manualization’ more than a decade ago, but that process is rarely used in studies of education programs.^x

Using a two-stage model for generating evidence on effective, implementable interventions will help put experiments and improvement science into balance. In the first stage,

districts and schools would be able to learn about recent research on effective practices, how to implement those practices, and their costs. Currently there is no organization doing all that is envisioned here for the first stage of the model. The [What Works Clearinghouse](#) and Best Evidence Encyclopedia provide information on evidence of effects but little information about implementation and cost.^{xi} This is not for lack of interest on their part; most research studies and reports provide too little information about implementation and cost, and standards are not in place for how to assess what is provided. Efforts to document implementation and cost need to be increased for this stage to be useful.

The second stage is improvement science. Districts that use evidence from the first stage but are not satisfied with the results or do not meet targets can work with improvement scientists to adapt interventions with evidence of effectiveness and monitor the results. The second stage needs only enough improvement-science capacity to work with districts that are committed to it. This may still be too many districts and not enough capacity, but starting from the total number of districts certainly overwhelms capacity, as noted above, whereas thinking of improvement science as targeted moves it into the realm of practicality.

The federal role in the two-stage model

The [‘Every Student Succeeds Act’](#) gives states responsibility to develop accountability structures. These structures need to include ‘comprehensive support and improvement plans’ for schools that need improvement, and these plans must include evidence-based interventions. Using the two-stage approach—with districts and schools moving to the second stage if improvement targets are not met—is a sensible means to develop a pool of evidence-based interventions that meet state needs.

Schools also might move into the second stage if they fall

in the 5 percent of schools for which Congress is requiring states to intervene. (States can choose to intervene in more schools, but not less than 5 percent.) The constellation of issues these schools face is an opportunity for educators and researchers to work together to identify improvements and implement new approaches. Having improvement-science teams working with schools that are most in need of improvement is a reasonable way to blend the strengths of the two approaches.

The two-stage model also can be connected to the new ‘Innovation Research’ section of the Act (section 4611). The section calls for the U.S. Department of Education to fund research to develop, test, and scale effective practices. The language does not indicate how funding priorities are to be identified.

One way to do so is to ground priorities in expressed state and local needs. For example, if language acquisition is a need in rural areas of the Southwest, and research identified in the first stage is thought to be inadequate, some of the innovation grants could be used to fill that gap. Similarly, states and districts might express needs to bolster reading, math, or kindergarten readiness, or any number of other objectives. The regional lab network already has an infrastructure for assessing needs at state and district levels. It can have a role in tying these needs to innovation priorities and monitoring whether needs change over time. The Institute of Education Sciences within the department, which operates the labs under contract, is well positioned to work with the labs to identify innovation priorities emerging from local needs. Labs also can conduct research to meet needs that do not become innovation priorities.

Education research is sparsely funded and is unlikely to enjoy the resources of the National Institutes of Health any time soon. Effectiveness research and improvement science need to be deployed in concert to make the best use of these scarce resources.

ⁱ The most common theme that emerged from a survey of educators about hurdles in their use of evidence was ‘localism.’ See Nelson, S., J. Leffler, and Barbara Hansen. “Toward a Research Agenda for Understanding and Improving the Use of Research Evidence. Accessed December 8, 2015: <http://educationnorthwest.org/sites/default/files/toward-a-research-agenda.pdf>.

ⁱⁱ See Bryk, A., and L Gomez. “Ruminations on Reinventing an R&D Capacity for Education Improvement.” Accessed December 8, 2015: http://cdn.carnegiefoundation.org/wp-content/uploads/2014/09/DED_paper.pdf.

ⁱⁱⁱ See ‘Learning To Improve: How America’s Schools Can Get Better At Getting Better,’ by Anthony Bryk, Louis Gomez, Alicia Grumow, and Paul LeMahieu (2015), Bryk and Gomez’s 2008 [paper](#), and the National Academies [monograph](#) laying out a design for the Strategic Education Research Partnership. Cohen-Vogel and her colleagues provide a succinct recounting of the emergence of improvement science against the backdrop of experiments. See L. Cohen-Vogel et al., “Implementing Educational Innovations At Scale: Transforming Researchers into Continuous Improvement Scientists.” *Educational Policy*, vol. 29(1), 257-277. Roots of improvement science can be found in the work of Donald Berwick in health-care, beginning in the eighties. Berwick, D. “The Science of Improvement.” *Journal of the American Medical Association*, 2008, 299(10), pp.1182-1184.

^{iv} Full disclosure: I directed the What Works Clearinghouse from 2008 to 2010, when practice guides were first released, and I chaired a panel that produced one of the first guides. I continue to be involved with the Clearinghouse in various roles.

^v IES recently funded two research centers to explore these topics but it will be several years before findings are known.

^{vi} Jeffrey Pfeffer and Robert Sutton. ‘The Knowing-Doing Gap.’ Harvard Business School Press: 2000.

^{vii} <http://www.carnegiefoundation.org/in-action/student-agency-improvement-community/>

^{viii} http://peabody.vanderbilt.edu/departments/tl/teaching_and_learning_research/mist/index.php

^{ix} For examples, see the Reading First study reported [here](#), the study of teacher induction programs reported [here](#), and the study of supplemental services reported [here](#).

^x For more on manualization, see <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4126244/> and Carroll KM, Nuro KF. One size cannot fit all: A stage model for psychotherapy manual development. *Clinical Psychology: Science and Practice* 2002; 9(4): 396–406.

^{xi} <http://www.bestevidence.org/>