CHIEFS for CHANGE

2016 July 7

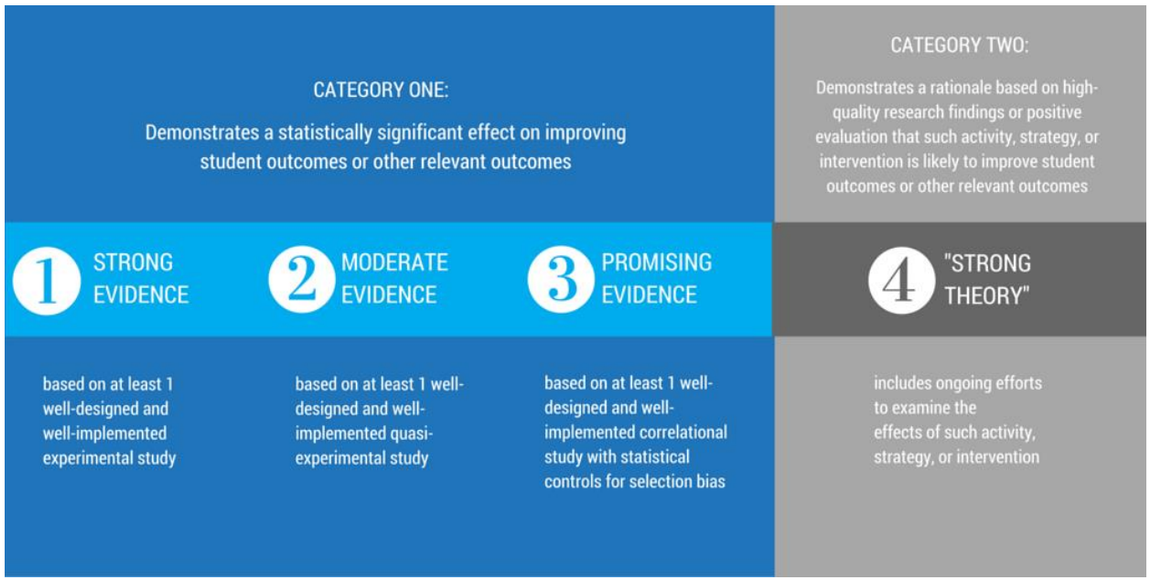
ESSA & EVIDENCE

Why It Matters

POLICY BRIEF
ESSA AND EVIDENCE: WHY IT MATTERS[1]

KEY TAKEAWAY
ESSA incentivizes states to use evidence-based programs and interventions in districts and schools. Doing so will lead to stronger student outcomes at reasonable cost; not doing so throws dollars after uncertain or even negative outcomes for students and schools.

OVERVIEW
ESSA's evidence-based provisions offer states an important opportunity to work with their school districts to select and implement research-based interventions. ESSA tiers research across four standards that embody varying degrees of methodological rigor, with Tier 1 representing the strongest, and Tier 4 the weakest, levels of research. ESSA lays out a number of funding streams that are only available if used to support activities that have research support. In some cases, states are confined to selecting interventions supported by Tiers 1-3; in other cases, they are free to select activities that have only Tier 4 research support.



The purpose of this policy brief is to suggest how state Chiefs might navigate ESSA's evidence-based provisions effectively. We will clarify what ESSA requires in terms of the use of evidence-based policies, what each standard of evidence means, and how state Chiefs might create a climate in which research-based policies become the default choice for district interventions.

WHY DOES IT MATTER?
Taken at face value, the evidence-based provisions of ESSA are not especially demanding. It is not very difficult to identify interventions that satisfy Tier 3 or Tier 4 levels of research. Although the proposed regulatory language from USED suggests that the Department will try to nudge states toward interventions with the strongest available research base, it is most unlikely, in our judgment, that a future USED would micromanage the process. Given thousands of potential interventions, and the fact that meeting the lower tiers of evidentiary support is relatively unproblematic, state Chiefs might understandably regard this entire aspect of ESSA as low on their priority lists.

In our judgment, doing so would be a serious mistake on educational, fiscal, moral, and political grounds.

Guiding districts and schools *toward* programs that have clear positive effects and *away* from trends without track records has great potential to change the academic trajectory for PK-12 students across the country. One example: replacing a math curriculum which lacks all research support with one deemed strong by Tier 1 research could add as much as seven months' worth of student learning – and for a very modest outlay (G. Whitehurst, 2009).

Given this country's static, or even widening, achievement gaps (Reardon, 2011) (Barton, Coley, & ETS Policy Information Center, 2010), this is also a matter of equity: Where research indicates the potential of a given intervention to significantly reduce achievement gaps, we surely have a pedagogical and moral imperative to employ it if we are able to do so.

---

[1] Chiefs for Change is grateful to David Steiner and his colleagues at the Institute for Education Policy at the Johns Hopkins University School of Education for providing the research and analysis in this brief.

From a fiscal point of view, committing public dollars to a policy that lacks research support, or to one with a scantier evidentiary base than a viable alternative, makes no sense.

At a time of constrained education spending, America's school districts waste billions of dollars each year with no clear correlation between expenditures and student performance – indeed, some wealthy districts are singularly unproductive (Boser, 2014).

An example may clarify why this is so important. Focusing once again on the domain of curricula, we now have high-caliber research indicating that some instructional materials produce better learning outcomes than others, and that the positive impacts of the best curricula are clear and sizable (Bhatt & Koedel, 2012), (Bhatt, Koedel, & Lehmann, 2013), (Agodini & Pendleton, 2009). Further, high-quality curricula do a better job of boosting student achievement than other popular interventions such as expanding preschool programs, giving merit pay to teachers, decreasing class sizes, or increasing the number of charter schools in a district – and at a much lower cost (G. Whitehurst, 2009), (Hattie, 2015), (G. J. Whitehurst, Chingos, Gallaher, & Brookings Institution, 2013). In addition, schools *have to* choose a curriculum, and the cost difference between curricula is small, thus making a change to stronger curricula relatively inexpensive (Polikoff, 2014). Finally, switching to higher-quality curricula is less contingent upon context than interventions such as universal preschool or whole school reforms. Using high-quality instructional materials, then, satisfies the criteria above: The research base is strong, the program effects positive, the implementation relatively easy, and the cost minimal. However, according to a recent national study, very few states evaluate the strength of curricula effectively so as to make cost-effective, academically impactful, decisions (Boser, Chingos, & Straus, 2015).

What happens when leaders fail to take evidence into account? Students learn less, and states waste money. Examples abound. According to a 2015 report from TNTP, school districts spend an average of $18,000 per teacher, per year on professional development, with no discernible impact on teacher effectiveness (TNTP, 2015). Reducing class size is expensive: California spent $25 billion on its program, only to have researchers conclude that the "attribution of gains in scores to CSR is not warranted" (California Department of Education, 2002). Indeed, a counter-argument is to *expand* the class sizes of highly effective teachers (Hansen, 2013).

There is another benefit to the use of evidence-based interventions: It can be politically advantageous. Programs that rest upon strong evidence are publically defensible and may be more readily transferred between districts. Here, the advantages depend upon sensitive implementation. Pressing a particular intervention on a district in a heavy-handed way, simply on the grounds that it has a higher research "score," is unlikely to be effective. But there are other strategies – of coalition-building, partnerships, and the thoughtful use of websites and other informational channels – that can, as suggested below, leverage the language of evidence-based policy making into an effective tool for district-level reform. Being transparent about data can build stakeholder support and thus enable state and district leaders to sustain a given intervention.

Thus, there is much to gain from understanding, comparing, and acting upon initiatives that are supported by strong research. Given the urgent needs of our most disadvantaged students, doing otherwise is irresponsible.

WHAT ESSA SAYS ABOUT EVIDENCE AND WHAT IT MEANS

ESSA endorses the use of evidence in federal, state, and district education policy, mentioning the term "evidence-based" 61 times and categorizing evidence-based upon the strength of the research base. The organization Results for America estimates that ESSA could shift as much as $2 billion per year, for the next four years, to evidence-based education interventions ("Michele Jolin Praises Senate Leaders on Evidence-Based Policy Provisions in ESSA," 2015).

The federal emphasis on evidence is not new. The previous iteration of the Elementary and Secondary Education Act (ESEA) did so, too, as Martin West notes: "NCLB also sought to make the American education system more data-driven, famously using the term 'scientifically based research' some 110 times in an attempt to limit the use of federal funds to activities with proven results." What has changed? The way evidence is defined and understood. "[NCLB] defined scientifically based research narrowly, emphasizing the need for experimental or quasi-experimental studies (and expressing a clear preference for the former)" (West, 2016). In contrast, ESSA lays out four standards of evidence rather than two, and makes the use of one the three a requirement for certain funded activities. The definition of "evidence-based" is given in section 8101(21) of the ESEA, as amended by the ESSA (United States Congress, 2015).[2] ESSA delineates "evidence-based" actions according to four categories that reflect strength of evidence.

[2] Pp. 393-394.

EVIDENCE-BASED. —

(A) IN GENERAL. —Except as provided in subparagraph (B), the term 'evidence-based', when used with respect to a State, local educational agency, or school activity, means an activity, strategy, or intervention that—

(i) demonstrates a statistically significant effect on improving student outcomes or other relevant outcomes based on—

(I) strong evidence from at least 1 well-designed and well- implemented experimental study;
(II) moderate evidence from at least 1 well-designed and well-implemented quasi-experimental study; or

(III) promising evidence from at least 1 well-designed and well-implemented correlational study with statistical controls for selection bias; or

(ii)(I) demonstrates a rationale based on high-quality research findings or positive evaluation that such activity, strategy, or intervention is likely to improve student outcomes or other relevant outcomes; and

(II) includes ongoing efforts to examine the effects of such

activity, strategy, or intervention.

(B) DEFINITION FOR SPECIFIC ACTIVITIES FUNDED UNDER THIS ACT. —When used with respect to interventions or improvement activities or strategies funded under section 1003, the term 'evidence-based' means a State, local educational agency, or school activity, strategy, or intervention that meets the requirements of subclause (I), (II), or (III) of subparagraph (A)(i).

The most important funding stream impacted by these provisions is the 7% of Title I, Part A, funds that states must set aside to help school districts improve their lowest-performing schools. These funds can be used only for interventions that are supported by the top three tiers of research as defined above. ESSA also includes a number of other formula grants (such as recruiting high-quality teachers) that can be used only on activities that meet this same standard of research. For a full list of these funds, see the very useful summary created by Results for America here.

FOUR TIERS OF EVIDENCE

The basic goal of education research is to isolate the effects of an intervention by rendering all background factors equal. A finding that students in small classes have higher achievement than those in large classrooms, if the first group of students happens to come from high-income families or if all their teachers happen to be ranked as "highly effective," would be of little interest. The finding only "matters" if the two student groups are identical in significant ways. ESSA's four standards are about this very thing: How effectively does the research behind a particular intervention isolate that intervention from conditions that would invalidate its findings?

The first tier is considered "strong" because it employs a randomized control trial (RCT). RCTs assign students to an intervention or to a control group randomly. The only expected difference between the control and intervention groups in an RCT is the outcome of the variable being studied.

An example of an intervention that was found to be statistically significant under the RCT model is a three-year study of Success for All's literacy program. The research team, funded under an i3 grant, randomly assigned schools to implement the Success for All reading program. Students in the "treatment" schools demonstrated the equivalent of two to three months' worth of learning beyond those in the control schools, and the Black-White achievement gap narrowed substantially in the treatment schools (Borman et al., 2007).

Other important areas of education research do not randomize the treatment directly, but rather use the randomization that occurs naturally. For instance, a direct comparison of test results for students in district and charter schools is of little value, because the families of the two groups are different in at least one significant way: Parents of charter school students *choose* to enroll their children in a non-zoned school. A study that mimics an RCT would compare the test scores of students who applied via lottery and won a charter school seat, and those of students who applied via lottery and did not. Such a method, which is often used in this context, would isolate the school effect (Furgeson, Mathematica Policy Research, Daniel J. Evans School of Public Affairs, & Center on Reinventing Public Education, 2012), (Hoxby, Murarka, & National Bureau of Economic Research, 2009).

RCTs provide the highest level of certitude about program effects and are considered the "gold standard" of research. Despite this benefit, RCTs carry inherent liabilities: They are costly and time-consuming, and they raise ethical questions about assigning children to less-favorable educational conditions. There are also many interventions that are difficult to study under these conditions, so the pool of interventions with this level of evidence may be limited.

The second standard ("moderate" or quasi-experimental) does not involve randomization. Quasi-

experimental research still compares control and treatment groups, but the two groups are not randomly distributed. Rather, researchers match the groups as well as possible, such as by demographics, age, gender, and other factors that might otherwise explain different results. For background on the use and risks of this kind of research, see a 2014 paper here (Walser, 2014). For an example of this form of research (on class size), see here (Lavy, Angrist, & National Bureau of Economic Research, 1997). Stanford's Center for Research on Education Outcomes also employs matching techniques – in this case, by creating "virtual twins" – to compare the effects of charter schools on students' academic achievement (Stanford University & Center for Research on Education Outcomes, 2013).

The third standard for research is often known simply as "correlational." In such studies, researchers try to determine whether or not two variables are related but cannot establish whether one causes the other. This means studying whether an increase or decrease in one variable corresponds to an increase or decrease in the other variable. Thus, to give a familiar example, we might be able to show that there is a relationship between aggressive behavior and watching violent TV programs, but not whether one behavior leads to the other. For an example of correlational research (on the teaching of mathematics), see the study here (Haciomeroglu, 2013).

Another example of correlational research: Harvard's Strategic Data Project found that teachers in the Los Angeles Unified School District who had earned a certificate from the National Board for Professional Teaching Standards (NBPTS) were able to increase student learning by an additional one to two months compared with peers who had otherwise similar backgrounds but had not become board certified (Strategic Data Project, 2012). The study could not establish that the certification process *caused* teachers to be more effective; those teachers could have been more effective before certification. Other studies come closer to establishing causation, although still imperfectly (Cavalluzzo, Barrow, Henderson, CNA Education, & National Board for Professional Teaching Standards, 2014).

The final standard refers to promising programs that have not yet been scientifically researched. Rather, this standard requires states to reference a positive evaluation of some kind in relationship to a chosen intervention and to track its effects in the field. A relevant example would be competency-based education. Competency-based education is used extensively in higher education and medicine (Frank JR et al., 2010), (Voorhees, 2001) and has foundations in education theory (Carr, 2003). However, it has not been rigorously studied in a primary or secondary setting. Another timely example would be the use of micro-credentials or badges. Micro-credentialing presents an innovation in career readiness; students can signal to potential employers that they have acquired relevant skills. However, the effects of micro-badges have not been rigorously studied (Devedžic, Jovanovic, Devedžic, & Jovanovic, 2015).

While the focus of this paper is on the use of school improvement funds in Section 1003– which can only be spent on interventions that meet one of the top three evidentiary standards – ESSA names numerous funding streams that need to satisfy only Tier 4's requirement of a positive evaluation.

In practice, these rankings are important. Imagine that you are picking from two school-turnaround models. The first is supported by quasi-experimental research and has a strong 0.50 effect size (roughly equivalent to up to a full year of extra learning). The second has been the subject of a well-designed randomized control trial, and the effect size is a 0.25. All else being equal (we will come back to that point), one should pick the second intervention. This choice becomes even clearer if several randomized control trials point to the same positive effects. This is because you can be far more certain that those effects will in fact occur. While a stronger impact is, of course, preferable, the risk that the intervention supported by weaker evidence will achieve much less when you actually use it is much greater. The further you move down the rankings that validate a particular strategy, the more likely it is that the positive findings of that research will not be replicated in a new intervention. In short, from a purely statistical point of view, you should aim as high on that evidentiary scale as possible.

All else is of course not equal. Districts, with state approval, will need to make complex choices that balance not only the tier of the research available but also the *quality* of that research. Not all Tier 1 research is equally well done: if there is only one study, or studies that point in different directions, statistical expertise will be needed. Below, we will come to non-statistical issues such as cost and politics that will also enter into district and state policy decisions.

U.S. DEPARTMENT OF EDUCATION DRAFT REGULATIONS

To date, USED has given notice of Proposed Rulemaking that includes the major evidence-based provisions in ESSA. Some of the proposed language is available for public comment until August 1 and may be referenced here.

There are more than 100 references to "evidence-based" in the Proposed Rulemaking. The focus is on the standards that SEAs should employ in dispensing funds to districts for targeted or comprehensive school support. Two key clauses suggest how USED is hoping to use regulations to nudge states toward the most rigorously researched interventions where such research is available. In general, USED maintains that:

"The purpose of these proposed regulations is to increase the likelihood that funds are awarded to LEAs that will successfully implement interventions in schools identified for comprehensive or targeted support and improvement. Specifically, the use of more rigorous evidence-based interventions and strong support from the local community are likely to increase a school's chances of significantly improving student achievement and outcomes."

Then it goes into detail: "Because the evidence base for interventions in low-performing schools is relatively nascent and still growing, proposed §200.21 would help support LEAs in making prudent, smart choices when selecting among evidence-based interventions by encouraging the use of interventions that are supported by the strongest level of evidence that is available and appropriate to meet the needs of the school, including, where possible, evidence suggesting that the intervention was effective for an overlapping population or in an overlapping setting to those of the identified school." The Department is clearly trying to use the regulatory regime to steer states toward interventions with the highest available level of research support.

This is potentially of some importance. States that wish to choose a Tier 3 intervention will have little difficulty finding a correlational study, somewhere, to justify that decision. The Department has limited leeway: It is simply not realistic to expect USED to make thousands of judgments as to whether a state used the "strongest level of evidence available." ESSA does provide important support for state Chiefs who choose to take the evidence-based provisions seriously, but *de facto*, offers limited accountability if state Chiefs do not. As we have suggested above, state Chiefs should choose interventions that rely upon the best possible research – because it is the right thing to do for students and most likely to achieve the intended outcomes.

RECOMMENDATIONS

It is tempting to recommend a decision logic for state Chiefs as follows: Every time you dispense funds to districts under the constraints of ESSA's evidence-based provisions, ensure that the district uses an intervention that is supported by strong positive findings, validated by a randomized control trial (Tier 1), rather than one backed by quasi-experimental or correlational studies (Tiers 2 and 3).

Based on strong research, state Chiefs could, for instance, use ESSA's Title I provisions to press districts to use stronger instructional materials in schools determined to be chronically under-performing (§200.21) and in schools determined to have chronically under-performing subgroups (§200.22). State Chiefs could try to mandate that all programs for English Language Learners reflect research findings on the amount of time required to learn a second language (§200.13).[3] State Chiefs could leverage research (Papay, West, Fullerton, & Kane, 2012) on the value of teacher residency programs in their application for Title II funds (§ 299.18b) (United States Department of Education, 2016).

But this kind of direct action may not be realistic or advisable - for several reasons. A top-down approach to district-level actions with regard to their lowest-performing schools would run afoul of ESSA's injunction to engage stakeholders in policy actions and might elicit resentment from district leaders.

[3] The USDOE Regulations proposed on May 31, 2016, even cite specific research on this (p. 12).

Additional factors might also – rightly – impact the choice of a particular intervention. These include:

- ***Cost.*** Cost is not correlated with levels of evidence. In some cases, interventions supported by the strongest evidence are cheaper than those with little evidentiary support. In other cases, however, the "best" policy may indeed be deemed unaffordable.
- ***Impacted Population*****.** Interventions in low-performing schools present multiple options. State Chiefs may need to choose between a program that greatly benefits a modest sub-section of the school population, and one that yields a more modest benefit – but to a much larger group of students.
- ***Politics*****.** Not all policies commended by strong research are politically acceptable. For example, research from CREDO might suggest that the Maryland SEA direct Baltimore to re-constitute its lowest-performing district schools with charter schools managed by nationally-successful CMOs (CREDO, 2015). Such a policy, however, is politically unavailable to the state at this time.
- ***Local circumstances and capacity*****.** Some interventions may be cost-effective and politically viable but not operationally possible. What appears strong in abstract terms may make little sense on the ground in a particular district or school. Factors such as school culture, leadership limitations, or the lack of locally available expertise, may constrain a state Chief's options.

The key action goal for state Chiefs and their leadership teams is thus not necessarily to press districts into interventions with the highest level of evidentiary support. Rather, **state Chiefs should focus on building "a culture of the use of evidence-based education policies."** This is a long-term strategy that aims to change the decision-making routine at the district and school levels. By this way of thinking, ESSA's evidence-based provisions can become a powerful means to an important end.

Such an approach does not imply a "hands-off" stance to the dispensing of funds tied to ESSA's evidence-based provisions. Rather, it acknowledges that Chiefs have multiple opportunities to champion high-impact, high-evidence programs – as discussed below. As a general matter, however, and with exceptions in situations of deep urgency, Chiefs should not, in our view, be in the business of dictating district-level choices of school-level interventions. Instead, they can use this remarkable opportunity to provide leadership in changing the policy culture.

To do so, Chiefs will need both reliable sources of information about the research support for possible intervention strategies and also mechanisms for moving the dial on the use of research as a key component of policy decision-making. This will require specific structures and processes at the SEA level to ensure that consideration of evidence becomes embedded in the culture of the organization.

Perhaps the most important starting point in building this culture can be the leverage provided by the needs assessment that districts must provide for each school that is subject to Comprehensive Support and Improvement under ESSA. This documentation is intended to identify the specific challenges faced by the school that form the basis of the district-recommended intervention. But that intervention, if it is to be evidence-based, must rely on the best possible analysis of school-level data – one that specifies the issues, the targeted population(s) and the priority-order of the issues to be addressed. Unless this underlying analysis is done well, the subsequent choice of a research-based intervention may be of limited value – just as a rigorous medical treatment is useless if it fails to correlate to the patient's precise diagnosis.

We strongly encourage state Chiefs to work with Chiefs for Change and the research community to develop a needs-assessment template for use by schools subject to interventions – and the districts in which they are located. Such a template could help form a key component in all subsequent actions to bring better results to the schools.

It is no small task for SEAs to analyze these needs assessments (which should include assisting districts in conducting them) and judge the efficacy and research support for proposed district-level action, much less to draw upon the plethora of funding opportunities under ESSSA that require evidence-based support. Where and how should states find the necessary support and expertise?

Assuming that funding for them is supported when the Education Sciences Reform Act is reauthorized, the most obvious source of help for state Chiefs seeking information will be the Regional Educational Laboratory (REL) that serves their states. State Chiefs, with the assistance of USED and organizations such as Chiefs for Change and CCSSO, may have to re-think their

use of these Labs: To date, they have often been seen as peripheral to state policy choices. Under ESSA, they could play a much more central role in supplying SEAs with actionable research data responding to such questions as:

- Has this intervention been subject to serious research analysis?
- Does that research show positive program effects?
- How large are the effect sizes, and how cost-effective, compared to other interventions?
- How easy is the intervention to implement with sufficient fidelity to remain effective?

State Chiefs who are serious about the use of evidence who need to rely heavily on RELs may need to designate one or more senior staff members with the primary responsibility to serve as the liaison with the REL. The challenge here is that the 10 RELs, which serve multiple states, were not set up to serve all the needs states may potentially have under ESSA – such as assessing the multiple interventions districts may want to employ for their compliance with ESSA's evidence requirements.

For states with greater in-house capacity, state Chiefs might consider creating something like an Evidence Resource Center. Such a center would work on all aspects of the evidence-based agenda. Among its tasks would be to curate and distribute education research of immediate relevance to school districts. The center would draw from such sources as "What Works Clearinghouse," USED's guide to research-based educational interventions. What Works Clearinghouse is searchable by topic and includes contextual information: what is known about what works, for whom, and under which conditions. To give a relevant example, a search under "School Organization and Governance" uncovers several studies of school turnaround models and their effects. A second source is the Best Evidence Encyclopedia, which is, in some ways, more user-friendly than the Clearinghouse, with an intuitive grading scale and grade for each reviewed intervention. This fall, the team producing this encyclopedia will release a new website that translates its existing evaluation data into ESSA's evidence tiers. For a broader source of information that covers multiple "out-of-school" topics, many of which intersect with student performance, see the "Results First Clearinghouse Database" from Pew.

To maximize the effects of its work, the center would develop ongoing relationships, wherever possible, with designated representatives from districts with low-performing schools. In certain cases, such as with the many rural districts in New York, organizations representing multiple districts (such as the BOCES in NY), could appoint a representative.

The Evidence Resource Center would also have primary responsibility for tracking outcomes of the districts' evidence-based policy interventions and for alerting SEA leadership about strong findings, either positively or negatively, on a regular basis.

Such a center could potentially use existing SEA staff members who have a background in research. Should this option prove impossible for fiscal or personnel reasons, Chiefs could engage the business and philanthropic communities with a vision for long-term reform, and solicit financial support to enable a university-based research center to provide the required work.

Building upon these efforts, a state Chief could lead the creation of an "Evidence Coalition," a working group composed of representatives from the Evidence Resource Center, the REL, districts with schools designated as low-performing; principals, teachers, and parents from such schools, and education researchers. Such a coalition would meet ESSA's stakeholder requirements and provide a vehicle for reinforcing the ongoing use of evidence. Supported by staffing from the Evidence Center, the coalition would meet regularly to review important policy interventions in the lowest performing schools, hear short briefings from SEA and/or scholars about new evidence around possible interventions, and report "from the ground" on the impact of current intervention strategies.

An early example of the kind of coalition that could produce this work can be found in the growing network of Researcher Practitioner Partnerships that receive funding and expertise from IES (the Institute of Education Sciences, which is the statistics, research, and evaluation arm of USED). IES supports the construction of a team drawn from the SEA, district offices, and research community whose purposes are twofold: to identify and research key educational challenges at the ground level, and to develop research-based interventions, track results, and disseminate findings.

New research is not always required, however: Often, the challenge may be to disseminate what is known in an effective form. Louisiana provides an example of the kind of work a state Chief could enjoin. Louisiana's Department of Education evaluates commonly used curricula and ranks them according to their alignment with state content requirements, partners with LearnZillion to create educator resources, and supports schools and districts in using high-quality instructional materials. The process of evaluating curricula is transparent and allows for public comment. The program is new, and its outcomes unstudied; the state cannot force its districts to choose strong

materials. However, the public nature of Louisiana's initiative illustrates how policymakers can use research to set a high standard for the entire state.

Another example illustrates what a coalition could accomplish. The Education Endowment Foundation (EEF), the UK's equivalent of the Institute of Education Sciences (IES), entered into a partnership with the Suffolk County Council called "Scaling up Evidence-based Programmes." Suffolk County's intention is to incentivize schools to use evidence-based practices and materials. Its Council provides matching funds for schools to implement one of nine instructional programs, all of which demonstrated positive effects on student learning in EEF evaluations. The nine programs include both whole class and also targeted interventions and encompass science, literacy, and mathematics instruction. The partnership's explanatory documents also list the additional months of progress associated with each intervention and the robustness of the research on a scale of 1-5, with 5 being the highest. For instance, "Mathematical Reasoning" is a whole-class program developed by Oxford University for Year 2 students.[4] EEF determined with a high level of confidence (level 5) that faithful implementation yields three additional months' worth of learning per year. "Switch On," a targeted reading and writing intervention delivered by assistant teachers in Years 3-7[5], also yields three additional months' worth of learning per year, with a confidence level of 3 (modest research base). Both partners work with schools to ensure fidelity of implementation, and the EEF is studying the results of the model. Schools are thus empowered to choose from a menu of scientifically vetted programs according to their students' needs.

In sum, ESSA's new language about evidence-based interventions can offer state Chiefs truly important leverage in moving districts and their schools through a major paradigm shift. Where once policies and the dollars that paid for them were based on habit or happenstance, the states, districts, and most importantly the schools and children of tomorrow, could be the beneficiaries of research-based interventions that will directly and measurably improve learning outcomes for students. This is not an opportunity to be squandered.

[4] The equivalent of the United States' 1st Grade.

[5] The equivalent of the United States' 2nd-6th Grades.

CITATIONS

Agodini, R., & Pendleton, A. (2009). Achievement effects of four early elementary school math curricula: findings from first-graders at 39 schools. Washington, DC: United States Department of Education. Retrieved from http://purl.access.gpo.gov/GPO/LPS117612

Barton, P. E., Coley, R. J., & ETS Policy Information Center. (2010). The black-white achievement gap: when progress stopped. Princeton, NJ: Policy Evaluation and Research Center, Policy Information Center, Educational Testing Service.

Bhatt, R., & Koedel, C. (2012). Large-Scale Evaluations of Curricular Effectiveness: The Case of Elementary Mathematics in Indiana. Educational Evaluation and Policy Analysis, 34(4), 391–412.

Bhatt, R., Koedel, C., & Lehmann, D. (2013). Is curriculum quality uniform? Evidence from Florida. Economics of Education Review, 34, 107–121. http://doi.org/10.1016/j.econedurev.2013.01.014

Borman, G. D., Slavin, R. E., Cheung, A. C. K., Chamberlain, A. M., Madden, N. A., & Chambers, B. (2007). Final Reading Outcomes of the National Randomized Field Trial of Success for All. American Educational Research Journal, 44(3), 701–731.

Boser, U. (2014). Return on Educational Investment 2014: A District-by-District Evaluation of U.S. Educational Productivity (p. 47). Washington, DC: The Center for American Progress.

Boser, U., Chingos, M., & Straus, C. (2015). The Hidden Value of Curriculum Reform (p. 51). Washington, D.C.: Center for American Progress. Retrieved from https://cdn.americanprogress.org/wp-content/uploads/2015/10/06111518/CurriculumMatters-report.pdf

California Department of Education. (2002). What we Have Learned About Class Size Reduction in California (Capstone Report). Sacramento, CA. Retrieved from http://www.classize.org/techreport/CSRYear4_final.pdf

Carr, D. (2003). Making sense of education: an introduction to the philosophy and theory of education and teaching. London; New York: RoutledgeFalmer.

Cavalluzzo, L., Barrow, L., Henderson, S., CNA Education, & National Board for Professional Teaching Standards. (2014). National Board Certification as Professional Development: An Empirical Study of the National Board for Professional Teaching Standards Process, Final Report. CNA Corporation.

CREDO. (2015). Urban Charter School Study: Report on 41 Regions. Stanford, CA: CREDO: Center for Research on Education Outcomes.

Devedžic, V., Jovanovic, J., Devedžic, V., & Jovanovic, J. (2015). Developing Open Badges: A Comprehensive Approach. Educational Technology Research and Development, 63(4), 603–620. Retrieved from https://www.learntechlib.org/p/161786/

Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, … Harris KA. (2010). Competency-based medical education: theory to practice. Medical Teacher, 32(8), 638–45.

Furgeson, J. R., Mathematica Policy Research, I., Daniel J. Evans School of Public Affairs, & Center on Reinventing Public Education. (2012). Charter-school management organizations diverse strategies and diverse student impacts. Princeton, N.J.; Bothell, Wash.: Mathematica Policy Research ; University of Washington, Center on Reinventing Public Education.

Haciomeroglu, G. (2013). Mathematics Anxiety and Mathematical Beliefs: What Is the Relationship in Elementary Pre-Service Teachers? Issues in the Undergraduate Mathematics Preparation of School Teachers, 5.

Hansen, M. (2013). Right-Sizing the Classroom: Making the Most of Great Teachers. Washington, DC: Thomas B. Fordham Institute. Retrieved from http://edex.s3-us-west-

2.amazonaws.com/publication/pdfs/20131119-Right-Sizing-the-Classroom-Making-the-Most-of-Great-Teachers-FINAL.pdf

Hattie, J. (2015). The applicability of Visible Learning to higher education. Scholarship of Teaching and Learning in Psychology Scholarship of Teaching and Learning in Psychology, 1(1), 79–91.

Hoxby, C. M., Murarka, S., & National Bureau of Economic Research. (2009). Charter schools in New York City who enrolls and how they affect their students' achievement. Cambridge, Mass.: National Bureau of Economic Research. Retrieved from http://papers.nber.org/papers/w14852

Kane, T. J., Rockoff, J. E., Staiger, D., & National Bureau of Economic Research. (2006). What does certification tell us about teacher effectiveness? evidence from New York City. Retrieved April 3, 2013, from http://papers.nber.org/papers/w12155

Lavy, V., Angrist, J. D., & National Bureau of Economic Research. (1997). Using Maimonides' Rule to Estimate the Effect of Class Size on Student Achievement. Cambridge, Mass.: National Bureau of Economic Research. Retrieved from http://papers.nber.org/papers/w5888

Michele Jolin Praises Senate Leaders on Evidence-Based Policy Provisions in ESSA. (2015, December 9). Retrieved June 22, 2016, from http://results4america.org/press-room/michele-jolin-praises-senate-leaders-evidence-based-policy-provisions-essa/

Papay, J. P., West, M. R., Fullerton, J. B., & Kane, T. J. (2012). Does an Urban Teacher Residency Increase Student Achievement? Early Evidence From Boston. Educational Evaluation and Policy Analysis, 34(4), 413–434. Retrieved from http://www.jstor.org/stable/23357021

Polikoff, M. (2014, August 18). At last, accountability for textbook publishers? Retrieved from http://edexcellence.net/articles/at-last-accountability-for-textbook-publishers

Reardon, S. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), Whither opportunity?: rising inequality, schools, and children's life chances.

Stanford University, & Center for Research on Education Outcomes. (2013). National charter school study 2013. Stanford, Calif.: Center for Research on Education Outcomes. Retrieved from http://credo.stanford.edu/research-reports.html

Strategic Data Project. (2012). SDP Human Capital Diagnostic: Los Angeles Unified School District. Harvard University: Center for Education Policy Research. Retrieved from http://cepr.harvard.edu/files/cepr/files/sdp-lausd-hc.pdf?m=1430928898

TNTP. (2015). The Mirage: Confronting the Hard Truth About our Quest for Teacher Development. New York, NY: TNTP.

United States Congress. (2015, December 10). S.1177 - 114th Congress (2015-2016): Every Student Succeeds Act. Retrieved June 22, 2016, from https://www.congress.gov/bill/114th-congress/senate-bill/1177

United States Department of Education. Elementary and Secondary Education Act of 1965, As Amended by the Every Student Succeeds Act - - Accountability and State Plans (2016).

Voorhees, R. A. (2001). Competency-Based Learning Models: A Necessary Future. New Directions for Institutional Research, 2001(110), 5–13. http://doi.org/10.1002/ir.7

Walser, T. M. (2014). Quasi-Experiments in Schools: The Case for Historical Cohort Control Groups. Practical Assessment, Research & Evaluation, 19(6).

West, M. R. (2016, February 5). From evidence-based programs to an evidence-based system: Opportunities under the Every Student Succeeds Act. Retrieved June 22, 2016,

from http://www.brookings.edu/research/papers/2016/02/05-evidence-based-system-opportunities-under-essa-west

Whitehurst, G. (2009). Don’t Forget Curriculum (Brown Center Letters on Education) (p. 12). Brookings.

Whitehurst, G. J., Chingos, M. M., Gallaher, M. R., & Brookings Institution. (2013). Do School Districts Matter? Brookings Institution.