# REVISITING SIG

## WHY CRITICS WERE WRONG TO WRITE OFF THE FEDERAL SCHOOL IMPROVEMENT GRANT PROGRAM

BY ALAN GINSBURG AND MARSHALL S. SMITH

SEPTEMBER 2018

FutureEd
GEORGETOWN UNIVERSITY

## About the Authors

Alan Ginsburg served as Director of Policy and Program Studies at the U.S. Education Department. Marshall S. Smith was Under Secretary of Education in the Clinton administration. He is a senior fellow at FutureEd and at the Carnegie Foundation for the Advancement of Teaching.

## About FutureEd

FutureEd is an independent, solution-oriented think tank at Georgetown University's McCourt School of Public Policy, committed to bringing fresh energy to the causes of excellence, equity, and efficiency in K-12 and higher education. Follow us on Twitter: @FutureEdGU

## Use

# REVISITING SIG

## WHY CRITICS WERE WRONG TO WRITE OFF THE FEDERAL SCHOOL IMPROVEMENT GRANT PROGRAM

### Table of Contents


| | |
|---|---|
| | Foreword |
| 1 | Report |
| 5 | Results of SIG Studies |
| 8 | Recommendations |
| 10 | Appendix |
| 16 | Endnotes |

FutureEd
GEORGETOWN UNIVERSITY

# FOREWORD

With a multi-billion-dollar price tag and a reform menu set by the federal government, the Obama administration's School Improvement Grants program was nothing if not controversial. When a 2017 evaluation commissioned by the U.S.Department of Education research arm reported that the three-year SIG grants didn't move the achievement needle in the low-performing schools the program targeted, critics pounced. Education Secretary Betsy DeVos used the study as a cudgel again federal involvement in school improvement and to tout private school vouchers.

But after studying the 2017 SIG evaluation and analyzing the 17 state or district SIG studies they could find through an in-depth literature review, FutureEd Senior Fellow Marshall Smith and researcher Alan Ginsburg have concluded that the federally funded SIG study was wrong to conclude that SIG didn't make a difference.

Smith, the former dean of the Stanford University Graduate School of Education, was U.S. Under Secretary of Education in the Clinton administration and served as a senior counselor to Education Secretary Arne Duncan early in the Obama administration, though he was not involved in the development of the SIG program. Ginsburg was Director of Policy and Program Studies at the U.S. Education Department, where he worked for nearly four decades under both Republican and Democratic administrations. Like Smith, he didn't participate in the SIG program or in the 2017 evaluation, which was commissioned by the Education Department's independent research arm, the Institute of Education Sciences.

FutureEd asked several of the nation's leading education researchers to review one or more of Smith and Ginsburg's drafts, including Future advisors Jane Hannaway of Georgetown University's McCourt School of Public Policy and Tom Dee of Stanford University's Graduate School of Education. In recent years, Dee studied the impact of the SIG program in California.

Susanna Loeb, director of the Annenberg Institute for School Reform at Brown University, and Dan Goldhaber, director of the Center for Education Data & Research at the University of Washington, also reviewed drafts. Goldhaber simultaneously serves as director of the federally funded National Center for Analysis of Longitudinal Data in Education Research (CALDER) at the American Institutes of Education (AIR), a research organization that co-authored the 2017 SIG study. Hannaway is a former CALDER director and has an appointment at AIR.

These researchers offered valuable feedback and agreed with Smith and Ginsburg that it unwarranted to conclude that the SIG program was unsuccessful. The reviewers, however, are not responsible for the content of this report. FutureEd Editorial Director Phyllis Jordan played a central role in the report's development, and Molly Breen and Jackie Arthur made many editorial contributions.

We think it is particularly important to cast the SIG program in a clearer light as states and localities begin to implement the Every Student Succeeds Act. The new federal law replaces the SIG program with a requirement that 7 percent of state Title I funds be spent on turning around states' lowest performing schools. FutureEd's mission is to help policymakers and practitioners make informed policy decisions on behalf of the nation's students. We offer this report in that spirit.

**Thomas Toch**

**Director, FutureEd**

**In 2010, the Obama administration invested $3.5 billion in a federal grant program to turn around roughly 1,250 of the nation's lowest-achieving public schools.[1] Seven years later, a high-profile study commissioned by the U.S. Education Department found that the three-year School Improvement Grants (SIG) had made "no significant impacts on math or reading test scores, or high school graduation."[2]**

Critics of the SIG program, in which the federal government eventually invested some $7 billion, have used the results to cast doubt on the program's turnaround strategies. Others have seized on the findings as evidence that low-achieving public schools cannot be reclaimed, even with substantial resources. Education Secretary Betsy DeVos has repeatedly used the study's conclusions to make her case for expanding private school vouchers as a better way to help students in the weakest schools. "We have invested billions and billions and billions of dollars from the federal level, and we have seen zero results," she said in a 60 Minutes interview last March.[3]

But a close examination of the IES-funded study and a comprehensive analysis of 17 separate state or district studies of SIG initiatives by independent researchers and state education agencies that we identified through a wide-ranging literature review suggests that the autopsy on the grant program is flawed and its core conclusion faulty. Our research suggests that many SIG programs did indeed produce significant improvements in student achievement.

The federally funded study, conducted by Mathematica Policy Research and the American Institutes for Research (AIR), both nonprofit research organizations, looked at too small a sample of schools to draw dependable conclusions. The study's small sample size required unrealistically large gains to detect "statistically significant" improvement in student achievement at SIG schools when compared to others.

To achieve the study's benchmark, schools would have needed to show growth equivalent to at least 60 percent of the typical annual math achievement gain at the elementary school level and a full year's gain at the middle- and high-school levels.[4] Such results are rarely seen in education research that measures the effects of overall school interventions on student achievement—and are greater than those achieved by several highly regarded education programs.

Second, the schools included in the federal study were not nationally representative of the SIG school population, a reality that the study's authors acknowledge.[5] The study included a far higher

percentage of urban schools and students from disadvantaged backgrounds than the SIG program did nationally, skewing the school sample.[6]

By comparison, 12 of the 17 state or local studies on the impact of SIG interventions showed statistically significant gains in math, reading or high school graduation when compared to the performance of similar schools that didn't participate in the SIG initiative.[7] While their methodologies varied, these studies collectively looked at more than twice as many SIG schools as the federally funded study, which was commissioned by the Education Department's Institute of Education Sciences (IES).

Further, some of the state and local studies contained important lessons about what works in turning around struggling schools. They also provided educators with more timely information on how interventions were working than did the federally funded study, which was completed years after the SIG program ended.

To be sure, the Mathematica and AIR researchers faced the daunting task of evaluating a nationwide program administered very differently from state to state. The limitations of the study and the positive findings by a dozen state and local research teams raise the question of whether the Department of Education should change its approach to evaluating federal school-improvement initiatives.

Rather than rely on national studies to judge the performance of federal initiatives, the Education Department would better serve the education sector by drawing more heavily on state and local studies, including results released before programs have run their course. That would allow the Education Department to provide schools and school districts with early insights into what is working and what is not.

It is particularly important to cast the SIG program in a clearer light as states and localities begin to implement the Every Student Succeeds Act (ESSA). Under SIG, the federal government invested an average of $2.8 million per school, spread over three years. The new federal law replaces the SIG program with a requirement that 7 percent of state Title I funds, over $1 billion a year, be spent on turning around states' lowest performing schools. While states will have wide leeway to design, implement, and monitor this new generation of school-improvement work, a more accurate accounting of the nation's return on its investment in SIG will help state leaders set the best possible path forward for students in struggling schools.

## A Lack of Statistical Power

The IES-funded study's bleak conclusion that schools in the SIG program showed no significant gains belies the fact that achievement scores did rise in most of the SIG schools.[8] The key issue was not whether the SIG schools improved but whether they improved more than similar schools that didn't receive grants. Answering that question required identifying similar schools for comparison.

The authors of the IES-funded study used a sophisticated statistical methodology called regression discontinuity to come as close as possible to a fair comparison between the schools that received SIG grants and those that didn't. The methodology involved a several-step process: establishing a common national award threshold, such as a school being in the bottom 5 percent of school performance; identifying schools that fall within a band on either side of the threshold; and then comparing the three-year achievement and graduation rate performance of the schools on either side of that band. This process created a number of challenges.

First and most importantly, the design of the SIG study left researchers with a sample size that made it hard to detect gains reflecting improved student achievement in the SIG schools compared to schools that didn't receive

grant support. In research parlance, the study was severely "underpowered." Brown University researcher Susanna Loeb estimates that the study's "statistical power" was so weak that the student performance gains in SIG schools "would have to be unrealistically large for the study to have been able to detect them."

The schools would have had to achieve an "effect size" of 0.3 for the schools' accomplishments to be judged statistically significant, Loeb wrote in a 2017 blog post. That means, she concluded, "the reform would have had to close the gap in math achievement between the students in SIG schools and the average student ...by a third in the first year of the program and 40 percent by the third year in order to be detected by the study."[9]

Such gains are rarely achieved by education reforms under the best school conditions, and could not reasonably be expected within three years in the deeply troubled SIG schools. "The difference between the report's conclusion that there was no effect and the more appropriate conclusion that they were not able to detect an effect is an important one," Loeb noted.[10]

By contrast, a recent federally funded study of the highly regarded KIPP schools touted significant growth of only two-thirds of the effect size that Loeb wrote was needed for likely significance in the SIG study.[11] Likewise, a 2017 multi-year evaluation of the New Teacher Center's teacher induction model found significant improvements equal to only a third of the SIG benchmark.[12]

The regression discontinuity approach, comparing just the schools on either side of the threshold, also had the effect of leaving out some of the lowest-performing schools that received help under SIG. In at least one state, these schools turned out to have some of the biggest student achievement gains. A pair of North Carolina studies illustrates the problem. One, using the regression discontinuity methodology that omitted some of the lowest performing schools, found no effects from the program.[13] A second study using a different methodology, which allowed researchers to examine results from all SIG schools, found significant gains when compared to the control group.[14]

Perhaps due to the lack of a common eligibility threshold for SIG schools across states, the IES-funded study also misidentified substantial numbers of schools.[15] As the authors acknowledge, an estimated 15 percent of the schools they included in the SIG cohort ultimately didn't receive grant funding, while 10 percent of the schools in the comparison group did receive SIG support.[16] While some misidentification is not uncommon when using the regression discontinuity method, the proportion of misidentified schools in the IES-funded SIG study clearly exceeds the 5 percent limit experts suggest for obtaining reasonable treatment estimates.[17]

Although the Mathematica and AIR researchers acknowledged the positive results of a few of the state and local studies we analyzed, they stuck by the accuracy of their own findings—that SIG made no difference in student achievement.[18] They also assert their findings could be applied more generally, in part

## Characteristics of Schools Receiving SIG Support and Schools in Federally Funded Study

| School Characteristics | Schools Receiving SIG Support | Schools in Federally Funded Study |
|---|---|---|
| **Setting** | **%** | **%** |
| ▪ Urban | 53 | 88 |
| ▪ Suburban | 24 | 7 |
| ▪ Town or rural | 23 | 6 |
| **Race/Ethnicity** | **%** | **%** |
| ▪ White | 27 | 9 |
| ▪ Black | 42 | 55 |
| ▪ Hispanic | 27 | 31 |
| ▪ Asian | 2 | 2 |

SOURCE: National Center for Education Evaluation and Regional Assistance

because they were based on a broader sample of schools from 60 school districts in 22 states.

Yet that argument does not hold up to close scrutiny. The researchers did not draw a random sample of SIG schools. As a consequence, schools from urban districts were significantly overrepresented in their study: 88 percent of schools in the IES-funded study came from urban school districts, compared to only 53 percent in the SIG program as a whole. The authors of the study acknowledged this misalignment, noting, "Because the SIG sample is not representative of [SIG] schools nationwide the findings here may not apply to all [SIG] schools nationally."[19]

## State and Local Studies

In light of these weaknesses in the federally funded study, the question becomes how strong were the state and local studies that found significant improvements in student achievement or graduation rates in SIG schools? Twelve of the 17 studies show statistically positive outcomes, including all of the district-level reports. Among the five studies with no significant findings, four came from states in which other studies of SIG schools were positive. This left only one state, Michigan, without a study showing a statistically significant positive outcome.

Three of the state and local studies—of SIG programs in California, Ohio and North Carolina—used the same regression discontinuity methodology that the IES-funded study deployed. In California, researchers reported gains that reduced by almost a quarter the gap between SIG school scores and the state's academic targets in the first year of the grant.[20] In Ohio, researchers found positive effects in both reading and math by the second year of the program, effects that grew larger in its third year, when they reached about twice the size of the estimated first-year effects of the California SIG intervention.[21] Only in North Carolina was there no significant improvement, though a separate report on the state's SIG schools using a different methodology showed significant gains in math.[22] In Houston, where a researcher used treatment and control schools chosen randomly for the elementary grades, the study found positive SIG math gains in two years.[23]

The other 13 studies used what's known as a difference-in-differences methodology, a respected approach, but one that has some limitations. Among them is the assumption that trends in SIG and comparison schools would have been the same without the grant program and that there are not other factors influencing SIG outcomes. But the method offers an important advantage in including all SIG schools in its calculations, a feature that produces a fuller picture of the impact of the SIG interventions.[24]

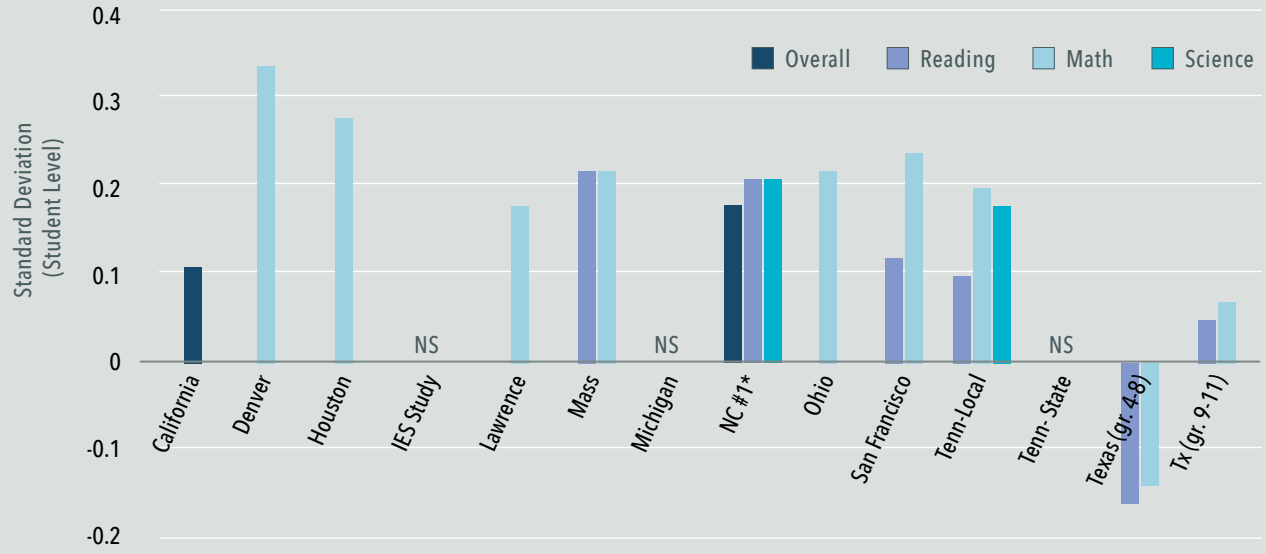
Of the studies using that methodology, nine found significant student-achievement gains for SIG schools when compared to other schools. These improvements averaged nearly .2 of the standard deviation of students when compared to similar schools.[25] For a typical 3rd or 4th grader, this effect represents 40 percent of the expected yearly gain in reading achievement and half the expected gain in math.[26]

Three of the state and local studies also included high school graduation rates in assessing gains. Two of the three observed significant gains.[27]

By their very nature, these state and local studies are not nationally representative. But they do represent a geographically diverse group of nine states and four school districts. That includes two of the largest states, California and Texas, as well as several large states such Ohio, North Carolina, Michigan, New Jersey and Massachusetts. We found no studies from small states. The urban districts that studied SIG results range from very large (Houston) to the small (Lawrence, Mass). Altogether, the state and local studies reflect results from 449 SIG schools, more than twice the 190 SIG schools studied in the IES-funded report.

# SIG'S IMPACT ON STUDENT ACHIEVEMENT

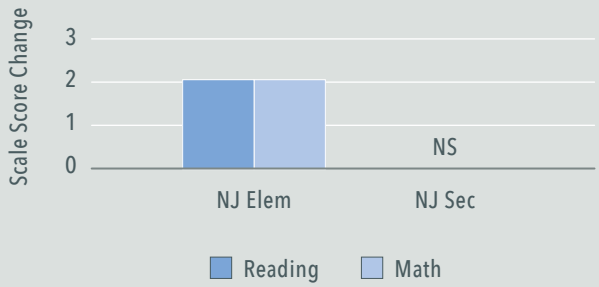
## Net Gains in Test Scores at SIG Schools Measured in Standard Deviations



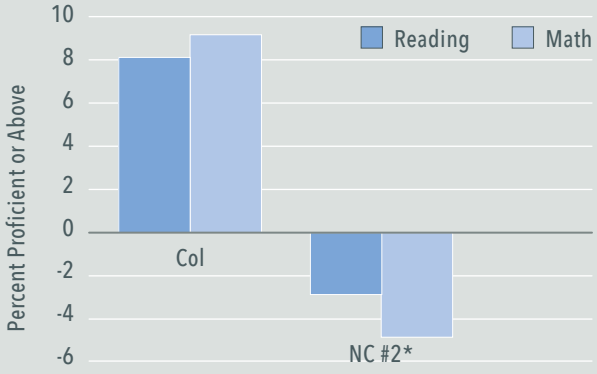

In each of the 17 studies, net gains are the average score gain of the SIG schools minus the average score gain of the comparison schools. The 17 studies used a variety of measures to assess student achievement. Thirteen measured net gains expressed as "effect sizes," where standard deviations are the unit of measurement. A standard deviation is a number used to tell how scores for a group are spread out from the average. Three studies assessed the proportion of students who scored proficient or higher on state tests. Two examined raw test scores and two looked at changes in high school graduation rates.

SOURCE: State and local studies detailed in Appendix (page 11)

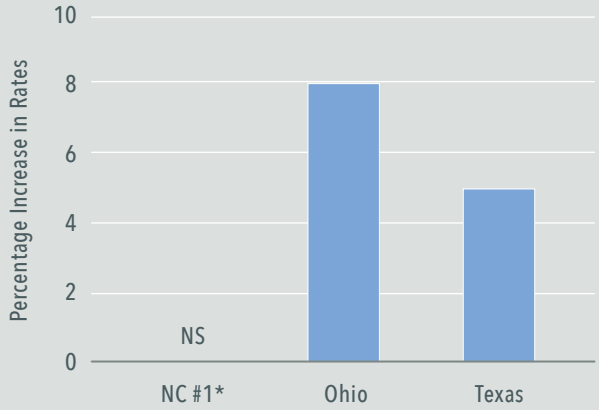
## Net Gains Measured in Test Scores




## Net Gains Measured in the Proportion of Students Scoring Proficient or Above




## Net Gains in High School Graduation Rates



## Learning Lessons

In addition to providing a clearer picture of the impact of the SIG interventions on student achievement and graduation rates, the state and local studies in many instances provided valuable answers to the important question of which school-level improvement practices make a difference.

The Education Department required SIG schools to implement one of four improvement strategies. The most frequently adopted and least invasive intervention was "transformation," the replacing of a failing school's principal a model selected by 75 percent of SIG schools. "Turnaround," used by 20 percent of SIG schools, called for replacing principals and at least half the teachers in failing schools.

The "restart" model, adopted by 5 percent of the schools in the program, required turning the school into a charter or bringing in an education management organization. The fourth option, "closing," which entailed shuttering the school and sending students elsewhere, was rarely chosen.

The IES-funded study and four state and district studies compared the impact on test scores and graduation rates of the most commonly used SIG strategies. Studies in California, Ohio and the San Francisco Unified School District found the turnaround model to be more effective than the transformation model, while a New Jersey report found the transformation model was more effective than turnaround.

The federally funded researchers compared the impact of the turnaround and restart models with the transformation model and also conducted separate comparisons at the elementary and high school levels. They found the turnaround and restart models performed best at the secondary level, while there was no difference at the elementary school level. Overall, these findings suggest that it would be useful to understand the conditions under which each of these models is effective.

The federal program also mandated that states and school districts provide SIG schools with improvement practices beyond staffing changes, ranging from increased learning time to increased parental involvement. States and districts were also responsible for providing SIG schools with experts to monitor and support their improvement efforts.

**The state and local studies in many instances provided valuable answers to the important question of which school-level improvement practices make a difference.**

The IES-funded study gathered data on the implementation of these specific improvement practices only in the second and third years of the SIG interventions, providing an incomplete picture of new practices beyond merely replacing principals and teachers. As a result, the study couldn't reliably gauge how much SIG schools had changed or strengthened their practices under the grant program.

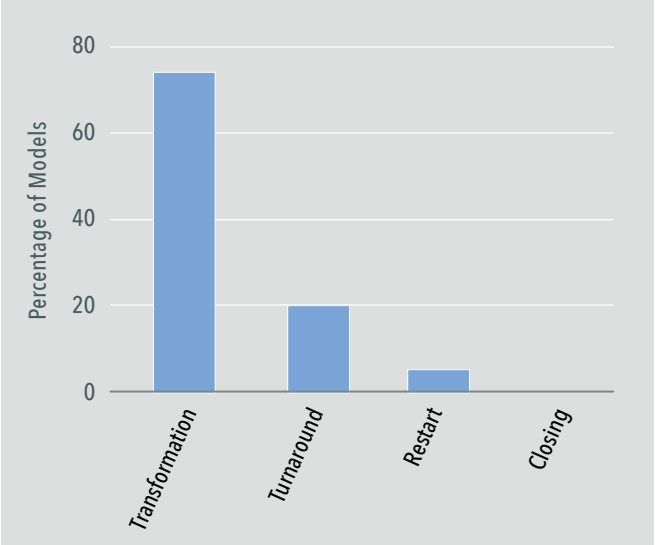
Compounding the problem, the authors of the IES-funded study often simply counted changes in the number of practices, without reflecting on the scope or quality of the reforms. They described "extended learning time" simply in terms of whether schools offered block scheduling, before- or after-school instruction, weekend instruction, or summer instruction, without identifying the numbers of students served or the additional learning time provided under each approach.[28]

Under a separate federal grant, AIR compiled case studies of several SIG practices.[29] Unfortunately, the AIR researchers did not use the information on changes in student achievement in SIG schools for the case studies. As a result, AIR couldn't gauge the effectiveness of various evidence-based improvement practices or contrast what was happening in high-performing and low-performing SIG schools, shortcomings that diminished the value of the case studies.

As an example, AIR's case studies surveyed school personnel about their perceptions of the professional assistance that SIG schools received and found some suggestion that support provided by school districts on average appeared less useful than support from external professional organizations, including universities and charter management organizations. This finding might have had more meaning if matched with SIG school outcomes.

In contrast, several of the state and district SIG studies demonstrated the effectiveness of different improvement practices. Adopting methods from the University of Chicago Consortium of Chicago School Research, the San Francisco Unified School District found that practices focused on student and parent engagement—which included extended learning time and early warning systems for students, and community workshops for parents—contributed to a reduction in unexcused absences and a greater likelihood that families would choose SIG schools.[30] The district also found that changes to hiring and evaluation practices, including one-on-one coaching and interim assessments, contributed to retaining more effective teachers in math and reading.[31]

**Percentage of Improvement Models Chosen by SIG Schools**




SOURCE: IES-Funded SIG Study

Studies in Denver, Houston, and Lawrence, Mass., showed sizeable SIG gains in math using the five strategies that Harvard economics professor Roland Fryer, Jr. gleaned from practices in successful charter schools: increased instructional time, a more rigorous approach to building human capital, high-dosage tutoring, frequent use of data to inform instruction, and a culture of high expectations.[32]

In North Carolina, researchers learned that the required replacement of principals had no effect on teachers' perceptions of the quality of leadership, perhaps because many of the new principals were inexperienced.[33] Such findings point to the importance of understanding how interventions work within local contexts.

Many of the state and local SIG studies also revealed trends in improvement, showing a pattern of gradual rather than sudden gains. In all six of the studies measuring interim gains over three years, improvement was greater in year two than in year one.[34] However, in three of those studies, SIG school achievement levels were not substantially better in year three than in year two, suggesting the importance of understanding the keys to sustaining growth beyond initial gains.[35]

## A New Federal-State Evaluation Partnership

The failure of the IES-funded study to capture any significant improvement in SIG schools and its inability to evaluate specific school-level improvement practices suggest that the U.S. Education Department needs to rethink its traditional approach to evaluating the impact of federal programs. The SIG evaluation was released in 2017, seven years after the grants were first awarded and four years after they expired.[36] Now that most states annually gather extensive education data down to the classroom level, federal program evaluations could be designed to take better advantage of available state data and to support continuous improvement in schools while simultaneously evaluating longer-term impacts.

Such changes would be particularly important as states begin assessing their own approaches to lifting up struggling schools and students. The latest version of the major federal elementary and secondary education law allows states to set their own accountability metrics and determine how to administer school improvement. Rather than just issue a single national study, the Education Department could add more value by aggregating the results of state and local studies, providing states with timely information to improve schools and synthesizing what research tells us about school improvement.

There are legitimate questions of whether the SIG program represented the best way to use federal funding to improve struggling schools. But it is wrong to suggest that there was no return on the SIG investment. More broadly, the SIG experience suggests that a single, nationwide analysis is not the best way to evaluate a program playing out differently in schools and districts across the country. State and district studies, the SIG results suggest, paint a richer picture of how effectively federal education dollars are spent.

# RECOMMENDATIONS

Policymakers could take three key steps to develop a new evaluation approach that provides information for frequent tracking of school progress and building the improvement capacity of states and districts.

1. **The U.S. Department of Education should use the data it already collects from states on school outcomes to generate state-specific studies of school improvement efforts.**

The federal government could support more effective evaluations of improvement programs with its existing data resources. EDFacts, the Education Department's data-collection arm, already gathers a rich body of data.[37] It provides a central repository of student-performance information submitted by states and combines the information with other state-submitted school data, including student population characteristics and Title I and special education participation.[38]

The Education Department could build on this repository with the annual production of state progress reports on struggling schools receiving school improvement funds under ESSA. Reports could also be generated on districts and the state and district studies could be aggregated to form a composite national picture.

This strategy could produce a less expensive, more rapid, and more informative system for tracking and reporting the progress of nationwide education interventions.

*continued*

## RECOMMENDATIONS *continued*

**2. The Education Department could help states use the results of program evaluations to strengthen school-improvement efforts.**

A January 2018 review of state ESSA plans by the nonprofit Results For America found that only four states expect to study their use of federal school improvement funds "to deepen understanding of what works and build an evidence base for school improvement."[39] Yet, the states have unique knowledge and often rich longitudinal data to identify what is working and why in their states.

The authors of the Ohio Education Research Center SIG study, Stephane Lavertu and Deven Carlson, made a similar recommendation, urging researchers to consider "conducting separate evaluations across [states] and then attempting to identify the contextual factors associated with policy success or failure after a sufficient number of evaluations have been conducted."[40]

There are several resources federal education officials could use to help states and districts do this work, including EDFacts state coordinators, the department's 10 regional education laboratories, and its Center on School Turnaround, run by the research organization WestED.[41] Further, as three state research directors propose, IES could make grants available to states to form research partnerships focused on state needs. The onus would then be on states to convince IES of the appropriateness of the research.[42]

**3. The Education Department should initiate a National Research Council (NRC) committee to synthesize the research in the United States and internationally on turning around low-achieving schools.**

In addition to generating new evaluation information on school improvement, an NRC synthesis would bring together in one place the best practices for improving low-performing schools, helping to inform the next generation of school-improvement work under ESSA.

Much has been learned in the decade since IES put out its 2008 guide, Turning Around Low-Performing Schools, which lamented the low quality of research supporting the institute's four improvement recommendations.[43] New resources for research continue to emerge.

Some state education agencies are launching independent reviews of their policies and programs.[44] District-level research consortia are tackling the practical challenges of school improvement in Baltimore, Chicago, Houston, Kansas City, Los Angeles, Philadelphia, New Orleans, New York City, and San Diego.[45] At the federal level, the "Investing in Innovation" program has spent nearly $1 billion since 2010 on learning how to improve low-performing schools.[46]

The NRC review would make sense of the growing academic literature on turning around low-performing schools, including the largely untapped international literature on how other countries have successfully addressed school improvement. Korea and Canada, for example, have found that collaborations between high-achieving schools and struggling peers with comparable demographics is a successful strategy.[47]

Collectively, these steps would encourage the widespread use of evidence-based improvement systems under ESSA.[48] They would also ensure that valuable education resources were being used effectively to improve academic success for our most vulnerable students.

# APPENDIX

## IES and State and District Studies of School Improvement Grants: Methods, Samples and Effects

| Jurisdiction | Causal Methodology | Impact Sample | Effect on Achievement and HS Graduate Rate |
|---|---|---|---|
| Institute of Education Sciences (IES)[1] | ▪ Regression discontinuity to estimate SIG impacts<br>▪ Note, Correlation analysis used to estimate gain scores of SIG schools (not net impacts) by type of intervention model (e.g.transformation). | ▪ SIG treatment for 2011, 2012, & 2013<br>▪ SIG treatment sample (regression discontinuity): purposively selected sample of 190 SIG award schools in 2011 from 22 states and 60 districts<br>▪ SIG treatment sample not nationally representative of SIG schools (overrepresented urban districts, under-sampled white students)<br>▪ Regression discontinuity comparison sample is of schools not eligible for SIG awards in 2011 and near the treatment estimated cutoff | ▪ 3 year regression discontinuity SIG net outcomes<br>– No significant impact on math or reading test scores<br>– No significant impact on high school graduation rate |
| California[2] | ▪ Regression discontinuity | ▪ SIG treatment for one-year 2011<br>▪ SIG treatment sample: 82 Tier 1 or Tier 2 SIG award schools among the lowest-achieving 5% and showing lack of achievement progress. The lowest achieving cutoff was set within grade-level strata of elementary, middle and secondary.<br>▪ Comparison sample: Localized comparison of schools that were above and near the treatment eligibility cutoff | ▪ One year regression discontinuity on California Academic Performance Index (API) covering core subjects<br>▪ Impact on API: +.32 school-level std. dev sig (about .11 student level std. dev)<br>▪ Reduced by 23% the size of the SIG school gap below the state school performance target (800 on California Academic Performance Index)<br>▪ Achievement growth significantly greater in turnaround schools than transformation or restart schools<br>▪ Teacher experience in SIG schools dropped about two years (two-thirds standard deviation) |
| Colorado[3] | ▪ Difference-In-differences | ▪ SIG treatment sample for cohort 1: 2010 baseline to 2014; Cohort 2: 2011 baseline to 2014<br>▪SIG treatment sample: 19 SIG award schools<br>▪ Comparison sample: Title I schoolwide, excluding SIG schools | ▪ Three-year difference-in-differences in percent at least proficient on State assessment compared to Title I schoolwides excluding SIG schools:<br>– Math: +9 percentage points (significance computed by this report)<br>– Reading: +7 percentage points; (significance computed by this report) |
| Denver Colorado[4] | ▪ Difference-In-differences | ▪ SIG treatment for 2012 and 2013<br>▪ Seven treatment schools with tested grades 3, 4, 6, and 9.<br>▪ Comparison sample: all other students in the Denver Public Schools in the same grades as treatment grades | ▪ Annual net outcome change for each of two treatment years (ordinary least squares)<br>▪ Math: +.17 std dev (significant); two-year net outcome +.34 std dev (significant)<br>▪ Reading: +.08 std dev (not significant) |

# APPENDIX *continued*

| Jurisdiction | Causal Methodology | Impact Sample | Effect on Achievement and HS Graduate Rate |
|---|---|---|---|
| Houston Texas ISD (Apollo 20)[5] | ▪ Experimental design<br>▪ Quasi-experimental design (Difference-In-differences ) | ▪ SIG treatment for 2011,<br>▪ SIG treatment sample: 20 (11 elementary and 9 secondary) of the lowest performing schools.<br>▪ 16 elementary schools randomly assigned: 8 to treatment and 8 to no-treatment | ▪ Annual net outcome change for each of two treatment years (ordinary least squares)<br>▪ Experimental elementary<br>– Math (intent to treat): +.14 std dev (significant); two-year net outcomes +.28 std dev<br>– Reading (intent to treat): +.04 std dev (not significant)<br>▪ Quasi-experimental elementary<br>– Math: +.20 std dev (significant); two-year net outcomes +.40 std dev<br>– Reading: +.08 std dev (signifcant); two-year net outcomes +..16 std dev<br>▪ Quasi-experimental secondary<br>– Math: +.10 std dev (signficant); two-year net outcomes +.20 std dev<br>– Reading: -.01 std dev (non significant)<br>▪ Effect of high intensity tutoring<br>– Elementary: +.09 std dev (experimental); +.06 std dev (quasi-experimental)<br>– Middle & Secondary: +0.09 std dev (quasi-experimental) |
| Lawrence, Mass. Public Schools[6] | ▪ Difference-in-differences | ▪ SIG treatment 2013 and 2014 (Part of Massachusetts state takeover of Lawrence school district for chronic underperformance)<br>▪ SIG treatment sample: all students in 28 schools in the Lawrence district taken over by the state<br>▪ Comparison: students in districts with a majority of low-income students | ▪ Two-year difference-in-differences estimates covering state assessments grades 3-8 and 10:<br>▪ Math: +.17 to .19 std dev (significant)<br>▪ Reading: .02 to .03 std dev (not significant) |
| Massachusetts[7] | ▪ Difference-in-differences (Comparative interrupted time series) | ▪ SIG treatment for: 2011, 2012, & 2013<br>▪ SIG treatment sample: 47 schools receiving Federal SIG funds through State Redesign Grants<br>▪ Comparison: schools in same district and serving about the same grade range as SIG schools but not receiving a SIG grant. | ▪ Three-year difference-in-differences in SIG schools on state assessments in grades 3-8 & 10:<br>▪ Math :+22 std dev sig (yr 1 =0.17 and yr 2 = 0.21)<br>▪ Reading: +22 std dev sig (yr 1=0.16 and yr 2 = 0.20)<br>▪ English language learners in SIG schools closed achievement gap in math and reading |
| Michigan[8] | ▪ Difference-in-differences (estimated using hierarchical linear modeling – HLM) | ▪ SIG treatment for: 2011, 2012 & 2013<br>▪ Treatment: 23 schools that received a SIG for three years<br>▪ Comparison: All SIG eligible schools not receiving an award | ▪ Three-year difference-in-differences using HLM on state assessments grades 3-9<br>▪ Grades 3-8: No statistically significant impacts in scores on MEAP in math, reading or science<br>▪ HS (primarily grade 11): No statistically significant impacts in scores on MME in math, reading or science |
| New Jersey[9] | ▪ Multiple regression | ▪ SIG treatment for: 2011, 2012, 2013 & 2014<br>▪ SIG treatment sample: All 20 schools receiving SIG funding<br>▪ Comparison: All eligible schools not receiving SIG funding | ▪ Three-year multiple regression on state assessment scores controlling for students demographics:<br>▪ Elementary & middle SIG schools: small significant gains of about +2 points on math and reading scale scores (effect sizes not shown)<br>▪ High schools: No significant differences on math or reading assessment |

## APPENDIX *continued*

| Jurisdiction | Causal Methodology | Impact Sample | Effect on Achievement and HS Graduate Rate |
|---|---|---|---|
| North Carolina #1[10] | ▪ Two methodologies compared: regression discontinuity and difference-in-differences (The difference-in-difference reflects the full SIG sample, while the regression discontinuity only reflects schools close to the cutoff.) | ▪ SIG treatment for 2011 to 2014<br>▪ SIG treatment sample: All 118 schools below bottom 5%.<br>▪ Comparison: non-treated schools whose 2010 performance composites were closest to the assignment threshold (next-lowest performing schools). | ▪ The very lowest performing SIG recipients performed differently over time than the SIG recipient schools.<br>▪ Difference-in-differences in final SIG year on student proficency.<br>– Overall school proficiency rates: 18 std. dev (significant); (yr 1 = 0.09 std dev not sig; yr 2 = .16 std dev sig)<br>– Math: +.21 std dev (significant)<br>– Reading: +.08 std dev(not significant)<br>– Science: +.21 std dev (significant)<br>– Graduation rate: .29 std dev (not significant but small HS sample of 17 schools)<br>– Estimates included three years of pre-SIG grant performance data to "mitigate potential positive bias of regression to the mean." (p.11)<br>▪ Four-year regression discontinuity had no significant overall effect on school proficiency rates |
| North Carolina #2[11] | ▪ Regression discontinuity | ▪ SIG treatment for 2011 to 2014<br>▪ Treatment sample: 85 K-8 schools receiving an award below: bottom 5% and around regression discontinuity cut off band.<br>▪ Comparison: non-treated schools whose 2010 performance composites were closest to the assignment threshold (next-lowest performing schools). | ▪ Four-year regression discontinuity on state assessments for K-8 grades:<br>– Math passing rates: overall -5.1 percentage points and significant negative effects for female and Hispanic students<br>– Reading passing rates: overall -3.2 percentage points and significant negative effects for Black students<br>▪ Increased concentration of low-income students in SIG schools<br>▪ Substantial increase in principals with three or less years of experience and no effect on teachers' perceptions of the quality of their schools leadership<br>▪ Some delayed (after first year) increase in teacher turnover, but not in inexperience |
| Ohio[12] | ▪ Regression discontinuity (student achievement)<br>▪ Difference-in-differences (HS graduation) | ▪ SIG treatment for: Cohort 1 2011-2014; Cohort 2: 2012-2014<br>▪ SIG treatment sample: 25 SIG I schools identified and included in the analysis<br>▪ Comparison: Schools near the performance cutoff not receiving a SIG grant | ▪ Regression discontinuity on state assessments measuring cumulative improvement in student achievement in year three (avg of cross-sectional and dynamic model estimates).<br>– Math: +.24 std dev (significant)<br>– Reading: +0.24 std dev (significant)<br>▪ Difference-In-differences for graduation rates:<br>– HS grad unit rate +7 to +9 percentage points<br>▪ Significant positive effect on school spending $1,500-$3,000 per pupil depending upon grant year (per pupil grants varied by year)<br>▪ Achievement effects primarily attributable to turnaround model.<br>▪ No significant effects on teacher turnover rates |

## APPENDIX *continued*

| Jurisdiction | Causal Methodology | Impact Sample | Effect on Achievement and HS Graduate Rate |
|---|---|---|---|
| San Francisco[13] (District) | ▪ Difference-in-differences | ▪ SIG treatment for 2011, 2012, 2013<br>▪ SIG treatment sample: nine SIG funded schools in San Francisco<br>▪ Comparison: non-SIG schools in San Francisco (Note: a smaller sample of non-SIG schools with SIG eligibility produced similar results) | ▪ Difference-In-differences in student achievement for all starters after 3 years:<br>– Math: Yr 3 +24 std dev (sig) (yr 2 +.14; yr 1 +.12)<br>– Reading: .increase 12 std dev (sig) (yr 2 +.07; yr 1 +.09)<br>– Reduced gap between SIG and non-SIG schools from .8 std dev to .5 std dev.<br>– Narrowed gap between SIG schools and rest of San Francisco schools by 3% (from, .80 std dev to .50 std dev)<br>▪ Turnaround schools outperformed transformation in math, but similar scores in reading<br>▪ Difference-in-differences in student achievement for all stayers after 3 years are somewhat larger than for all starters<br>▪ Achievement improvement mediators included reduced unexcused (not excused) absences, increased family preferences for schools, improved retention of effective teachers, and and greater development of teacher professional capacity |
| Tennessee[14] | ▪ Difference-In-differences | ▪ SIG treatment for 2013-2015<br>▪ SIG treatment sample: 42 Schools in two distinct strategies – (1) 26 schools run by district through establishing locally run district-within-a-district Innovation Zones (iZone), with substantial extra pay incentives for teachers; (2) state-run Achievement School District (ASD) ( 5 schools directly run by ASD & 11 schools under ASD but run by Charter Management Organization<br>▪ Comparison: similarly low-performing schools (lowest 5%) not receiving an ASD or iZone | ▪ Three-year difference-In-differencess effect sizes on state assessments math, reading, science<br>– Local Innovation Zones IZones (t26 schools): Math .20 std dev (significant), Reading .10 std dev (significant), and Science .18 std dev (significant)<br>– Achievement School District (ASD) math, reading, and science not signicant overall and for state-run and charther management run idirect (5): Math .07 std dev (not sig)<br>– IZones net outcomes were significant each year but showed no pattern of increase or decrease |
| Texas[15] | ▪ Difference-in-differences | ▪ SIG treatment only for first year of implementation 2011<br>▪ SIG treatment sample: 65 SIG schools in first cycle – 53 of the 65 SIG recipients were high schools<br>▪ Comparison: All primary and secondary schools in Texas not receiving a SIG grant with Texas state outcomes data | ▪ One-year difference-In-difference results developed by grade for math and reading average scores and percent proficint (Study presents net outcomes only for each grade between 3 to 11 and not for grade groupings or for all students)<br>– Elementary/middle grades: mostly many negative effects some approaching one-fifth standard deviation.<br>– HS. Mostly lpositive effects although smaller( .1 std dev).<br>▪ HS graduation rates improved by a significant five percentage points |

# APPENDIX *continued*

## CITATIONS


[1] Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., Graczewski, C., Boyle, A., Upton, R., Tanenbaum, C., & Giffin, J. (2017). School Improvement Grants: Implementation and Effectiveness (NCEE 2017- 4013). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Online: https://ies.ed.gov/ncee/pubs/20174013/.

[2] Dee, T. School Turnarounds: Evidence from the 2009 Stimulus. Cambridge, MA: National Bureau of Economic Research, 2012. Online: http://www.nber.org/papers/w17990.

[3] Colorado Department of Education (2015). Evaluation of SIG Schools (Cohorts 1 and 2). Unit of Federal Programs. Available May 2017. Online:http://www.cde.state.co.us/fedprograms/brief-report-for-cohort-1-and-2-sig-schools.

[4] Roland G. Fryer Jr. "Injecting Charter School Best Practices into Traditional Public Schools: Evidence From Field Experiments." Quarterly Journal of Economics. 2014;129 (3) :1355-1407. Online: https://scholar.harvard.edu/fryer/publications/injecting-charter-school-best-practices-traditional-public%C2%A0schools%C2%A0evidence-field. (Note: Denver Public Scholos used $3.8 million SIG funds in 2012 and 2013 to support school improvement project, as reported by Center for American Progress (2018) Strategies to Improve Low-Performing Schools Under the Every Student Succeeds Act. p.21.).

[5] Roland G. Fryer Jr. Injecting Charter School Best Practices into Traditional PublicSchools:Evidence From Field Experiments*. Quarterly Journal of Economics (2014). 2014;129 (3) :1355-1407. Available April 2018: https://scholar.harvard.edu/fryer/publications/injecting-charter-school-best-practices-traditional-public%C2%A0schools%C2%A0evidence-field. The Houston intervention (Apollo 20) used federal SIG funds to fund over half the proprams cost. (Center for American Progress (2018) Strategies to Improve Low-Performing Schools Under the Every Student Succeeds Act. p.20.).

[6] Schueler, B., J. Goodman, and D. Deming (2016). Can States Take Over and Turn Around School Districts? NBER Working Paper No. 21895. January 2016. JEL No. I20,I21,I28. Also, Evidence from Lawrence, Massachusetts Lawrence schools received over $7 million in school redesign grants funded through the federal School Improvement Grant (SIG) program (Center for American Progress (2018) Strategies to Improve Low-Performing Schools Under the Every Student Succeeds Act. p.21.).

[7] LiCalsi, C., M. Citkowicz, L. B. Friedman, and M. Brown (2015). "Evaluation of Massachusetts Office of District and School Turnaround Assistance to Commissioner's Districts and Schools: Impact of School Redesign Grants." Washington, DC: American Institutes for Research. Onlne: http://www.air.org/resource/evaluation-massachusetts-district-and-school-turnaround-assistance-impact-school-redesign.

[8] Rice, John, J. Carlos Bojorquez, M. Diaz, S. Wend, &t J. Nakamot (2014).Evaluation of Michigan's School Improvement Grant: Outcomes After Three Years.WestED. Available May 2017 online: https://www.michigan.gov/documents/mde/Final_MDE_SIG_Outcome_After_Three_Years_483504_7.pdf.

[9] Kyse,Eden N.,R. Swann-Jackson, J. Marini, J. Benton, A. Byrne, A. Sceppaguercio,& K. Wilson.(2014). Final Evaluation Report for the School Improvement Grant (SIG) Evaluation: Summary and Recommendations (2014). Montclair State University. Online: http://www.nj.gov/education/title1/sig/EvaluationReportC1and2.pdf.

[10] Henry, Gary T., J.Guthrie & L. W. Townsend, (2015). Outcomes and Impacts of North Carolina's Initiative to Turn Around the Lowest-Achieving Schools. Consortium for Educational Research and Evaluation–North Carolina. Online: http://cerenc.org/rttt-evaluation/district-and-school-transformation/.

[11] Heissel, Jennifer A. and H. F. Ladd (2016). School Turnaround in North Carolina: A Regression Discontinuity Analysis. National Center for the Analysis of Longitudinal Data in Education Research (CALDER). Working Paper 156. Online: http://www.caldercenter.org/sites/default/files/WP%20156.pdf.

[12] Student achievement study by Carlson, D. and S. Lavertu (2018). "School Improvement Grants in Ohio: effects on student achievement and school administration." Educational Evaluation and Policy Analysis, forthcoming; High school graduation study by Carlson, D. and S. Lavertu (2017). "Ohio School Turnaround Interventions: Impacts on Achievement, Attainment, and Administration" in Munn, S., J. Lindsey, D. Carlson and S. Lavertu. Evaluating the Ohio Improvement Process. Ohio Education Research Center. Winter, 2017.

[13] Sun, Min, Emily K. Penner, Susanna Loeb (2017). "Resource- and Approach-Driven Multidimensional Change: Three-Year Effects of School Improvement Grants." In American Educational Research Journal. Forthcoming. Online early publication: DOI: 10.3102/0002831217695790.

[14] Zimmer, Ron, G. T. Henry & A. Kho (2016). "The Effects of School Turnaround in Tennessee's Achievement School District and Innovation Zones." Educational Evaluation and Policy Analysis Volume: 39 issue: 4, page(s): 670-696. Online: http://journals.sagepub.com/doi/abs/10.3102/0162373717705729.

[15] David Dickney-Griffith(2013). "Preliminary Effects of the School Improvement Grant Program on Student Achievement in Texas" in Georgetown Public Policy Review. Online: http://gppreview.com/wp-content/uploads/2014/02/Dickey-Griffith-D.pdf.

# REPORT ENDNOTES

[1] Hurlburt, S., K. C. Le Floch, S. B. Therriault, and S. Cole. (2011) "Baseline Analyses of SIG Applications and SIG-Eligible and SIG-Awarded Schools." NCEE 2011-4019. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2011. In 2010 the Department of Education made a special one-time $3.0 billion SIG award on top of the .5 billion already appropriated. The Department spent the entire $3.5 billion on SIG programs that ran from 2010 to 2013. Later appropriations brought the total federal investment in the SIG program to over $7 billion. See https://www.air.org/resource/school-improvement-grants-implementation-and-effectiveness.

[2] Dragoset,L., J. Thomas, M. Herrmann, J. Deke, S. James-Burdumy, C. Graczewski, A. Boyle, R. Upton, C. Tanenbaum, and J. Giffin. (2017). School Improvement Grants: Implementation and Effectiveness (Final Report). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Available November 2017 online: https://ies.ed.gov/ncee/pubs/20174013/.

[3] Interview with Education Secretary Betsy DeVos on 60 Minutes, March 11, 2018.

[4] This is based on a minimum detectable effect size of .3 at 80 percent power for the SIG study (Table A.9) by Dragoset, et. al.and an interpretation of effect sizes SOMETHING MISSING HERE from See Table 5 (p.28) in Lipsey, M., K. Puzio., C. Yun, M. Hebert, K. Steinka-Fry, M. Cole, M. Roberts, K. Anthony, M. Busick.(2012)). Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms. (NCSER 2013--33000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. Available https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf. A grade 3-4 student in math has an average annual achievement gain of about .5 standard deviations and a grade 7-8 student in math of about .3 standard deviations.

[5] Dragoset, et. al. (2017). pp.7 & 29.

[6] The study oversampled urban districts because of their higher concentration of schools eligible for SIG and the proximity to comparison schools.

[7] A binomial formula was applied to compute the probability of obtaining at least 12 positive outcomes, assuming the chance of a significant positive outcome was from the same statistical population distribution and was .1 or less. At a probability of .1, the chances of obtaining 12 positive outcomes among 17 outcomes is less than one in a million trials.

[8] For example, the SIG program office sample of the universe of useable outcome data on SIG schools shows at least an eight-percentage point improvement in the percent of students proficient or advanced on state math assessments in just three years 2010 to 2013. Online: (https://www2.ed.gov/programs/sif/data/school-impro-grants-nat-sum-sy1213.pdf.).

[9] Loeb, S. (2017). "Continued Support for Improving the Lowest-Performing Schools." Evidence Speaks Reports. Vol 2, #8, Feb 9. Online: https://www.brookings.edu/wp-content/uploads/2017/02/es_20170209_loeb_evidence_speaks.pdf.

[10] Loeb, S., op. cit . p.1.

[11] In 2010, the KIPP Foundation received a $50 million, multiyear grant from the U.S. Department of Education Investing In Innovation fund to scale-up KIPP schools. A rigorous Mathematica evaluation of KIPP students score improvement relative to controls found three-year standard deviation impacts of 0.21 at elementary schools (experimental design), 0.16 standard deviations at middle school (experimental design) and 0.20 at high school (matched comparison for new students). Tuttle, C., P. Gleason, V. Knechtel, I Nichols-Barrer, K. Booker, G. Chojnacki, T. Coen, and L. Goble (2015). Understanding the Effect of KIPP as it Scales: Volume I, Impacts on Achievement and Other Outcomes. Available August 2018: https://www.mathematica-mpr.com/news/kipp-i3-scale-up.

[12] Schmidt, R. V. Young, L. Cassidy, H. Wang, and K. Laguarda (2017). Impact of the New Teacher Center's New Teacher Induction Model on Teachers and Students. SRI Education. Online: https://www.sri.com/sites/default/files/publications/ntc_i3_validation_eval_brief.pdfThe average of the reading and math gains for students of teachers participating in the New Teacher Center evaluation relative to the experimental comparison over two years was 0.12 standard deviations.

[13] Heissel, J. A. and H. F. Ladd (2016). "School Turnaround in North Carolina: A Regression Discontinuity Analysis." National Center for the Analysis of Longitudinal Data in Education Research (CALDER). Working Paper 156. Online: http://www.caldercenter.org/sites/default/files/WP%20156.pdf.

[14] Henry, G.T., J.E. Guthrie and L. W. Townsend, (2015). Outcomes and Impacts of North Carolina's Initiative to Turn Around the Lowest-Achieving Schools. Consortium for Educational Research and Evaluation–North Carolina. Online: http://cerenc.org/wp-content/uploads/2015/09/0-FINAL-Final-DST-Report-9-3-15.pdf.

[15] The states distributed the money in a variety of ways. In Massachusetts, for instance, 100 percent of SIG grants went to the state's lowest-achieving schools serving low-income students, i.e. schools that receive funding under the federal Title I program that are also among the lowest-performing 5 percent of schools in the state, or high schools with graduation rates under 60 percent. In contrast, in Tennessee upwards of 80 percent of SIG funds went to schools that were less challenged than those in the Massachusetts SIG program, while only 14 percent were in the lowest-performing tier.

[16] Dragoset, et al., 2017, p18. While the regression discontinuity approach is designed to address situations where there is not a sharp award cutoff on an assignment criterion estimation error arises nonetheless from misclassification with the size of the error varying by the percent of misclassified schools in the sample. See Price, Kelci M., "Effects of Misassignment on Bias in the Fuzzy

## REPORT ENDNOTES *continued*

Regression Discontinuity" (2009). Dissertations. Paper 236. Online: https://ecommons.luc.edu/luc_diss/236.

[17] Shadish, D., T. Cook, and D. Campbell (2002). Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Boston, MA.: Houghton Mifflin Co. p.229.

[18] Asked about the study's lack of statistical power, lead author Lisa Dragoset of Mathematica responded in an August 17, 2018, email: "We believe that our study did have sufficient statistical power to detect an effect that would be meaningful from a cost-effectiveness framework, because there are cheaper interventions than SIG that targeted similarly disadvantaged populations and for which studies have found impacts much larger than our minimum detectable effects."

[19] Dragoset.,et. al., p. ES-6.

[20] Dee, T. "School Turnarounds: Evidence from the 2009 Stimulus." Cambridge, MA: National Bureau of Economic Research, 2012. online:: http://www.nber.org/papers/w17990.

[21] Carlson, D. and S. Lavertu (2018). "School Improvement Grants in Ohio: Effects on Student Achievement and School Administration." Educational Evaluation and Policy Analysis, vol. 40, 3: pp 287-315. The math and reading standard deviation effects are an average of the study's year three cross-sectional and dynamic model estimates.

[22] Heissel, et al., 2016.

[23] Fryer, R.G. "Injecting Charter School Best Practices into Traditional Public Schools: Evidence From Field Experiments. "Quarterly Journal of Economics (2014). 2014;129 (3) :1355-1407. Online: https://scholar.harvard.edu/fryer/publications/injecting-charter-school-best-practices-traditional-public%C2%A0schools%C2%A0evidence-field.

[24] The Tennessee difference-in-difference studies tested pre-treatment SIG and comparison school trends for parallelism on scores over time during the period prior to the treatment. The study of Massachusetts used an interrupted time series to estimate the difference-in-difference approach. Some of the state studies further refined their calculations by including student achievement trends for schools prior to receiving SIG support (Massachusetts, North Carolina, Tennessee) or using multi-level modeling (Michigan). On the other hand, Colorado used a simple comparison group of other SIG eligible schools or Title I schoolwide schools and did not adjust for differences in other student characteristics across schools. On balance, perhaps the approach of one of the North Carolina studies of comparing results under both regression discontinuity and difference-in-differences methods is sensible.

[25] The standard deviation is a measure of the typical spread in student test scores from the average score. A greater standard deviation indicates a greater score dispersion. The outcomes reported in standard deviation units are for studies with interventions of different durations. California and Texas had studies lasting for one year. Denver, Houston and Lawrence two years and the rest of the studies were over three years. We decided it was most useful to assess interventions by showing actual outcome changes rather than adjusting the outcomes for different intervention durations.

[26] See Table 5 (p.28) in Lipsey et. al.

[27] Studies in Ohio and Texas showed significant gains in high school graduation rates, while one of the North Carolina studies looked at high school graduation rates and found no significant gains.

[28] Dragoset, et al., 2017, Table F.4, p.F-12.

[29] Cle Floch, K., J. O'Day, B. Birman, S. Hurlburt, M. Nayfack, C. Halloran, A. Boyle, S. Brown, D. Mercado-Garcia, R. Goff, L. Rosenberg, and L. Hulsey. (2016). Case Studies of Schools Receiving School Improvement Grants: Final Report (NCEE 2016-4002). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. pp.75-78. Available November 2017 online: http://files.eric.ed.gov/fulltext/ED565615.pdf.

[30] Bryk, A. S., P.B. Sebring, E. Allensworth, J.Q. Easton, & S. Luppescu. (2010). Organizing Schools for Improvement: Lessons from Chicago. Chicago, IL: University of Chicago Press.

[31] Sun,, et al., (2017).

[32] Fryer, op. cit., Abstract.

[33] Heissel et al. (2016) p.19.

[34] All six studies with year-three significant positive outcomes and that reported annual achievement outcomes found that year-two net outcomes were greater than year one improvements: Colorado, Massachusetts, North Carolina (difference-in differences), Ohio, San Francisco, and Tennessee.

[35] Colorado, Ohio and Tennessee are three states for which year three student achievement outcomes are not consistently greater than year two.

[36] Dragoset, et.al., 2017.

[37] The IES study used annually state achievement test data but only reported SIG school results once, seven years after the study started. The EdFacts Initiative is available at https://www2.ed.gov/about/inits/ed/edfacts/index.html. Also see U.S. Department of Education (2017). EDFacts Workbook SY 2017-18. Version 14. Available March 2018: https://www2.ed.gov/about/inits/ed/edfacts/eden/17-18-workbook-14-0.pdf.

[38] Ibid.

[39] Results For America (2018). ESSA Leverage Points: 50-State Report on Promising Practices For Using Evidence to Improve Student Outcomes. Available March 2018: https://results4america.org/tools/essa-leverage-points-50-state-report-promising-practices-using-evidence-improve-student-outcomes/.

[40] Carlson, D. and S. Lavertu (2018). "School Improvement Grants in Ohio: Effects on Student Achievement and School Administration." Educational Evaluation and Policy Analysis, vol. 40, 3: pp 287-315;

## REPORT ENDNOTES *continued*

High school graduation study by Carlson, D. and S. Lavertu (2017). "Ohio School Turnaround Interventions: Impacts on Achievement, Attainment, and Administration," in Munn, S., J. Lindsey, D. Carlson and S. Lavertu. Evaluating the Ohio Improvement Process. Ohio Education Research Center. Winter, 2017.

[41] West Ed. The Center on School Turnaround. Available November 2017 online: http://centeronschoolturnaround.org/.

[42] Conaway, C., V. Keesler, N. Schwartz (2015). "What Research Do State Education Agencies Really Need? The Promise and Limitations of State Longitudinal Data Systems." Educational Evaluation and Policy Analysis, vol. 37, 1_suppl: pp. 16S-28S. , First Published May 1, 2015.

[43] Herman, R., P. Dawson, T. Dee, J. Greene, R. Maynard, S. Redding, and M. Darwin. (2008). Turning Around Chronically Low-Performing Schools: A Practice Guide. Institute for Education Sciences. Available March 2018 https://ies.ed.gov/ncee/wwc/PracticeGuide/7.

[44] States supporting studies of their SIG interventions reviewed in this report include: Colorado, Massachusetts, Michigan, New Jersey, Ohio, and Tennessee.

[45] For a listing of city-based and other research partnerships see: The Research Alliance for New York City Schools. Available November 2017 online: https://steinhardt.nyu.edu/research_alliance/research/resources.

[46] Source is the website for the Invest in Innovation Fund, accessed at: https://www2.ed.gov/programs/innovation/funding.html.

[47] "In Shanghai (China), policies support collaboration between better- and lower-performing schools with the aim of transferring leadership capacity from the former to the latter. One aspect is called empowered administration, a school-custody programme in which the government asks higher-performing public schools to administer weaker schools. Under this scheme, the high-performing school appoints its experienced leader, such as the deputy principal, to be the principal of the weaker school and sends a team of experienced teachers to lead in teaching. In this way, the ethos, management style and teaching methods of the good schools are transferred to the poorer-performing school. In addition, a consortium of schools is established, where strong and weak schools, old and new, public and private, are grouped into a consortium or cluster, with one strong school at the core."OECD (2013). What Makes Schools Successful? Resources, Policies and Practices. Volume IV. p. 130. https://www.oecd.org/pisa/keyfindings/Vol4Ch4.pdf. For Canada, Successful Focussed Intervention in Low Performing Schools March 26, 2018. Informal note: MaryJean Gallagher and Michael Fullan (not published).

[48] West, M. (2016). "From Evidence-Based Programs to an Evidence-Based System: Opportunities Under the Every Student Succeeds Act." Brookings Evidence Speaks Series. Feb. 5. Available March 2018: https://www.brookings.edu/research/from-evidence-based-programs-to-an-evidence-based-system-opportunities-under-the-every-student-succeeds-act/.

FutureEd
GEORGETOWN UNIVERSITY

# REVISITING SIG

## WHY CRITICS WERE WRONG TO WRITE OFF THE FEDERAL SCHOOL IMPROVEMENT GRANT PROGRAM