

School Interventions That Work:

Targeted Support for Low-Performing Students

July 2017



Table of Contents

School Improvement Under the Every Student Succeeds Act	1
School Needs Assessments	2
Elements of an Effective School Needs–Assessment Tool	3
District Action: Link to Research-Based Practices	4
Putting It All Together: The School Improvement Plan (SIP)	6
Site Visits	7
Conclusion	8
Appendix	9
Endnotes	13

Acknowledgments

*This report was prepared by the **Johns Hopkins Institute for Education Policy** (the Institute) for the Alliance for Excellent Education (the Alliance). Cover photo by **Allison Shelley/The Verbatim Agency for American Education: Images of Teachers and Students in Action**.*

*The Alliance acknowledges **Carnegie Corporation of New York** for its generous financial support for the development of this paper. The findings and conclusions presented are those of the Alliance and the Institute and do not necessarily represent the views of Carnegie Corporation of New York.*

*The **Alliance for Excellent Education** is a Washington, DC–based national policy, practice, and advocacy organization dedicated to ensuring that all students, particularly those traditionally underserved, graduate from high school ready for success in college, work, and citizenship. www.all4ed.org*

*The **Johns Hopkins Institute for Education Policy** (the Institute) believes that building partnerships across different constituencies is necessary to advance excellence and equity for all of America’s children. The Institute operates on the understanding that education policy must be informed both by real-world conditions and also by excellent research; that it is possible to translate the technical vocabularies of research into a language that is accessible and useful to policy experts, principals, teachers, and parents; and that in our richly diverse nation, education must be driven and sustained by evidence about what works and what does not. <http://edpolicy.education.jhu.edu/wordpress/>*

© Alliance for Excellent Education, July 2017.

School Improvement Under the Every Student Succeeds Act

The Every Student Succeeds Act of 2015 (ESSA) gives states and districts much more flexibility in determining how to turn around their lowest-achieving schools. ESSA stipulates that states must identify schools for improvement based on the performance of all students and student subgroups. However, the U.S. Department of Education (ED) no longer may specify particular school intervention models, as it did under ESSA's predecessor, the No Child Left Behind Act (NCLB). Moreover, funds previously set aside for School Improvement Grants (SIGs), a federal program that provided states with funds to support persistently low-achieving schools, now flow through the regular Title I formula.¹

For state and district leaders faced with the challenges of low-performing schools, not much has changed. It is not as if a new intervention strategy now has become available for district and state use. Once states identify the bottom-performing 5 percent of Title I schools, high schools with a graduation rate at or below 67 percent, and those schools with continually underperforming subgroups of students, as stipulated in ESSA, leaders of schools, districts, and states must act. In the case of the bottom 5 percent of schools and high schools with low graduation rates, districts must lead "comprehensive" interventions. In the case of underperformance of student subgroups, schools must lead "targeted" interventions. In both cases, the intervening agent must adopt one or more "evidence-based" strategies, where evidence is defined as a practice supported by the findings of one of three categories of research: randomized control trials (known as Tier 1 research), quasi-experimental (Tier 2 research), or correlational (Tier 3 research).²

ESSA also outlines a broad process for intervening in schools with underperforming student subgroups that references "targeted support and improvement plans," referred to throughout this report as "school improvement plans":³

- Schools, in partnership with stakeholders, must develop and implement an improvement plan for each underperforming student subgroup. This plan must be informed by all data indicators and include evidence-based interventions.⁴

The plan must be approved and monitored by the local education agency (LEA).

- The improvement plan also must identify resource inequities for student subgroups performing at the same level as the lowest-performing 5 percent of schools.
- If the student subgroup fails to improve within a state-determined number of years, the state steps in with additional action.



Photo by Allison Shelley/The Verbatim Agency for American Education: Images of Teachers and Students in Action

Essentially, ESSA envisions targeted intervention to be the purview of schools and districts. Under the law, states outline the parameters used to identify schools with consistently underperforming student subgroups and identify schools for intervention. But schools, in partnership with stakeholders, develop the improvement plans that districts must approve. Districts can apply to the state for funding to support school-level intervention.

There are several key elements of this process for identifying and intervening in schools that likely are required to raise the performance of under-achieving student subgroups. This report explores those steps, starting with a discussion of effective school needs assessments and school improvement plans (SIPs). The appendix offers examples of evidence-based interventions for historically underserved students.

School Needs Assessments

Just as a doctor's prescribed course of treatment fails if the patient's condition has been diagnosed incorrectly, a school intervention will fail if it addresses the wrong problem. The school needs assessment is the diagnosis, and the SIP sets out the treatment. ESSA requires that the "thermometer" used to identify a low-performing student subgroup includes indicators of student academic proficiency and growth or another academic indicator (for elementary and middle schools), high school graduation rates (for high schools), and English language proficiency, together with at least one state-selected indicator of school quality or student success. For instance, data that shows poor performance in a student subgroup's math scores does not alone identify the cause for those low scores, which could range from poor instruction, poor curriculum, chronic absenteeism, poor morale in the school with attendant discipline issues, wrong professional development, lack of timely and disaggregated data, and more.

ED's nonregulatory guidance makes the essential link between the needs assessment and the schoolwide improvement program explicit:

Comprehensive improvement plans must be based on a school needs assessment, include evidence-based interventions, and identify resource inequities. Stakeholders must be involved in planning. In schools identified for

comprehensive support and improvement, ESSA guidance emphasizes the importance that improvement plans address all learners in the school, especially the lowest-achieving students.⁵

The nonregulatory guidance issued by ED further specifies that the school needs assessment must address how the selected school improvement methods and strategies will

- strengthen the academic program in the school;
- increase the amount and quality of learning time; and
- provide an enriched and accelerated curriculum, which includes programs and activities needed to provide a well-rounded education.

States and districts can follow this same process for implementing targeted interventions to ensure that the interventions are matched properly to the specific challenges encountered by the relevant schools. But how should a school develop its needs assessment to ensure, to the greatest degree possible, maximum accuracy and that the needs of all learners are met?

Current Needs-Assessment Practice and Its Limitations

Schools have been doing needs assessments for years, and there is no shortage of guidance on how best to conduct them. Several proprietary tools enable districts to organize data to facilitate analysis and presentation. One example is the "Adaptive System of School Improvement Support Tools™ (ASSIST™)"⁶ from AdvancEd; a second is from Indistar.⁷ Some states have created resources to guide districts in constructing their needs assessments. Virginia,⁸ Arkansas,⁹ and Colorado¹⁰ are among the most comprehensive.¹¹

Unfortunately, while these various tools and state resources help with the needs-assessment process, they cannot get "inside the school" to ensure that the process diagnoses the correct problem. The directions for completing the needs assessment include such phrases as, "Develop the priorities, goals, benchmarks based on disaggregated data, which focuses on student needs," and "Design evidence-based action geared towards enhancing student achievement."¹² These tools also suggest methods for setting priorities without helping districts or states do the prioritizing.

The question at the heart of the process is this: *What exactly is causing the problematic results in student achievement?* This is no mean task. For instance, identifying a general misalignment in eighth-grade math results between the curriculum and the state assessment does not tell a school leader if the problem occurs across the entire academic skill set or within a particular domain such as algebraic reasoning. A remedy that involves a full-scale curriculum replacement might not only be unnecessary, but it also could cause general disruption and a further lowering of student outcomes

In short, the evidence-based intervention supplied by the district and supported by the state can be effective only if the diagnosis of the problem is accurate at a sufficiently granular level to distinguish one or more root cause(s) for students' poor academic performance.

It is true that general data points are a key starting point. For example, high levels of student and/or teacher absences strongly correlate with weaker student results, so finding high levels of either is a warning sign. But the next stage means digging deeper; is there an especially high rate of absenteeism among certain grade levels, or certain student subgroups, and/or at certain times of the year? Different findings point to different interventions.

In the end, a strong needs assessment should produce a fine-grained analysis of what is happening in student learning and an effort to correlate that analysis with adult behaviors. Such an assessment can enable the school to select the intervention that has the best chance of changing adult behavior that research has shown will increase student performance where that performance is most seriously challenged.

Elements of an Effective School Needs–Assessment Tool¹³

Based on the theory of action outlined above, an effective school needs–assessment tool would include at least these four critical elements: (1) organized display of data; (2) granular performance data; (3) theory of action for data collection; and (4) crosswalk between adult actions and student achievement data.

Organized Display of All High-Level, Reliable Data Available to the School

No accurate school needs assessment can be done without relevant empirical information. An effective tool will generate a data-rich profile of students' academic performance with accurate school climate indicators included. All the ESSA-required data, as interpreted by the state, and additional data will be here: (1) multiple indicators of student academic achievement and growth; (2) history of gap-closing results for student subgroups (if available); (3) comparisons to other schools; and (4) full transparency on student subgroup achievement. There is no single “best in class” example needs-assessment tool, but schools in Ohio,¹⁴ New Jersey,¹⁵ and Colorado¹⁶ offer promising examples that require rich data sets to inform the starting point of the needs assessment.

Granular Performance Data

Beyond excellent school-level assessment and climate data, chief state school officers and superintendents need to understand what is going on *inside* the numbers. For instance, what aspects of English language arts or math, exactly, cause students to struggle and how many students struggle with them? What percentage of students belongs to one or more student subgroup? Louisiana offers one example of how a state education agency (SEA) uses additional descriptors to provide schools with information beyond basic performance categories, which SEAs provide to schools in template form. Schools in other states that have access to diagnostic, formative, and interim assessment data can use [Louisiana's example](#) as a model or create their own template. The key principle is the same: schools cannot intentionally improve what they cannot identify.¹⁷

Theory of Action for Data Collection

Once a school has basic student-performance data on hand, the school can dig into the crucial aspects of adult behavior that potentially impact student achievement results. Because there is a potentially unlimited universe of possible data, schools need to set organizing and operating assumptions at the front end to guide their inquiry. The New York State Education Department (NYSED) offers a self-diagnostic tool that contains useful elements to guide this work; however, the tool is purely qualitative. A more helpful model would include a Likert scale

to weight the level of performance in each domain. This would enable a school needs assessment to prioritize the areas for evidence-based interventions.

Instruction is at the core of this school-based adult behavior diagnostic. What are teachers teaching and how effectively are they teaching it and, in both cases, why? An effective needs-assessment tool will prompt granular consideration—by teacher, subject, and grade level—of what school leaders know and do not know about enacted curriculum; degree of alignment horizontally across grades; and extent of vertical alignment (from one grade to the next) in subject matter. The tool also can answer questions about teacher effectiveness. For example, have the principal's observation results been correlated with student academic growth data? What has changed in the last year in terms of instruction—content in subject or grade level, teaching or leadership personnel, key mandates, professional development, and parental outreach and involvement—and what has *not* changed in these domains?

An effective needs-assessment tool will set out what the school thinks it is doing in these domains but restrict the reporting to areas in which data supports the findings.

Crosswalk: Adult Actions and Student Results

The final operation the tool will prompt and support is the crosswalk—the school's best effort to link specific adult behaviors to specific problematic student academic outcomes. With the

assistance of a theory of action/best practices metric, the school will define adult action both positively and negatively, as in "X is doing, or is not doing, A" (where X is one or more adults and A represents a specific action). The key is to avoid impressionistic evaluation, such as "X is not doing A very well." This is a limitation of the NYSED diagnostic tool, which describes teacher actions with words such as "beginning to" rather than in concrete terms such as "in eighth grade, teachers meet for two and one-half hours a week to share their standards-aligned lesson plans for the following week and make modifications to ensure alignments and subject-matter coherence across grade" or "teachers in eighth grade do not share their lesson plans with each other."

No needs-assessment tool fully focuses on the granular data in areas that research shows matter the most for effective instruction, which requires the close alignment of the following critical factors:

- Formal curriculum supported by research that shows a strong student impact (see the next section about researched-based practices)
- Evidence that teachers teach that curriculum and share best practices and challenges regularly (ideally weekly) with granular data on student performance as their key input
- Alignment between regular in-class assessments embedded in that curriculum and more formal formative, interim, and state-mandated summative assessments

District Action: Link to Research-Based Practices

When the needs assessment is done effectively, districts, with possible assistance from their SEAs and regional educational laboratories (RELs), can match research-based interventions to the most critical challenges identified by a low-performing school, following these broad steps:

- **Literature review.** Find relevant, research-based literature that identifies strategies to address the weakness that showed up in the data. Keep in mind that context matters for strategies to succeed. School leaders should compare the research settings to those in their state or district.
- **Benchmarking.** Examine how similar states or districts implemented solutions to handle similar needs.



Photo by Allison Shelley/The Verbatim Agency for American Education: Images of Teachers and Students in Action

- **Multi-attribute utility technique (MAUT).** Develop criteria against which each solution strategy is rated. MAUT results in a summary of ratings for each solution strategy, which school leaders can translate into a list of strategies with the highest potential for addressing their specific need. This is a helpful approach because it requires that school leaders identify the most important criteria for addressing each of their identified needs.
- **Quality function deployment.** Identify key components of program design features and determine the likelihood that they will resolve the needs the school has identified.¹⁹

The capacity of schools and districts to research best-evidenced interventions will vary immensely. Below are some of the best resources to use.

What Works Clearinghouse

[What Works Clearinghouse](#) (WWC) offers a summary of pre-K–12 educational interventions that have been the subject of strong research that matches Tier 1 (randomized control trial) or Tier 2 (quasi-experimental) as defined by ESSA. WWC recently redesigned its web pages to be user-friendly.²⁰ Readers can home in on subject areas, and the site will rank-order interventions according to the research strength that supports them.²¹

Unfortunately, the research base is very thin, or even non-existent, for some student subgroups, including English language learners (ELLs) and students with disabilities. It is only when one turns to interventions that apply to a specific racial group that the WWC becomes helpful. For example, a search for the topics of literacy, high school grades, and “black,” yields ten programs listed as “fully relevant” to this student population. However, drilling down into the actual research studies shows that even this promising list turns out to be modest, with interventions included in which the impact is so minor that one could hardly adopt them with confidence.

Best Evidence Encyclopedia

[Best Evidence Encyclopedia](#) (BEE) is similar to WWC in using a high bar to identify strong research studies. It has launched a user-friendly tool, [Evidence for ESSA](#), that explicitly links the research base for a variety of interventions to the ESSA evidence

requirements. Readers can search for interventions using the same filters as with WWC—subject, grade level, evidence rating, and population served—but with new filters that will enable users to further narrow their search. New filters include objectives (e.g., phonics or reading comprehension), community (i.e., rural, urban, suburban), and distinctive features (e.g., technology, professional development, family focused, or textbook/curriculum).

More importantly, the enhanced BEE includes research findings that qualify for Tier 3 of ESSA’s research standards—interventions defined as “promising,” those correlational studies that cannot generate causal claims but indicate that student outcomes moved in a particular direction after the intervention was introduced. When interpreted with due caution, such studies could result in the inclusion of far more research-based interventions, thus offering users a more real-world tool, especially in domains such as special education and ELL where a paucity of research makes evidence-based interventions difficult to identify. For specific research studies that identify interventions targeted to these student subgroups, see Appendix: Studies on English Language Learners, Students with Disabilities, and Students from Low-Income Families.

Evidence from Best Practices

Finally, school leaders might turn to evidence from best practices in addition to the major clearinghouses and to individual studies for examples of promising targeted interventions. For instance, Claude Goldenberg’s “Teaching English Language Learners: What the Research Does and Does Not Say” lays out the findings from two meta-analyses of research on teaching ELLs.²² However, it is difficult for readers to know how much to rely upon the stated findings, since the level of rigor is variable and would not be included in WWC or BEE.

On the other hand, the fact that an instructional approach has not yet been subjected to rigorous study does not mean it is not worth considering. Schools rightly hesitate to implement an intervention where a study finds a connection between that intervention and gains in scores or high school graduation rates without having controlled for other factors. Nevertheless, when faced by an almost empty ESSA Tier 1 or Tier 2 research record, studies showing robust, multi-year demonstrations of success make the practices worth considering. A few examples follow.

English Language Learners

One such example comes from the record of the International Network for Public Schools (International schools) in New York, where each school is comprised of 100 percent ELLs. Findings from a 2005 study show extremely positive evidence of increased high school graduation rates and low drop-out rates. Among the best practices the researchers report are language and content integration; language skills are most effectively learned in context and emerge most naturally in purposeful, language-rich, experiential, and interdisciplinary study. At the International schools, all learners—faculty and students—experience the same model, which maximizes their ability to support each other.²³

Students with Disabilities

A meta-analysis by Thomas Scruggs from George Mason University and his colleagues²⁴ examines seventy experimental and quasi-experimental studies involving interventions for secondary-grade students with special needs. Their analysis shows that both explicit and mnemonic instruction have positive effects on student achievement.²⁵

Explicit instruction uses practices such as teaching in small steps, rules-based instruction-guided practice, and independent practice to convey learning. Two studies show positive achievement for students' learning under this model. Each study employs a different explicit-instruction technique.

- One study of rules-based instructions for teaching the scientific method finds that students with disabilities receive higher achievement gains than those using more constructivist instructional methods.²⁶
- A separate study²⁷ finds learning gains when students with learning disabilities are taught social studies using smaller instructional units, following each unit with student practice and teacher feedback.

Mnemonic instruction has a more robust research base of twenty-one studies that show potential effectiveness. Mnemonic instruction links new information to prior knowledge through visual and/or acoustic cues. Keyword and peg word are two commonly used approaches within mnemonic instruction. Keyword uses familiar sounds to the word being taught, then uses pictures to link the two words; peg word uses rhyming words to help students remember sequential order.

Both techniques are effective in teaching English vocabulary words,²⁸ characteristics of animals,²⁹ and U.S. history.³⁰ Scruggs and his colleagues note that mnemonic instruction “is target[ed] specifically to verbal associative learning and does not address all aspects of content area learning.”³¹

Putting It All Together: The School Improvement Plan (SIP)

As is the case for schools' needs assessments, SIPs are, in themselves, nothing new. Under NCLB, the federal government required that schools officially designated as “in need of improvement” use SIPs. ESSA requires the use of SIPs as well, as outlined at the beginning of this report.³²

However, in the absence of the needs-assessment steps outlined previously, the SIP is only of modest use. Without an exacting, granular self-assessment, a check by an external visiting team, and the careful matching of strongly researched interventions to the specific identified needs of schools, SIPs will capture best intentions but will fail to improve student learning.

Given this, the evidence on SIPs is not surprising: the quality of the SIP and its relationship to ongoing school improvement correlate to student outcomes. There is little overall evidence about the effects of SIPs on student learning.³³ However, there is some evidence of a positive correlation between high-quality SIPs and student learning. A study of Clark County School District (CCSD) in Nevada, at the time the fifth largest school district in the country, finds a consistently significant relationship between the quality of the SIP and student learning in math and reading, even after controlling for many school characteristics. SIP quality in this study was determined by a rubric created by CCSD and Nevada's Leadership and Learning Center,³⁴ which graded SIPs on seventeen indicators³⁵ on a scale of 1 to 3 (the highest rating). The study finds that “how frequently the SIP was monitored” and “the timeliness of SIP plan goals” are the two dimensions of SIPs most strongly associated with increased student learning.³⁶

Another study finds that higher-quality SIPs positively correlate with student achievement, but once student and school factors are controlled for, there is no statistically significant relationship between higher-quality SIPs and student achievement.³⁷ Meanwhile, a separate study examines the relationship between SIP quality and plan implementation by studying a school

improvement initiative in the Los Angeles Unified School District (LAUSD) that relies on formal school planning. The researchers' findings show that the average SIP quality was low throughout the initiative. They did, however, find a positive correlation between the relative quality of particular SIPs and reported implementation outcomes in the first phase of LAUSD's initiative.³⁸

The effects of a high-quality SIP are compounded when the school undertakes an ongoing improvement process. Research findings show more powerful effects from ongoing school improvement processes, rather than from SIPs alone.³⁹ For example, a four-year evaluation of schools' experiences in implementing the Scaffolded Apprenticeship Model (SAM)⁴⁰ for inquiry-based school improvement in New York City finds that schools that heavily invested in SAM developed a culture of inquiry and that student achievement improved. For example, schools with continued teacher involvement in the SAM credentialing program had a significantly larger proportion of students on track for high school graduation and college readiness, and a smaller proportion of students off track for high school graduation by their junior year, than other schools with similar student populations.⁴¹

Site Visits

In many instances, the issuance of the SIP is the last formal requirement of the school-reform process until interim and/or summative performance data must be produced. However, there are two critical steps in the process at which an external, on-site review of the workings of the school could add important quality assurances to what otherwise is an entirely local effort: (1) when the school is finalizing its needs assessment, and (2) after a reasonable number of months have passed and the implementation of the selected intervention(s) is well underway.

The site visit is an idea borrowed in part from the inspection process conducted by the Office for Standards in Education, Children's Services and Skills (OFSTED) in the United Kingdom.⁴² In the British case, the purpose of sending a carefully prepared team into a school is partly for accountability and partly for diagnostic purposes. New York adopted a somewhat similar model but strictly for the diagnostic purpose. The key value-add in holding such a visit prior to completing the needs assessment is to enable an outside analysis of the degree to which the school accurately has self-diagnosed its unique challenges. The

purpose of the second visit is to evaluate the degree to which the selected interventions are being implemented effectively, with the requisite support and leadership from the school. In both cases, the aim is to maximize the potential for the entire school improvement effort to show positive impact on student learning.

NYSED states that "[e]ach year, the Commissioner will appoint an Integrated Intervention Team (IIT) to conduct an on-site diagnostic district review and school reviews of selected Priority and/or Focus Schools within the district to inform the development of the District Comprehensive Improvement Plan and School Comprehensive Education Plan"⁴³ NYSED also provides the following guidance:

The review will be carried out over a period of one, two, or three days by an Integrated Intervention Team (IIT) composed of NYSED staff; an Outside Educational Expert selected by the district and approved by NYSED; a district representative; and in some instances experts in the education of English language learners and/or students with disabilities. The length of the review will depend on the size of the school and its accountability status. The process of conducting the reviews will focus on collecting and assessing low-inference data, (data of what is actually observed and heard, absent of added meaning, assumptions, conclusions and beliefs) to evaluate school and district practices based on six tenets. The six tenets are:

*Tenet 1: District Leadership and Capacity
Tenet 2: School Leader Practices and Decisions
Tenet 3: Curriculum Development and Support
Tenet 4: Teacher Practices and Decisions
Tenet 5: Student Social and Emotional Developmental Health
Tenet 6: Family and Community Engagement⁴⁴*

NYSED provides a set of [high-level guiding questions](#) to orient its visiting teams prior to their school visits.

Conclusion

Too often, the approach to improving underperforming schools resembles a doctor trying to treat a patient's condition without an accurate thermometer and with little knowledge of medical research. While school districts know which schools are struggling to bring students to grade level, they often lack the fine-grained knowledge that would give them a truly accurate "reading" of the school among the many factors that can impact student performance. But when armed with accurate granular data instruments supported by on-site analysis from experienced teachers and school administrators, school and district leaders can identify and prioritize the most critical causes of students' poor performance. Once equipped with an accurate diagnosis, leaders can vet potential interventions. In each case they can consider whether the intervention has been compared to alternatives based on the best available research in terms of its relevance to the school's population and context, intervention impact, implementation challenges, and costs.

Finally, because top-down interventions so easily can alienate those adults (namely teachers and school leaders) whom the interventions are intended to support, district leaders have the opportunity, with the support of the state, to embed the process of diagnosis and intervention selection in a comprehensive structure involving school personnel and parents, among others. The time has come for transparency, research, and commitments from schools, districts, and states to work together to bring the most effective assistance possible to those who are, every day, trying to make a difference for students who need it the most.



Photo by Allison Shelley/The Verbatim Agency for American Education: Images of Teachers and Students in Action

Appendix

Studies on English Language Learners, Students with Disabilities, and Students from Low-Income Families

It is possible for education leaders to carry the What Works Clearinghouse (WWC) and Best Evidence Encyclopedia (BEE) further by investigating individual studies the clearinghouses reviewed. The research cited from these two sources is not, however, exhaustive of all that is available. In particular, the redesigned BEE tool (Evidence in ESSA) uses selection criteria that are narrower than the pre-existing BEE database from which the tool is drawn. The researched interventions included in the following sections boosted learning for English language learners (ELLs), students with disabilities, and low-income-status middle and high school students. While this list prioritizes those studies that find a statistically significant impact and large effect size for a given intervention, it also includes examples of high-quality studies that do not reach statistical significance.⁴⁵ Statistical significance is tied to such things as sample size, and there are well-designed research studies with modest sample sizes that are nevertheless worthy of attention.

English Language Learners

TABLE 1: Pathway Project					
Study	Intervention ⁴⁶	Research Design	Context	Students Studied	Effect Size ⁴⁷
Olson et al. (2017) ⁴⁸	Pathway Project (professional development) Whole school (2 years)	Cluster randomized	16 secondary schools in Anaheim, CA 68% Latino, 18% Asian, 12% white, 21% ELL, 71% free or reduced-price lunch	575 10th-grade students (313 experimental group, 262 control group)	+.19

Pathway Project is a professional development program for teachers whose students include ELLs who can participate in standard English classes. Pathway focuses primarily on writing but includes reading comprehension strategies.⁴⁹ In this study, teachers received forty-six hours of training each school year (via six, six-hour released days scattered throughout the school year and five, two-hour after-school sessions) over a two-year period.

Researcher Carol Booth Olson, director of the UCI Writing Project at the University of California, Irvine, and her colleagues employed a multi-site, cluster, randomized field trial design. They selected sixteen secondary schools and randomly assigned the teachers to either the Pathway Project or the control group. The research team randomly selected one of each teacher's classes to participate in the study, and randomly assigned students to that class for each teacher. Using the California High School Exit Examination English language arts (ELA) scores as a measure, Pathway students scored non-significantly higher at the cluster level (+.19), controlling for California Standards Test pretests. This effect size suggests that the average student, starting at the 50th percentile, would move to roughly the 58th percentile after his or her teacher received Pathway professional development. An effect size of .19 also translates into roughly 143 days of additional ELA learning among tenth graders. This means that these students gained an additional 143 days of learning beyond what they would have learned had their teachers not received Pathway professional development. These findings largely replicated another randomized control trial of the Pathway Project that took place in California's Santa Ana Unified School District.⁵⁰

TABLE 2: Academic Language Instruction for All Students (ALIAS)

Study	Intervention	Research Design	Context	Students Studied	Effect Size ⁴⁷
Lesaux et al. (2010) ⁵¹	Academic Language Instruction for All Students (ALIAS) Whole class/school (18 weeks)	Cluster quasi-experimental	7 middle schools in an urban southwestern district 49% Latino, 73% ELL	476 6th-grade students (296 experimental group, 180 control group) 21 classes (13 experimental group, 8 control group)	+.15 [†]

ALIAS is a vocabulary intervention designed to be used forty-five minutes a day in standard English language arts classrooms with significant numbers of ELLs. Each series of lessons is based on one informational text. From that text, students work on small numbers of high-utility and abstract words. The intervention includes whole-group, small-group, and independent activities. Students have opportunities for listening, speaking, reading, and writing with the targeted words.

Lead researcher Nonie Lesaux of the Harvard University Graduate School of Education and her colleagues pre- and post-tested students using the Gates MacGinitie Reading Test (GMRT) Reading Comprehension exam. The effect of treatment on GMRT scoring was significant (effect size=.15/ $p < .06$) and suggests that the students gained an additional eighty-four days of learning by using ALIAS.

Students with Disabilities

TABLE 3: READ 180

Study	Intervention	Research Design	Context	Students Studied	Effect Size ⁴⁷
Sprague et. al (2012) ⁵²	READ 180 Targeted (1 year)	Student randomization	9th-grade students from 5 Title I high schools in western Massachusetts who tested between a 4th- and 6th-grade reading level 73% students of color, 19% students with disabilities, 72% free or reduced-price lunch	456 9th-grade students in 5 cohorts (231 experimental group, 225 control group)	+.18*

READ 180 is a supplemental program for struggling readers in which instructors conduct ninety-minute sessions each day with students. The program combines thirty-minute whole-group instruction with one hour of rotating twenty-minute activities that focus on independent reading, small-group direct instruction with the teacher, and use of READ 180's adaptive software. Teachers receive content-focused workshops and video instruction to develop pedagogical knowledge and skills.

Researcher Kimberley Sprague and her colleagues from the Education Alliance at Brown University compared READ 180 targeted interventions to control groups to determine effectiveness. The research team randomized cohorts of struggling-readers across five high schools serving predominantly students of color from low-income families. One-fifth of students in the study sample received special education services. Statistically significant, positive effects (+.18/ $p < .03$), or roughly 135 days of additional learning, were found for ninth-grade students in the READ 180 treatment cohort as compared to the control group.

TABLE 4: Butterfly Phonics

Study	Intervention	Research Design	Context	Students Studied	Effect Size ⁴⁷
Merrell and Kasim (2015) ⁵³	Butterfly Phonics Targeted, small group tutoring (4 months)	Student randomization	7th-grade students from 6 secondary schools in London, England, reading at least one year below grade level 78% white, 16% African American, 35% students with disabilities, 64% ELL, 51% free or reduced-price lunch	310 7th-grade students in 6 cohorts (161 experimental group, 149 control group)	+.30***

Butterfly Phonics is a small-group phonics instruction program for struggling readers. The program uses formal phonics instruction, comprehensive understanding of text, and class discussion of textual meaning to improve reading comprehension.⁵⁴ The program is delivered through a formal instruction style to small groups of six to eight students by a trained practitioner and an assistant.

Students who read at least one year below grade level were randomly assigned within schools. The control-group students continued to receive regular English lessons. The treatment-group students received two hours of Butterfly Phonics lessons per week and were withdrawn from English lessons, although they continued to receive lessons in classes that were taught in the English language. The New Group Reading Test (NGRT) measured effectiveness outcomes. The analysis by Christine Merrell and Adetayo Kasim, both from Durham University in the United Kingdom, found a significant effect size of +.30,⁵⁵ or an additional 235 days of learning, in favor of the treatment group on the NGRT test. Additionally, these findings were statistically significant at the $p < .001$ level.

TABLE 5: REACH

Study	Intervention	Research Design	Context	Students Studied	Effect Size ⁴⁷
Sibieta (2016) ⁵⁶	REACH (20 weeks)	Student randomization	Lowest readers in 27 disadvantaged secondary schools in or near Leeds, England 63% students with disabilities, 68% white, 32% students of color, 24% ELL, 31% free or reduced-price lunch	202 (70 REACH, 69 REACH-LC, 63 control group)	+.42*

REACH is a program from the United Kingdom that provides early-secondary school struggling readers with one-to-one tutoring. Students receive thirty-five-minute sessions once a week for twenty weeks with specially trained paraprofessionals. During these sessions, students read aloud from books at their level while tutors keep notes on student performance, which serves as the basis for remedial teaching. A variation of REACH, REACH-LC, adds a language comprehension element that emphasizes metacognitive skills, reading comprehension, inference, and writing.⁵⁷

Luke Sibieta, program director at the Institute for Fiscal Studies in London, evaluated REACH and REACH-LC in twenty-seven secondary schools in and around Leeds, England. The research team randomly assigned students who scored poorly on the Single Word Reading Test to REACH, REACH-LC, or control treatments. They were pre- and post-tested using NGRT. Post-test analysis found statistically

significant, positive effects with both the REACH cohort (effect size=+.33, p<.001) and the REACH-LC cohort (effect size=+.51, p<.001) over the control group. Averaging across the two variations, the mean effect size was +.42, which would translate into roughly 280 days of additional learning in the United States. Such translations should be regarded with caution, given that this intervention occurred in the United Kingdom.

Students from Low-Income Families

TABLE 6: Talent Development High School

Study	Intervention ⁴⁶	Research Design	Context	Students Studied	Effect Size ⁴⁷
Balfanz, Jordan, and Legters (2004) ⁵⁸	Talent Development High School (strategic reading and student team writing) Whole class/school (1 year)	Cluster quasi-experimental	6 nonselective high schools in Baltimore 89% African American, 9% white, >90% free or reduced-price lunch	457 9th-grade students (257 experimental group, 200 control group)	+.32

Talent Development High School (TDHS) is a whole-school reform model. Within the model, the strategic reading component aims to develop reading and writing comprehension skills. Teachers provide “read-aloud/think-aloud” demonstrations and mini-lessons on comprehension strategies. Time also is allocated for students to self-select reading and writing activities.

The research team of Robert Balfanz, Nettie Legters, and Will Jordan from Johns Hopkins University selected three well-matched schools. Eight teachers used strategic reading in twenty classes. The control schools provided double-dose classes in English and math but followed the district curriculum.⁵⁹ Post-test analysis, controlling for pretests and demographic factors, showed an effect size of +.32, which is equivalent to roughly 240 days of learning.

TABLE 7: One-to-One Laptop Use

Study	Intervention ⁴⁶	Research Design	Context	Students Studied	Effect Size ⁴⁷
Dunleavy and Heinecke (2007) ⁶⁰	1:1 Laptop use Whole school (2 years)	Cluster randomized	Middle school students (grades 6–8) in an urban school district in a mid-Atlantic state ⁶¹ 81% African American, 2% Asian, 3% Latino, 13% white, 59.6% free or reduced-price lunch	163 students in 14 classes in grades 6–8 (52 experimental group, 111 control group)	+.24*

Matt Dunleavy of Radford University and Walter Heinecke of the University of Virginia conducted a randomized trial to test the impact on student test scores of giving students individual laptops as opposed to their using school computers. The study randomly assigned 163 middle school students to one of two conditions: (1) treatment students received a laptop (Apple iBook/128MB hard drive/OS 10.2.8) loaded with online mathematics and science textbook access and laptop-based instruction, and (2) the control group had access to the same resources in the school’s computer lab. Researchers used the state standardized test (SST) to measure outcomes. Analysis revealed statistically significant positive effects were found on the SST scores favoring the treatment group (ES=.24/p<.03), or roughly an additional 160 days of learning.

Endnotes

- ¹ National Conference of State Legislatures, "Summary of the Every Student Succeeds Act, Legislation Reauthorizing the Elementary and Secondary Education Act" (Washington, DC: National Conference of State Legislatures, n/d), http://www.ncsl.org/documents/educ/ESSA_summary_NCSL.pdf (accessed March 30, 2017).
- ² For an analysis of these research standards and what they might mean in practice, see Chiefs for Change, "ESSA and Evidence: Why It Matters" (Washington, DC: Author, June 2016), <http://chiefsforchange.org/wp-content/uploads/2016/07/ESSA-and-Evidence-Why-It-Matters.pdf>.
- ³ ESSA states the following about the "targeted support and improvement plan": "(B) TARGETED SUPPORT AND IMPROVEMENT PLAN.— Each school receiving a notification described in this paragraph, in partnership with stakeholders (including principals and other school leaders, teachers and parents), shall develop and implement a school-level targeted support and improvement plan to improve student outcomes based on the indicators in the statewide accountability system established under subsection (c) (4), for each subgroup of students that was the subject of notification that—(i) is informed by all indicators described in subsection (c) (4)(B), including student performance against long-term goals; (ii) includes evidence-based interventions; (iii) is approved by the local educational agency prior to implementation of such plan; (iv) is monitored, upon submission and implementation, by the local educational agency; and (v) results in additional action following unsuccessful implementation of such plan after a number of years determined by the local educational agency. (C) ADDITIONAL TARGETED SUPPORT.—A plan described in subparagraph (B) that is developed and implemented in any school receiving a notification under this paragraph from the local educational agency in which any subgroup of students, on its own, would lead to identification under subsection (c)(4)(D)(i)(I) using the State's methodology under subsection (c)(4)(D) shall also identify resource inequities (which may include a review of local educational agency and school level budgeting), to be addressed through implementation of such plan." U.S. Department of Education, "Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act—Accountability and State Plans" (2016), <https://www.federalregister.gov/documents/2016/11/29/2016-27985/elementary-and-secondary-education-act-of-1965-as-amended-by-the-every-student-succeeds>. For further details and discussion, see J. Poiner, "School Identification and Intervention in the Age of ESSA," *Flypaper* (blog), Thomas B. Fordham Institute, June 21, 2016, <https://edexcellence.net/articles/school-identification-and-intervention-in-the-age-of-essa>.
- ⁴ ESSA defines "evidence-based" activity, strategy, or intervention as either (1) demonstrating a statistically significant impact on student achievement based on evidence from at least one well-implemented (i) experimental study (ii) quasi-experimental study, or (iii) correlational study that controls for selection bias; or (2) demonstrating a rationale based on high-quality research that such an activity will likely improve student achievement and includes an ongoing evaluation of the effect of the activity, intervention, or strategy. This report refers to research that meets the first standard using experimental study as Tier 1; research that meets the standard using quasi-experimental research as Tier 2; and research that meets the standard using a correlational study as Tier 3. Research that meets the second standard is referred to as Tier 4.
- ⁵ See U.S. Department of Education, "Supporting School Reform by Leveraging Federal Funds in a Schoolwide Program: Non-Regulatory Guidance" (Washington, DC: Author, September 2016), <https://ed.gov/policy/elsec/leg/essa/essaswpguidance9192016.pdf>.
- ⁶ AdvancEd, "ASSIST™ school improvement tool," <http://www.advanced.org/services/assist> (accessed March 30, 2017).
- ⁷ Indistar, Homepage, <http://indistar.org>. See tabs at <http://indistar.org/howto/> for further details about the tool. (accessed March 30, 2017).
- ⁸ K. Bond and L. Sodat, "Conducting a Comprehensive Needs Assessment for the Title I, Part A, Schoolwide Plan," (Richmond, VA: Virginia Department of Education, Office of Program Administration and Accountability Coordinators' Academy, July 2014), http://www.doe.virginia.gov/federal_programs/esea/tech_assistance_academy/2014/Conducting_a_Comprehensive_Needs_Assessment.pdf.
- ⁹ M. Smith and L. Spain, "Conducting a Comprehensive Needs Assessment In Your School," (Arkansas Department of Education, Department of Learning Science, July 2014) <http://centeronschoolturnaround.org/wp-content/uploads/2014/10/AR-Conducting-a-Needs-Assessment.pdf>.
- ¹⁰ WestEd, "A Guide for Comprehensive Needs Assessment" (Washington, DC: Author, April 4, 2008), http://www.cde.state.co.us/sites/default/files/documents/fedprograms/dl/consapp_na_guide.pdf.
- ¹¹ See also K. E. Fernandez, "Evaluating School Improvement Plans and Their Effect on Academic Performance," *Educational Policy* 25, no. 2 (2011): 338–367, 347, for a list of seventeen research-based indicators of high-quality SIPs for additional components to include.
- ¹² Such language appears across many of the school needs assessments. The second phrase comes from Arkansas's school needs assessment.
- ¹³ Elements of this section are drawn from material prepared for an internal memo to Chiefs for Change. The Institute is grateful to Chiefs for Change for permission to use this material.
- ¹⁴ Ohio Department of Education, Horace Mann Pre-K–8 School Report Card 2015–16, <http://reportcard.education.ohio.gov/Pages/School-Report.aspx?SchoolID=016840><http://reportcard.education.ohio.gov/Pages/School-Report.aspx?SchoolID=016840> (accessed March 31, 2017).
- ¹⁵ State of New Jersey Department of Education, Alexandria Township Middle School: School Report Card 2014–15, <http://www.state.nj.us/education/pr/1415/19190020005.pdf> (accessed March 31, 2017).
- ¹⁶ Colorado Department of Education, "SchoolView: Data & Accountability," <http://www.cde.state.co.us/schoolview> (accessed March 31, 2017).
- ¹⁷ The Louisiana data was provided by the Louisiana Department of Education in a private correspondence.
- ¹⁸ New York State Education Department (NYSED), "Office of Accountability: Diagnostic Tool for Effectiveness," <http://www.p12.nysed.gov/accountability/diagnostic-tool-institute/home.html> (accessed March 31, 2017); NYSED, "Diagnostic Tool for Schools and District Effectiveness FAQs," June 2015, <http://www.p12.nysed.gov/accountability/diagnostic-tool-institute/DTSDEFAQs>.

- ¹⁹ J. W. Altschuld and B. R. Witkin, *From Needs Assessment to Action: Transforming Needs into Solution Strategies* (Thousand Oaks, CA: Sage Publications, 2000).
- ²⁰ U.S. Department of Education, Institute of Education Sciences, "What Works Clearinghouse: Welcome Page," <https://ies.ed.gov/ncee/wwc> (accessed March 31, 2017).
- ²¹ WWC combines the quality of research, sheer number of research studies that made it over the bar for inclusion, and actual impact on learning that the research suggests occurs due to the intervention. Only by digging into the supporting materials can the user disentangle these elements.
- ²² C. Goldenberg, "Teaching English Language Learners: What the Research Does—and Does Not—Say," *American Educator*, Summer 2008.
- ²³ M. Fine, B. Stoudt, and V. Futch, *The Internationals Network for Public Schools: A Quantitative and Qualitative Cohort Analysis of Graduation and Dropout Rates* (New York, NY: The Graduate Center, City University of New York, June 2005), <http://internationalsnps.org/wp-content/uploads/2012/10/Fine-Report-for-Web-and-Print.pdf>.
- ²⁴ T. E. Scruggs et al., "Do Special Education Interventions Improve Learning of Secondary Content? A Meta-Analysis," *Remedial and Special Education* 31, no. 6 (November 1, 2010): 437–49, <http://journals.sagepub.com/doi/abs/10.1177/0741932508327465>.
- ²⁵ The mean effect for explicit instruction was 1.68 and the mean effect for mnemonic instruction was 1.47. The mean effect size indicates the magnitude of the effect an intervention had on the group that received the treatment versus a comparison group. Researchers do not usually translate effect sizes involving special education students into additional days of learning.
- ²⁶ J. A. McCleery and G. A. Tindal, "Teaching the Scientific Method to At-Risk Students and Students with Learning Disabilities Through Concept Anchoring and Explicit Instruction," *Remedial and Special Education* 20, no. 1 (January 1, 1999): 7–18, <http://journals.sagepub.com/doi/abs/10.1177/074193259902000102>.
- ²⁷ P. Hudson, "Using Teacher-Guided Practice to Help Students with Learning Disabilities Acquire and Retain Social Studies Content," *Learning Disability Quarterly* 20, no. 1 (1997): 23–32, <http://journals.sagepub.com/doi/abs/10.2307/1511090>.
- ²⁸ M. A. Mastropieri, T. E. Scruggs, and C. Whedon, "Using Mnemonic Strategies to Teach Information about U.S. Presidents: A Classroom-Based Investigation," *Learning Disability Quarterly* 20, no. 1 (1997): 13–21, <http://journals.sagepub.com/doi/abs/10.2307/1511089>.
- ²⁹ T. E. Scruggs et al., "Effects of Mnemonic Reconstructions on the Spatial Learning of Adolescents with Learning Disabilities," *Learning Disability Quarterly* 15, no. 3 (August 1992): 154–62, <http://journals.sagepub.com/doi/abs/10.2307/1510240>.
- ³⁰ T. E. Scruggs and M. A. Mastropieri, "Mnemonic Instruction of LD Students: A Field-Based Evaluation," *Learning Disability Quarterly* 12, no. 2 (1989): 119–25, <http://journals.sagepub.com/doi/abs/10.2307/1510727>.
- ³¹ Scruggs et al., "Do Special Education Interventions Improve Learning of Secondary Content?"
- ³² SIPs are required in ESSA. See endnote 3 for details.
- ³³ R. S. Ettinger, "Shifting from Plan to a Process: School Improvement Plans in Cambridge Public Schools" (doctoral dissertation) Harvard University, 2015.
- ³⁴ The rubric was designed to be used across the United States and Canada and was created by examining SIP templates in use in the United States and Canada and by reviewing the literature on best practices of SIPs.
- ³⁵ Example indicators include (1) comprehensive, which measures if goals are connected to concerns and causes identified in the SIP; (2) measurable goal, which measures if goals are quantifiable and include a baseline and target measure; and (3) monitoring plan, which measures if there are monitoring steps in places that explicitly describe what will be done to assess progress toward each goal.
- ³⁶ K. E. Fernandez, "Evaluating School Improvement Plans and Their Effect on Academic Performance."
- ³⁷ D. J. Huber and J. M. Conway, "The Effect of School Improvement Planning on Student Achievement," *Planning and Changing* 46, no. 1/2 (2015): 56.
- ³⁸ K. O. Strunk et al., "The Best Laid Plans: An Examination of School Plan Quality and Implementation in a School Improvement Initiative," *Educational Administration Quarterly* 52, no. 2 (2016): 259–309.
- ³⁹ V. Bernhardt, *Data Analysis for Continuous School Improvement* (New York, NY: Routledge, 2013).; G.J. Langley, *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance*, 2nd ed. (San Francisco: Jossey-Bass, 2009).
- ⁴⁰ SAM is a well-defined model of school change that emphasizes collaborative inquiry and teacher leadership in which a team of teachers uses student data to identify specific learning gaps and develop targeted interventions.
- ⁴¹ J. E. Talbert et al., *Inquiry-Based School Reform: Lessons from SAM in NYC* (Center for Research on the Context of Teaching: Stanford University, 2012).
- ⁴² "OFSTED Home Page," Government of the United Kingdom, <https://www.gov.uk/government/organisations/ofsted> (accessed March 31, 2017).
- ⁴³ NYSED, "Diagnostic Tool for School and District Effectiveness (DTSDE)" (Albany, NY: Author, October 2015), <http://p1232.nysed.gov/accountability/diagnostic-tool-institute/documents/DTSDEOrientationforSchoolBasedTeamsandLeaders-SeptemberandOctober2015.pdf>.
- ⁴⁴ ———, "Office of Accountability: Diagnostic Tool for Effectiveness."
- ⁴⁵ X. Fan, "Statistical Significance and Effect Size in Education Research: Two Sides of a Coin," *Journal of Educational Research* 94, no. 5 (2011): 275–82, http://www.jstor.org/stable/27542335?seq=1#fn-dn-references_tab_contents.
- ⁴⁶ Duration of the intervention appears in parenthesis.
- ⁴⁷ Statistical significance: †p < .10, *p < .05, **p < .01, ***p < .001.

- ⁴⁸ C. B. Olson et al., "Reducing Achievement Gaps in Academic Writing for Latinos and English Learners in Grades 7–12," *Journal of Educational Psychology* 109, no. 1 (2017): 1–21, <http://dx.doi.org/10.1037/edu0000095>.
- ⁴⁹ A. Baye et al., "Effective Reading Programs for Secondary Students" (Baltimore, MD: Johns Hopkins University School of Education's Center for Data-Driven Reform in Education, December 2016), <http://www.bestevidence.org/word/Secondary-Reading-12-06-16.pdf>.
- ⁵⁰ C. B. Olson and R. Land, "A Cognitive Strategies Approach to Reading and Writing Instruction for English Language Learners in Secondary School," *Research in the Teaching of English* 41, no. 3 (2007): 269–303, <http://www.jstor.org/stable/40171732>.
- ⁵¹ N. K. Lesaux et al., "The Effectiveness and Ease of Implementation of an Academic Vocabulary Intervention for Linguistically Diverse Students in Urban Middle Schools," *Reading Research Quarterly* 45, no. 2 (2010): 196–228, <http://www.jstor.org/stable/20697183>.
- ⁵² K. Sprague et al., "Springfield-Chicopee School Districts Striving Readers (SR) Program Final Report Years 1–5: Evaluation of Implementation and Impact" (Providence, RI: Research & Evaluation Division of the Education Alliance at Brown University, March 2012), <https://www2.ed.gov/programs/strivingreaders/springfieldyear5evalrpt.pdf>.
- ⁵³ C. Merrell and A. Kasim, "Butterfly Phonics: Evaluation Report and Executive Summary" (London, UK: Education Endowment Fund, February 2015).
- ⁵⁴ This study was included in the original research view using resources in BEE. This material is not included in the ESSA-related tool as this is international research.
- ⁵⁵ Effect size is a pooled effect size. The effect size for all students was .43. Separate analysis was conducted for students who qualify for free or reduced-price lunch. Effect size for this group was .16.
- ⁵⁶ L. Sibietta, "REACH Evaluation Report and Executive Summary" (London, UK: Education Endowment Foundation, February 2016), https://educationendowmentfoundation.org.uk/public/files/Support/Campaigns/Evaluation_Reports/EEF_Project_Report_REACH.
- ⁵⁷ Baye et al., "Effective Reading Programs for Secondary Students."
- ⁵⁸ R. Balfanz, N. Legters, and W. Jordan, *Catching Up: Impact of the Talent Development Ninth Grade Instructional Interventions in Reading and Mathematics in High-Poverty High Schools* (Baltimore, MD: Johns Hopkins University Center for Research on the Education of Students Placed at Risk, December 2004), <https://eric.ed.gov/?id=ED484524>.
- ⁵⁹ Baye et al., "Effective Reading Programs for Secondary Students."
- ⁶⁰ M. Dunleavy and W. F. Heinecke, "The Impact of 1:1 Laptop Use on Middle School Math and Science Standardized Test Scores," *Computers in the Schools* 24, no. 3–4, 7–22. (2007): 7–22, http://www.tandfonline.com/doi/abs/10.1300/J025v24n03_02.
- ⁶¹ Researchers in this study chose not to identify the state and district.



ALLIANCE FOR
EXCELLENT EDUCATION

all4ed.org