School Leadership Interventions Under the Every Student Succeeds Act: Evidence Review

Updated and Expanded
This report was updated in January 2017 to include Appendix C

Rebecca Herman, Susan M. Gates, Aziza Arifkhanova, Andriy Bega, Emilio R. Chavez-Herrerias, Eugeniu Han, Mark Harris, Jennifer Tamargo, Stephani L. Wrabel
Acknowledgments

We are grateful to staff of The Wallace Foundation who provided input on the report—
during a briefing on key findings and through feedback on an earlier draft of this report. 
Specifically, we appreciate the probing questions and insights from Ed Pauly and Jody 
Spiro. Kathryn Young, Dan Gordon, and Scott Palmer of Education Counsel provided leg-
islative analysis and framing suggestions that markedly improved the quality of the report. 
We also thank Catherine Augustine and William Johnston from RAND and Jason Grissom 
from Vanderbilt University for feedback on the many iterations of this report, as well as 
Chandra Garber of RAND for her contributions to the text and accompanying documents. 
We take full responsibility for any errors.

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation 
of RAND intellectual property is provided for noncommercial use only. Unauthorized post-
ing of this publication online is prohibited. Permission is given to duplicate this document 
for personal use only, as long as it is unaltered and complete. Permission is required from 
RAND to reproduce, or reuse in another form, any of our research documents for commer-
cial use. For information on reprint and linking permissions, please visit www.rand.org/
pubs/permissions.

For more information on this publication, visit www.rand.org/t/RR1550-2.

The RAND Corporation

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND’s publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

www.rand.org

© Copyright 2017 RAND Corporation
PREFACE

The reauthorization of the U.S. Elementary and Secondary Education Act, referred to as the Every Student Succeeds Act (ESSA), emphasizes evidence-based initiatives while providing new flexibilities to states and districts with regard to the use of federal funds, including funds to promote effective school leadership.

The RAND Corporation conducted a synthesis of the evidence base on school leadership interventions to better inform the use of school leadership interventions under ESSA; identify examples of improvement activities that should be allowable; and guide education policymakers, practitioners, and thought leaders on the use of research-based practices. This report describes the opportunities for supporting school leadership under ESSA, discusses the standards of evidence under ESSA, and synthesizes the research base with respect to those standards. The information can guide federal, state, and district education policymakers on the use of research-based school leadership interventions; help them identify examples of improvement activities that should be allowable under ESSA; and support the use of such interventions. The report expands on the version first released in April 2016 (School Leadership Interventions Under the Every Student Succeeds Act: Volume I—A Review of the Evidence Base, Initial Findings, by Rebecca Herman, Susan M. Gates, Emilio R. Chavez-Herreras, and Mark Harris) by including tier IV evidence in the review, further analyzing the evidence requirements in ESSA, and providing considerations for policymakers and practitioners.

This research has been conducted in RAND Education, a division of the RAND Corporation, with grant funding from The Wallace Foundation. The Wallace Foundation is committed to improving school leadership through better training, hiring, support, and evaluation of principals. For more than a decade, it has invested in research, initiatives, and evaluations to improve school and district leadership and contribute to an evidence base in this area.
## Contents

**Preface** ................................................................................................................................. iii

**Chapter 1** Introduction ........................................................................................................... 1

**Chapter 2** How Does the Every Student Succeeds Act Support School Leadership Improvement? ................................................................. 4

**Chapter 3** How Does the Every Student Succeeds Act Define Evidence? .................. 7

**Chapter 4** What Is the Evidence of Effects of ESSA-Eligible School Leadership—Improvement Activities? ................................................................. 16

**Chapter 5** Recommendations for Using This Review .......................................................... 28

**Appendix A** State Exemptions to ESSA Evidence Requirements ........................................... 30

**Appendix B** Evidence Requirements Across Federal Legislation and Policy ..................... 31

**Appendix C** Summary of Updates to Evidence Report in December 2016 Release ........ 34

**Notes** ........................................................................................................................................ 37

**Bibliography** ................................................................................................................................. 39
CHAPTER 1. INTRODUCTION

The Every Student Succeeds Act (ESSA) presents a renewed focus on school leadership and acknowledges the importance of school principals to school improvement and effective instruction (Public Law No. 114-95, 2015). The act allows states and districts to use federal funds for activities targeting the quality of school principals and other school leaders.

ESSA repeatedly calls for the use of evidence-based activities, strategies, and interventions (Public Law No. 114-95, 2015). This makes good sense: Investments in education must produce results. Students’ efforts, teachers’ time, and scarce financial resources are more likely to be well spent when education-improvement activities are selected because there is evidence that they are effective. To select education-improvement activities without considering their proven impact may be seen as an irresponsible use of limited resources.

In particular, ESSA establishes a framework with tiers of evidence for consideration and use by policymakers and educators. While helpful, this framework does not resolve all questions about which leadership activities would meet these tiers to qualify as evidence-based and be most useful to states, districts, and schools. In the face of such ambiguity, the U.S. Department of Education’s (2016c) Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments (hereafter referred to

- School leadership can be a powerful driver of improved education outcomes.
- Activities designed to improve school leadership demonstrate positive impact on student, teacher, and principal outcomes, based on research that is consistent with ESSA evidence tiers.
- ESSA expands opportunities for states and districts to use federal funding for initiatives that strive to improve the quality of school leaders.
- ESSA’s evidence tiers provide a framework for using evidence in school leadership policy and practice.
- ESSA’s framework with tiers of evidence, coupled with the U.S. Department of Education’s nonregulatory guidance, strongly emphasizes the use of evidence in setting direction for improving school leadership and prioritizes more-rigorous evidence.
- ESSA provides avenues to consider and build the evidence base for new and underresearched interventions.
- ESSA provides some flexibility for states to interpret and apply evidence requirements.
as Evidence Guidance) provides some needed clarity. However, there remain some points of ambiguity. Clarification on these points would support broader application of the evidence tiers to education decisionmaking.

We interpret the language of the legislation in the context of other federal initiatives on education evidence, as well as the Evidence Guidance.

Although we feel that this work is well grounded, we acknowledge that others may provide different—and equally valid—interpretations of the ESSA evidence tiers. Further, our interpretations may change depending on future guidance. This report aims to be transparent about how we have defined each tier and our process for applying the criteria to the evidence we reviewed.

The RAND Corporation conducted a synthesis of the evidence base on school leadership efforts to inform the use of school leadership activities and interventions under ESSA. This report is intended to help federal, state, and district education policymakers understand and implement efforts to improve school leadership that are consistent with ESSA. A more limited version of this report was released in April 2016. Please see Appendix C for an explanation of the changes between the releases of the first and second reports.

In this report, we first offer an overview of the ways in which school leadership may affect outcomes of interest and then describe how school leadership is addressed by ESSA funding streams and statutory provisions. The key questions for this topic are

- What is the evidence that school leadership matters for school improvement?
- What school leadership–improvement activities are supported under ESSA?

We then describe the ESSA-defined tiers of evidence that such funding streams will require. We compare ESSA evidence tiers with evidence requirements for other federal education programs to identify ambiguities in the ESSA tiers. The key question for this section is

- How are the ESSA evidence tiers defined, how does current guidance clarify these tiers, and what further guidance might improve the use of these evidence tiers for education decisionmaking?

Having laid out relevant foci of ESSA as context, we then describe the evidence review. We provide a brief description of our methodology in reviewing the literature and then present findings on improvement activities that could reasonably be interpreted to fit within ESSA’s evidence framework, given the statute and guidance. The key question for this section is

- What is the evidence of effects of school leadership–improvement activities, as judged against the ESSA evidence tiers?

Finally, we offer recommendations to guide education policymakers, practitioners, and thought leaders on the use of research-based practices.
Why Focus on School Leadership?

States and districts have multiple ways to promote school improvement. What would justify a focus on school leadership? Research points to the value and importance of school leaders in driving student achievement gains. This research suggests that school leadership could be an important lever for school-improvement strategies pursued by states and districts.

In their comprehensive review of the literature, Leithwood et al. (2004) concluded that principals are second only to teachers as the most important school-level determinant of student achievement. That finding is significant in view of the reality that there are far fewer principals than teachers in a district and that each principal has the potential to affect the outcomes of far more students. That review was conducted more than ten years ago, but subsequent research has reinforced that basic finding (see, for example, Coelli and Green, 2012; Dhuey and Smith, 2014; Grissom, Kalogrides, and Loeb, 2015). A principal scoring one standard deviation above the mean for principal effectiveness could move the mean student achievement from the 50th to the 58th percentile (Branch, Hanushek, and Rivkin, 2012). Research also demonstrates that principals are important to key teacher outcomes. Teacher turnover is lower in schools led by high-quality principals (Boyd et al., 2011; Branch, Hanushek, and Rivkin, 2012; Grissom, 2011; Ladd, 2011), and more-effective principals retain and hire higher-quality teachers and have teachers who improve faster (Loeb, Kalogrides, and Béteille, 2012). Research further indicates that principal turnover leads to lower teacher retention and lower gains for students (Béteille, Kalogrides, and Loeb, 2012; Miller, 2013). 3

In sum, there is substantial research evidence demonstrating that school leaders are a powerful driver of student outcomes. This evidence base justifies ESSA’s investment in school leaders as part of school improvement. Although ESSA does not approach this level of specificity, federal, state, and district policymakers might consider guiding resources toward principal-improvement activities that have demonstrated impact on principals’ actions and characteristics that are associated with improved student outcomes. In the next chapter, we explore the opportunities to improve school leadership through ESSA.
CHAPTER 2. HOW DOES THE EVERY STUDENT SUCCEEDS ACT SUPPORT SCHOOL LEADERSHIP IMPROVEMENT?

Linking Elements of ESSA to School Leadership Activities

School leadership is explicitly acknowledged as a valid target of educational-improvement activities across the titles in ESSA; in many areas of the act where school leadership is not explicitly called out (e.g., school improvement efforts under Title I), states and districts could still choose to support leadership-focused activities in pursuit of school-improvement objectives. In this chapter, we provide a brief description of which elements of ESSA could be used for school leadership–improvement initiatives.

Title I (Improving Basic Programs Operated by State and Local Educational Agencies)

Title I of ESSA authorizes approximately $15.0 to $16.2 billion per year (2017–2020) to states in formula funding to improve basic state and local education programs. Title I provides broad allowable uses of funds, including the ability to support leadership. In addition, Title I has historically included a substantial investment in identifying and improving low-performing schools. Most recently, the School Improvement Grant program invested billions of dollars into supporting fundamental change in the lowest-performing schools in each state (see, e.g., U.S. Department of Education, 2016a). ESSA has replaced the School Improvement Grants with School Improvement Funds, which require a 7 percent set-aside of state Title I allotments, still focused on the lowest-performing schools (including the lowest-performing 5 percent of schools in a state, as well as high schools that fail to graduate one-third or more of their students). All schools receiving Title I school improvement funds must include in their plans proposed school improvement activities that demonstrate strong, moderate, or promising evidence of effects (see ESSA’s evidence tiers I, II, and III, below). School Improvement Funds may be used to support activities to improve school leaders. In fact, under previous versions of the program, funds were frequently directed toward replacing or improving principals.

Title II, Part A (Supporting Effective Instruction)

Title II, Part A, authorizes approximately $2.3 billion per year (2017–2020) to states in formula funding to improve the quality of teachers,
principals, or other school leaders.7 States may select from a broad array of state-level allowable uses of Title II formula funds to help improve leadership quality, and states may also reserve up to an additional 3 percent of the amount set aside for district subgrants for those activities designed to improve the principal pipeline, such as

- improving principal certification (regular and alternative), evaluation, and support systems
- improving preservice (principal preparation programs and academies)—within certain limitations
- providing training or professional development on such topics as differentiating performance; evaluating teachers; cultural competency; instruction and student learning; postsecondary coursework for students through dual enrollment or early college high school; and science, technology, engineering, and mathematics and career and technical education instruction
- recruiting, retaining, and training school leaders (among others)
- improving induction and mentoring for early-career principals
- differentiating pay for hard-to-fill positions
- offering more-focused opportunities not explicitly targeting school leadership, such as transition to elementary school and school readiness, pre-K–3 alignment, implementing bullying prevention and restorative justice practices, and sexual-abuse prevention.

Although some of these allowable Title II, Part A, uses of funds require an evidence base (from evidence tiers I–IV), others do not. However, the Department of Education’s recent guidance on evidence encourages states and districts to use the strongest evidence appropriate to the need (U.S. Department of Education, 2016c).

Title II, Part B (National Activities)
Title II, Part B, authorizes $469 to $489 million per year (2017–2020) for all parts of Title II, Part B, National Activities (including support for both teachers and principals), through which the Department of Education administers several programs. This includes the competitive Teacher and School Leader

---

Title I Leadership Support Under No Child Left Behind and ESSA

Both No Child Left Behind (NCLB) and ESSA provide resources for comprehensive school support and improvement, which can involve whole-school initiatives that focus on principals. NCLB specifies that school improvement plans must include strategies based on scientifically based evidence. Similarly, comprehensive support and improvement plans under ESSA must include evidence-based interventions. So, both laws require evidence of effects. The major difference between ESSA and NCLB in Title I is the greater flexibility under ESSA school improvement funds for states and districts to select or develop interventions. NCLB’s School Improvement Grant program, as implemented under the American Recovery and Reinvestment Act, gave states and districts a select set of interventions; for schools that had persistent low performance, all of the prescribed interventions explicitly required or implicitly suggested removing the principal. ESSA does not have that requirement. Title I comprehensive school reform under ESSA, unlike School Improvement Grants under NCLB, can be used for a very broad array of improvement initiatives—including principal-focused initiatives.
Incentive Fund, which allows states and districts to develop human-capital management systems, including performance-based incentives, such as bonuses for teachers or principals based on improved student achievement. These performance incentives can be used with both school leaders and teachers; before ESSA, the incentives only explicitly targeted teachers.

The Supporting Effective Educator Development (SEED) Grant Program is another national activity. The SEED program is a competitive federal grant program to support the development of effective educators, including school leaders, through nontraditional certification programs, evidence-based professional development on several topics (e.g., literacy, numeracy, incorporating postsecondary coursework in the K–12 curriculum), and other learning opportunities (e.g., learning through partnerships, activities leading to credentials).

Finally, ESSA national activities include the School Leader Recruitment and Support fund, a competitive grant program to support efforts to improve the recruitment, preparation, placement, support, and retention of effective principals or other school leaders in high-need schools. Such activities could include traditional or alternative preservice training programs; recruiting, selecting, developing, and placing leaders in high-need schools, with the purpose of implementing reforms; continuous professional development; and developing and disseminating information on best practices.

Both the SEED program and the School Leader Recruitment and Support fund prioritize applications that include activities that meet evidence tiers I–III. Otherwise, Title II supports for school leadership improvement can call on evidence tiers I–IV.

**Summary**

ESSA provides opportunities to improve school leadership by supporting school improvement programs that have a strong leadership component and by improving steps in the principal pipeline, such as preparation programs, certification, professional development, and recruitment and placement. These investments include formula funding with broad allowable uses, some of which contain expectations that activities meet one of the four evidence tiers (e.g., some Title II, Part A, activities); targeted formula funds with specific requirements around evidence tiers I–III (e.g., School Improvement Funds); and competitive grants that may be used to advance leadership, several of which contain competitive priorities around activities that meet evidence tiers I–III (e.g., the School Leader Recruitment and Support fund). In the next chapter, we explore ESSA evidence requirements, especially for school leadership–improvement activities.
CHAPTER 3. HOW DOES THE EVERY STUDENT SUCCEEDS ACT DEFINE EVIDENCE?

States and districts must show evidence of prior success for some of the leadership-improvement activities allowed under ESSA. The level and type of evidence depends on the exact source of funding for the intervention. We discuss the different tiers of evidence required by ESSA and identify which tiers are needed for each relevant ESSA funding stream.

To frame this discussion, we begin by presenting a theory of action that indicates paths by which school leadership initiatives may be linked to desired student outcomes to understand what types of evidence may be acceptable under ESSA standards.

Identifying What Counts as a School Leadership Initiative

ESSA calls for the use of evidence-based activities, strategies, and interventions. A key question for our review is what counts as an evidence-based school leadership initiative? Figure 1 describes the simplified theory of action that guided our review. In this theory, a catalyst for change, such as a state policy, drives policymakers and educators to focus on improving school leadership. They select and implement activities, strategies, or interventions designed to improve school leadership. These improvement activities change school leaders’ behaviors, which improve instruction, the school climate, and other teacher and school outcomes, which then improve student outcomes. ESSA evidence tiers focus on two parts of the theory: activities and outcomes. (In ESSA, outcomes are somewhat loosely defined, as explored:

Defining Evidence-Based Under the Every Student Succeeds Act

“(2) 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 [Title I], 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).”

Introduction to ESSA’s Evidence Tiers
ESSA defines four tiers of evidence, in order of rigor, for judging whether an activity is evidence-based. To be eligible for Title I School Improvement Funds, and to meet a competitive priority under some discretionary grant programs (e.g., SEED or the School Leader Recruitment and Support Fund), an activity must demonstrate evidence in one of the first three tiers. Otherwise, an activity with an evidence-based requirement must demonstrate evidence under any of the four tiers to be approved. The Evidence Guidance encourages the use of the strongest available evidence within the three or four allowable tiers; the legislation does not require it. To be evidence-based, an activity, strategy, or intervention must show statistically significant positive effects on student or other relevant outcomes, based on one or more of the following:

- Tier I (strong evidence)—at least one well-designed and well-implemented experimental study (randomized controlled trial)
- Tier II (moderate evidence)—at least one well-designed and well-implemented quasi-experimental study
- Tier III (promising evidence)—at least one well-designed and well-implemented correlational study that controls for selection bias.

For most federally funded school leadership–improvement activities other than those mentioned above, tier IV is also generally considered sufficient evidence:

- Tier IV (demonstrates a rationale)—the activity, strategy, or intervention demonstrates a rationale based on high-quality research or a positive evaluation that suggests it is likely to improve student or other relevant outcomes. For tier IV activities, there must be ongoing efforts to evaluate the effects of the activity, strategy, or intervention.
Exactly where the evidence requirements apply can vary by program—ESSA’s evidence requirements can be complicated and sometimes unclear. Under Title I, Part A (School Improvement), comprehensive and targeted programs must include evidence-based interventions—using evidence from tiers I, II, or III—to be funded. In Title II, however, some entire programs are required to be evidence-based, some programs are required to be evidence-based for some but not all components, and some programs are required to be evidence-based but can be exempted by the state if there is insufficient evidence. For example, Teacher and School Leader Incentive Fund grant applications must propose evidence-based projects (using evidence from tiers I, II, III, or IV). Residency programs on school leadership must include evidence-based coursework, but clinical experience and mentoring are not required to be evidence-based. Further, the coursework can be exempted from the evidence requirement at the state’s discretion. Appendix A offers more detail on these distinctions.

Understanding ESSA Evidence Tiers in Relation to Other Federal Education Evidence Requirements

ESSA does not break new ground in setting an expectation that evidence should inform education decisions. No Child Left Behind (NCLB) also established that expectation, using the term scientifically based research 69 times (Public Law 107-110, 2002). The Department of Education’s Institute of Education Sciences was established in 2002 “to provide scientific evidence on which to ground education practice and policy” and developed resources such as the What Works Clearinghouse (WWC) to support that objective (Institute of Education Sciences, undated). (See Appendix B for a comparison of ESSA evidence requirements and requirements in other federal policy.)

Although federal policy has maintained a consistent emphasis on using evidence in education decisionmaking, the approach has evolved. The ESSA evidence tiers are new and differ substantially from prior evidence requirements, such as those used by the WWC (Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2014). For example, the WWC only allows evidence at the tier I and tier II levels, while ESSA provides opportunities to use tier III and tier IV evidence. Despite some differences in the language of the law, the Department of Education’s Evidence Guidance draws on the WWC and the Education Department General Administrative Regulations (EDGAR) to further specify the evidence tiers. The Evidence Guidance refers to EDGAR for definitions of research design (randomized controlled trial and quasi-experimental study), relevant outcomes, sample parameters (large, multisite), and logic models. The Evidence Guidance refers to the WWC as a potential source for evidence (original studies and reviews) and suggests that WWC evidence standards can help judge whether a study is well designed and well implemented. In this way, the Evidence Guidance provides useful—although nonregulatory—information to help interpret ESSA’s evidence tiers. Table 1 identifies ambiguities in the ESSA evidence tiers, clarifications provided by the Evidence Guidance, areas where the Evidence Guidance does not provide clarifications (middle column, italics), and sources for the clarifications.

The Evidence Guidance provides suggestions that address many of those areas of ambiguity (U.S. Department of Education, 2016c). However, further clarification will likely be needed—particularly with regard to evidence tier IV.
Table 1. Clarifying ESSA Evidence Tier Ambiguities

<table>
<thead>
<tr>
<th>Sample size</th>
<th>ESSA Ambiguity</th>
<th>Evidence Guidance</th>
<th>Guidance Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>Does not refer to sample size</td>
<td>Recommends large, multisite sample</td>
<td>EDGAR, WWC</td>
</tr>
<tr>
<td>Context</td>
<td>Does not indicate whether study context matters</td>
<td>Recommends that the sample and setting in the research overlap with that in the site</td>
<td>Not specified</td>
</tr>
<tr>
<td>Flawed studies</td>
<td>Requires that experimental and quasi-experimental studies be well designed and well implemented to qualify as tier I or tier II evidence but is silent on how to consider randomized controlled trials and quasi-experimental studies that have some design or implementation flaw</td>
<td>Includes—but downgrades—somewhat flawed randomized controlled trials(^a) Does not discuss flawed quasi-experimental designs</td>
<td>WWC</td>
</tr>
<tr>
<td>Rationale</td>
<td>The lowest category of evidence (tier IV) under ESSA includes “a rationale based on high-quality research findings or positive evaluation” No clear definition of rationale No clear definition of high-quality research findings or positive evaluation</td>
<td>Defines a rationale as a “well-specified logic model” Does not define high-quality research findings</td>
<td>EDGAR</td>
</tr>
<tr>
<td>Ongoing evaluation</td>
<td>Does not specify who should conduct evaluation</td>
<td>Indicates that evaluator could be outside implementing site</td>
<td>Not specified</td>
</tr>
<tr>
<td>Relevant outcomes</td>
<td>Does not specify eligible outcomes</td>
<td>Recommends focusing on student outcomes or outcomes associated with program goals</td>
<td>EDGAR</td>
</tr>
<tr>
<td>Important findings</td>
<td>Does not mention substantively important findings</td>
<td>Does not mention substantively important findings</td>
<td>EDGAR, WWC use substantively important findings</td>
</tr>
<tr>
<td>Body of evidence</td>
<td>Focuses on a single positive finding, rather than preponderance of evidence</td>
<td>Recommends that the favorable findings must not be countered by unfavorable findings</td>
<td>EDGAR, WWC</td>
</tr>
</tbody>
</table>

SOURCES: Public Law No. 114-95, Every Student Succeeds Act, Title VIII, Sec. 8101, Definitions, December 10, 2015; Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2014; Evidence Standards; Education Department General Administrative Regulations (EDGAR), Code of Federal Regulations, Title 34, Subtitle A, Chapter 1, Part 77.1, Definitions That Apply to All Department Programs, as amended December 19, 2014; Education Department General Administrative Regulations (EDGAR), Title 34, Subtitle A, Chapter 1, Part 77, Definitions That Apply to Department Regulations, as amended December 19, 2014.

\(^a\) Somewhat flawed randomized controlled trials can be considered to meet WWC standards with reservations under some circumstances, putting them on par with well-designed and well-implemented quasi-experimental designs.
Unpacking the Tiers
To conduct this review, we developed functional definitions of the design element of studies associated with each evidence tier (see Table 2). These definitions are consistent with the ESSA evidence tiers and with the Evidence Guidance, with one exception: study size. School leadership—improvement activities, and studies of these activities, tend to be focused on small groups. Requiring sample sizes of 50 or more schools or 350 or more students would be prohibitive and not appropriate for programs serving school leaders. For example, some principal preparation programs—especially those that were tailored to meet regional needs—may serve cohorts that are smaller than 350 participants.15

Evidence Requirements for Tiers I–III (Strong, Moderate, and Promising Evidence)
The definitions of strong, moderate, and promising evidence in ESSA, like the WWC definitions, prioritize experimental and quasi-experimental study designs. Experimental studies that involve the random assignment of participants to intervention and control groups are potentially tier I studies. Potentially tier II studies are those using a quasi-experimental design. These studies do not use random assignment to sort participants into treatment and control groups. The comparison group may be constructed by (1) assigning participants nonrandomly to groups and then administering the intervention to a group or (2) analyzing existing data, comparing those exposed to the intervention with those not exposed. In either case, analysis must demonstrate that the treatment and comparison groups were equivalent prior to the start of the intervention. Correlational studies with robust controls potentially constitute tier III studies. These studies use existing data to examine the relationship between the intervention and one or more relevant outcomes.

Table 2. Classifying Evidence Against ESSA Evidence Tiers: Study Design

<table>
<thead>
<tr>
<th>Design</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potentially Tier I (Strong)</td>
<td></td>
</tr>
<tr>
<td>Experimental study</td>
<td>Random assignment of participants to intervention and control groups</td>
</tr>
<tr>
<td>Potentially Tier II (Moderate)</td>
<td></td>
</tr>
<tr>
<td>Quasi-experimental study</td>
<td>Nonrandom assignment of participants to intervention and comparison groups by (1) providing intervention to one group or (2) using existing data, identifying a comparison group of nonparticipants. Must demonstrate that the groups were equivalent before the intervention started\a</td>
</tr>
<tr>
<td>Potentially Tier III (Promising)</td>
<td></td>
</tr>
<tr>
<td>Correlational study with controls</td>
<td>Using existing data, correlations between intervention status and outcomes must control for factors related to selection bias (e.g., participant demographics, prior associated outcomes)</td>
</tr>
<tr>
<td>Potentially Tier IV (Research-Based Rationale)</td>
<td></td>
</tr>
<tr>
<td>Rationale</td>
<td>Well-specified logic model that builds on high-quality prior research or a prior positive evaluation</td>
</tr>
</tbody>
</table>

NOTES: The study design only provides part of the requirements for each tier. For tiers I–III, the evidence must have the identified design and must be well designed and implemented. For tier IV, the evidence must have the identified design and included ongoing efforts to examine the effects.\a Quasi-experimental designs that introduce known pre-intervention differences (i.e., regression discontinuity designs) also qualify.
controlling for factors that might be related to selection bias, such as prior outcome levels or participant
background characteristics. According to ESSA, findings must be statistically significant and positive to be
considered for tiers I–III.16

**Evidence Requirements for Tier IV (Research-Based Rationale)**

While tiers I through III describe levels of evidence, tier IV does not provide direct evidence of impact.
Rather, tier IV requires a research-based rationale to believe that the intervention will have the desired
impact, coupled with ongoing evaluation of the intervention to build an evidence base on the impact
of that intervention. The Evidence Guidance defines *rationale* for tier IV: “*Logic model* (also referred to
as theory of action) means a well-specified conceptual framework that identifies key components of the
proposed process, product, strategy, or practice (i.e., the active ‘ingredients’ that are hypothesized to be
critical to achieving the relevant outcomes) and describes the relationships among the key components
and outcomes, theoretically and operationally.”17

Tier IV was particularly challenging to operationalize in this review. Numerous blogs, commentar-
ies, and articles posted during and since ESSA’s passage have highlighted concerns about the ambiguity
in the definition of tier IV evidence (see, e.g., Advanced Education Measurement, 2016; Slavin, 2015;
West, 2016). A broad interpretation of the ESSA tier IV language would permit more interventions to be
implemented under ESSA, potentially increasing opportunities for states and districts to develop interven-
tions for their own context or to build evidence on popular but underresearched interventions. A narrow
interpretation of the ESSA tier IV language would encourage greater use of interventions most likely to
be effective, those with strong research-based rationales. We aimed to walk a line between extremes. For
this review, we made the following assumptions, based as much as possible on the language of ESSA, the
Evidence Guidance, and sources cited in the Evidence Guidance:

- The logic model should be shown graphically (e.g., a series of boxes and arrows). According to the
  Logic Model Workshop Toolkit (Shakman and Rodriguez, 2015, p. 3), also cited in the Evidence
  Guidance, a “logic model is a visual representation of a theory of action or program logic guiding the
design and implementation of a program or policy.” We did not consider such statements as “X causes
Y” to be a logic model.
- The logic model should include—at minimum—key components of the intervention and outcomes,18
  and it may additionally include resources needed to implement the intervention, outputs, and different
  types of outcomes, such as short-term, intermediate, and long-term outcomes (Shakman and Rodri-
guez, 2015).
- The research findings or evaluations that support the logic model should show that at least one compo-
nent of the intervention corresponds to at least one desired outcome.19

Operationally, we made some decisions to facilitate the review:

- Our primary source of information was the set of documents collected for the evidence review, which
  mainly included studies but also some descriptive documents.20 There may be interventions that
  qualify for tier IV but did not appear in the review because they did not have studies.

  - *What does this mean for you?* If your intervention does not appear in the review, it might still have
    qualified as tier IV. Look for a logic model in descriptive materials (e.g., on the developer’s website,
in brochures or pamphlets).
We relied on an author review of the research and evaluation supporting the logic model, rather than reviewing all of the cited studies.

- *What does this mean for you?* If you want to have greater confidence in the potential impact of the intervention, you may want to look at the studies cited as support for the intervention’s logic model. Ideally, the supporting research would be strong (tier I) or promising (tier II).

- *We did not examine whether the intervention was part of an ongoing evaluation.*

- *What does this mean for you?* If you plan to implement an intervention under tier IV, you should determine whether others are currently studying it, or plan to evaluate it yourself.

Figure 2, a more detailed version of Figure 1, provides a graphical representation of tier IV evidence under ESSA, provided that there is ongoing evaluation of the intervention.

Evidence consistent with this definition would provide valuable—albeit not WWC-level rigorous—information where there is a gap in existing evidence and a practical need to implement reform.

As noted above, we developed and applied operational definitions for the ESSA evidence tiers based on the legislative language and prior federal initiatives; other researchers and policymakers might legitimately develop alternative definitions. Tier IV evidence is especially open to interpretation. Some interpretations might have a higher standard, and other interpretations might have a lower standard. In operationalizing tier IV, we aimed to balance consideration of newly developed interventions that have not yet established a rigorous evidence base with an expectation that education decisions should be informed with reasonably rigorous evidence.

### What Is Not Considered Evidence Under ESSA?

Following the functional guidelines described in Table 2 for this review, there were three common reasons a school leadership study did not qualify as evidence under ESSA. The first is because the study, while related to school leadership, does not examine an activity, strategy, or intervention. Examples of studies that would fail the evidence review for this reason include the following:

- Descriptive analyses of statistical patterns, trends, or relationships, such as a study of the correlation between principal leadership styles and student outcomes: In this example, the study does not provide information on interventions designed to improve principals.

- Case studies that are purposefully selected to identify common patterns or themes: In this situation, the outcomes in the successful cases might relate to one or more interventions, but the study design does not allow one to conclude that a single intervention caused the outcomes.
A second reason that some studies did not qualify as evidence under ESSA was because the intervention examined was not designed to improve principals. For example, an intervention designed to improve teachers’ instruction, in which the principal plays a role, is not expressly designed to improve the actions of principals.

The third reason some documents failed this review was because they did not include systematic analyses of evidence. Such studies could include purposefully selected anecdotes about the success of the improvement activity; analysis of untested, irrelevant, or not-validated outcomes (e.g., opinion surveys); and theory presented without any outcome analysis. In our review, these evidence limitations explained why many reviewed documents did not meet the ESSA evidence tiers.

Finally, states and districts faced with the challenge of conducting labor-intensive evidence reviews would benefit from being able to access existing rigorous reviews. Although ESSA does not indicate whether a research review that summarizes findings from a set of studies might itself be considered sufficient to meet standards, the Department of Education’s Non-Regulatory Guidance for Title II, Part A, suggests that this might be an option, by referring to an earlier version of the current report: “An additional resource that SEAs [state education agencies] and LEAs [local education agencies] may consider when selecting evidence-based interventions related to school leadership is School Leadership Interventions under the Every Student Succeeds Act from RAND Corporation” (U.S. Department of Education, 2016b, p. 15).

ESSA Ambiguity Regarding Outcomes

To be considered evidence-based according to the ESSA evidence tiers, an improvement activity must demonstrate “a statistically significant effect on improving student outcomes or other relevant outcomes.”

Although the legislation does not define student outcomes or other relevant outcomes, the Evidence Guidance suggests that outcomes should be consistent with Part 77.1 of EDGAR: “[T]he student outcome(s) (or the ultimate outcome if not related to students) the proposed process, product, strategy, or practice is designed to improve; consistent with the specific goals of a program.” Other relevant outcomes might include outcomes not necessarily at the student level. For this study of school leadership–improvement interventions, such outcomes as principal skills, teacher instruction, and school climate also are “relevant outcomes.” The Department of Education confirmed this interpretation in response to questions (U.S. Department of Education, 2016d).

School leader–improvement activities might be judged by their impact on teacher outcomes known to improve student outcomes, such as greater use of effective instructional practices or increased retention of highly effective teachers. This interpretation might be very appropriate for activities known to take some time to affect students. Principals’ impact on students is mainly filtered through changes to teachers and instruction (Hallinger, 2011; Heck and Hallinger, 2014). An intervention that improves instruction, which then improves student learning, can magnify the breadth of the impact but also may take longer than an intervention that focuses on an individual child. A study of the impact of a principal intervention on instruction may be feasible and informative where a study of the impact of the principal intervention on students is not.
ESSA on Interventions: Branded Versus Nonbranded

ESSA evidence tiers clearly and consistently focus on an “activity, strategy, or intervention.” Research is relevant in as much as it demonstrates that an education activity, strategy, or intervention is likely to produce the desired effect. ESSA does not, however, define what might be considered an activity, strategy, or intervention. Some interventions are created by developers and marketed as “branded” interventions. Others are developed locally and may not have a recognizable brand name. Although either can be supported by tiers I–IV evidence, branded interventions may be more likely to have a robust research base, because the developers are motivated to demonstrate impact to potential customers. This does not, however, mean that only branded interventions may be implemented under ESSA. In fact, a check-the-box approach to adopting only branded interventions named in this report runs counter to the spirit of ESSA, which provides states and districts new flexibilities to develop homegrown interventions most suited to their contexts and approaches.

States and districts may opt to replicate a branded program in their own contexts. If the unbranded program shares all of the components of the branded program and research on the branded program meets other ESSA requirements, then that research can be used to justify the unbranded replication. According to the U.S. Department of Education (2016d), “[t]he label or brand attached to a program or intervention included in a research study is less important than the activities, strategies, and practices that constitute that program or intervention.”

To help educators consider the research-based interventions that might be replicated in their own contexts, in Tables 6–11, we provide brief descriptions of each intervention for which we list findings.

Finally, an unbranded intervention that has tiers I–III evidence of effects or has a well-specified logic model and is subject to ongoing evaluation (tier IV) would meet ESSA evidence requirements.

Outstanding Questions Regarding Evidence Under ESSA

Although the Evidence Guidance addresses many of the ambiguities about ESSA evidence tiers, the following questions on evidence under ESSA remain:

- whether and how to require interventions to demonstrate substantively important findings (findings that have a large effect size)
- how to treat flawed quasi-experimental studies
- whether and how to consider evidence that is relevant to but not clearly about an intervention
- how to differentiate between “good” and “bad” tier IV evidence
- whether and how to use research reviews and syntheses.

We anticipate that federal or state departments of education might provide further suggestions or guidance to help apply the evidence tiers, because ESSA is silent on some points that have been important in past Department of Education evidence standards. Our current review of school leadership–improvement activities casts a broad net to include studies that meet ESSA’s evidence tiers, as specified in the legislation, Evidence Guidance, and in our Table 2. Findings may be subject to change, depending on guidance or information provided by federal or state departments of education.
CHAPTER 4. WHAT IS THE EVIDENCE OF EFFECTS OF ESSA-ELIGIBLE SCHOOL LEADERSHIP–IMPROVEMENT ACTIVITIES?

Because of the opportunities for funding school leadership interventions under ESSA discussed above, RAND conducted a critical literature review of the evidence regarding the effects of school leadership–improvement activities. The goal of this review was both to assess the current state of evidence for school leadership–improvement activities and to provide a model for administrators seeking to grow the evidence base around unproven interventions. In this chapter, we describe the methodology used to conduct our review and then report our key findings.

Methodology for Our Review of the Literature on School Leadership–Improvement Activities

The review of the evidence is framed by ESSA funding streams and evidence requirements. Table 3 shows the relationship graphically. ESSA provides funding for school leadership improvement through Title I; Title II, Part A; and Title II, Part B. Therefore, we review activities that fit the ESSA funding stream definitions. In addition, ESSA requires Title I school improvement activities to be supported by tiers I through III evidence, and Title II activities (when required to be evidence-based) should be supported by tiers I through IV evidence. Table 4 provides an overview of the literature review’s study-inclusion criteria. Again, although we drew on the best available resources to develop the process for applying ESSA evidence tiers, we recognize that reasonable people might take different approaches.

We conducted two waves of literature searches, adding terms and intervention names from the initial review to the second wave. We identified and prescreened more than 3,500 articles based on their titles, conducted a two-stage screening of more than 500 articles based on their abstracts, and fully reviewed 128 articles. For the final stage, two coders reviewed each article, and senior staff reconciled any differences. By the end of this process, studies that are reported here were examined by five researchers.

To help readers effectively process the findings from our literature review, we organized our discussion of evidence around three broad categories of school leadership–improvement activities that, together, include ESSA-allowable school leadership–improvement activities. First, states and districts can monitor

<table>
<thead>
<tr>
<th>ESSA Requirement</th>
<th>Evidence Required for Funding</th>
<th>Activity Eligible for Review (funded by ESSA and relevant to school leadership)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title I: School Improvement</td>
<td>Tiers I, II, III</td>
<td>Comprehensive and targeted school reforms with substantial school leadership component</td>
</tr>
<tr>
<td>Title II, Part A: Supporting Effective Instruction</td>
<td>Tiers I, II, III, IV</td>
<td>Pipeline activities for principals (certification, evaluation, mentoring, preservice, professional development, recruitment/retention, induction/mentoring, pay)</td>
</tr>
<tr>
<td>Title II, Part B: National Activities</td>
<td>Tiers I, II, III, IV</td>
<td>Performance-based human-capital management systems; pipeline activities spanning districts and states; competitive grants</td>
</tr>
</tbody>
</table>
whether principals are meeting performance expectations (i.e., state or district principal evaluation systems). Second, states and districts can take actions to improve the likelihood that school leaders actually meet those expectations through effective management structures, operations, and requirements. This category includes four subcategories: principal preparation programs, strategic staff management, professional learning, and working conditions. Finally, states and districts can improve school leadership through broader school improvement efforts that include leadership enhancements as a key component.

Findings from Our Review of the Evidence

In this section, we present the findings for school leadership–improvement activities that have evidence consistent with ESSA evidence tiers. Across the types of school leadership–improvement activities, we found in this review that several had tiers I through IV evidence. This includes tiers I through III evidence on comprehensive school improvement interventions that feature school leadership as a core component. Such activities would be eligible for funding under Title I School Improvement Funds. Table 5 provides a summary of the findings, and Tables 6–10 provide more detail by type of leadership improvement activity. Tables 6–10 also provide brief descriptions of the elements of each intervention, as of November 2016.²⁴

Leader-Evaluation Systems

Leader-evaluation systems are a set of processes, tools, and metrics designed to evaluate principals’ strengths and needs—for either accountability or developmental purposes. In theory and policy, these systems should be aligned with rigorous leadership standards (e.g., state standards or the Professional Standards for Edu-

---

Table 4. School Leadership–Improvement Activities Review: Study-Inclusion Criteria

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement activity</td>
<td>School leadership certification, evaluation, mentoring, preservice, professional development, recruitment/retention, induction/mentoring, or pay; School reform with substantial school leadership component; and other activities considered on a case-by-case basis</td>
</tr>
<tr>
<td>Sample and setting</td>
<td>U.S. K–12 public schools (including charter and magnet schools with open admission)</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Student achievement, progression, affective outcomes; teacher instruction, affective, impact, and mobility outcomes; school climate; principal leadership behaviors</td>
</tr>
<tr>
<td>Recency</td>
<td>Document released from 1995 to 2016</td>
</tr>
<tr>
<td>Study design</td>
<td>Tier I: random assignment&lt;br&gt;Tier II: nonrandom assignment&lt;br&gt;Tier III: correlational study with controls&lt;br&gt;Tier IV: research-based rationale</td>
</tr>
<tr>
<td>Study implementation (applies to tiers I–III)</td>
<td>Equivalence: (1) baseline similarities between groups on relevant factors (e.g., participant demographics, prior outcomes) or (2) analytic control on relevant factors&lt;br&gt;Confound: no conditions, other than those being studied, that affected one group that did not affect the other(s)</td>
</tr>
</tbody>
</table>
Table 5. Summary of Tiers I–IV Evidence on the Effects of School Leadership–Improvement Activities

<table>
<thead>
<tr>
<th>Activity</th>
<th>Evidence Base (number of studies)</th>
<th>Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leader-evaluation system</td>
<td>Tier IV (2)</td>
<td>Promising models with research-based theory of action</td>
</tr>
<tr>
<td>Principal preparation programs</td>
<td>Tier II (2), tier III (2)</td>
<td>Student-achievement gains, principal competency gains</td>
</tr>
<tr>
<td></td>
<td>Tier IV (3)</td>
<td>Models with research-based theory of action</td>
</tr>
<tr>
<td>Strategic staff management</td>
<td>Tier III (1)</td>
<td>Negative findings for principal change</td>
</tr>
<tr>
<td>Professional learning</td>
<td>Tier I (1), tier II (2), tier IV (3)</td>
<td>Positive or no effect on student achievement; reduced staff turnover; promising coaching model with research-based theory of action</td>
</tr>
<tr>
<td>Working conditions</td>
<td>Tier II (2), tier IV (3)</td>
<td>Mixed effects of autonomy on achievement; incentive/evaluation system correlates with higher student achievement; school administration manager corresponds to more instructional activity time</td>
</tr>
<tr>
<td>School improvement models</td>
<td>Tier I (1), tiers I &amp; II (2), Tier II (5)</td>
<td>Positive effects on student achievement</td>
</tr>
</tbody>
</table>

Table 6. Tiers I–IV Evidence on Leadership Evaluation Programs

<table>
<thead>
<tr>
<th>Activity</th>
<th>Study</th>
<th>ESSA Tier</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanderbilt Assessment of Leadership in Education: 360-degree principal evaluation tool designed to inform evaluation and professional growth</td>
<td>Porter et al., 2006</td>
<td>Tier IV</td>
<td>Prior research supports theory of action: feedback on performance improves leadership behaviors, which improve school performance and student success</td>
</tr>
<tr>
<td>Marzano School Leader Evaluation Model: Tool to evaluate school leaders</td>
<td>Carbaugh, Marzano, and Toth, 2014, 2015</td>
<td>Tier IV</td>
<td>Prior research supports the use of key components (see model); the model is hypothesized to improve student achievement</td>
</tr>
</tbody>
</table>
cational Leaders; see National Policy Board for Educational Administration, 2015) and draw on multiple perspectives (e.g., the principal’s supervisor, teachers, parents) and types of data (e.g., student-achievement data, observations, surveys). Often, the evaluation systems are developed by the state education agency or district and tap into nationally available tools. We found no research that met tiers I through III criteria on the impacts of state or district evaluation systems or evaluation tools on students or other relevant outcomes. However, we did identify two leader-evaluation tools that qualified as having tier IV evidence: the Vanderbilt Assessment of Leadership in Education (Porter et al., 2006) and the Marzano School Leader Evaluation Model (Carbaugh, Marzano, and Toth, 2014). Both of these tools are grounded in rigorous prior research that supports the theory of action.

**Principal Preparation Programs**

Principal preparation programs, broadly defined, involve a classroom-based program and some type of school-based internship and can lead to an advanced degree or certification. They may be provided by universities, districts, or independent organizations, or some combination of the three. ESSA defines principal preparation programs as operated by a public or other nonprofit organization (including or affiliated with an institution of higher education), containing a clinical preparation course (where the student is paired with an effective educator) and instruction in content areas, committed to producing a specified number of effective educators, and requiring demonstrated effectiveness to award a certificate or degree. ESSA also defines school leader residency programs—a type of preparation program—as school-based, with one year of learning and leading in an authentic school setting, as well as concurrent evidence-based coursework and mentoring from an effective principal.

There is substantial case-study research identifying components common to expert-identified effective preparation programs but less rigorous research on the effects of preparation programs overall or on specific programs. Table 7a summarizes four studies that provide evidence of effectiveness for principal preparation programs at tiers II and III. Two tier III studies showed positive relationships between characteristics of preparation programs and principal behaviors, teacher staffing, and achievement (Braun, Gable, and Kite, 2008; Fuller, Young, and Baker, 2011). One specific preparation program—New Leaders—has tier II evidence showing positive outcomes and would be considered evidence-based according to the ESSA definition (Gates, Hamilton, et al., 2014). One preparation program—the Texas Principal Excellence Program—has tier II evidence that shows no statistically significant relationship between student achievement outcomes and program participation after one year but significant improvements on three of nine principal competencies (Fouche, 2012).

An additional three studies provide evidence for principal preparation initiatives or components of those initiatives at the tier IV level, as summarized in Table 7b. These include studies of numerous preservice training programs (Darling-Hammond et al., 2007), a principal residency network (Braun, Billups, and Gable, 2013), and district-university preservice collaborations (Turnbull, Riley, and MacFarlane, 2013).

Few states currently require principal preparation programs to provide evidence of positive outcomes, such as principal retention rates or impacts on student learning, although some states (e.g., North Carolina, Ohio) are moving toward report cards for preparation programs (Briggs et al., 2013; Yoder, Freed, and Feters, 2014). Although there are no readily available ratings, there are tools for rating programs. For example, *Quality Measures™ Principal Preparation Program Self-Assessment Toolkit: For Use in Developing, Assessing,
Table 7a. Tiers I–III Evidence on Principal Preparation Programs

<table>
<thead>
<tr>
<th>Activity</th>
<th>Study</th>
<th>ESSA Tier</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texas Principal Excellence Program: Summit meetings, workshops, and webinars; 360-degree assessment; online learning tools; ongoing support from consultants</td>
<td>Fouche, 2012</td>
<td>Tier II</td>
<td>No statistically significant effects on student achievement outcomes for program participants after one year; statistically significant improvement in three of nine competencies according to 360-degree evaluation</td>
</tr>
<tr>
<td>New Leaders: Prepares principals to address the achievement gap and related challenges in high-need urban schools. Three core elements: selective recruitment and admissions, residency-based training, and endorsement and support for principals early in their tenures</td>
<td>Gates, Hamilton, et al., 2014</td>
<td>Tier II</td>
<td>Generally larger student-achievement gains in math and reading than in comparable schools, with differing effects across districts</td>
</tr>
<tr>
<td>Principal-preparation programs: Usually include classroom-based education, school-based internship, and advanced degree or certification</td>
<td>Fuller, Young, and Baker, 2011</td>
<td>Tier III</td>
<td>Positive association between some characteristics of principal-preparation programs (such as being housed at research or doctoral institutions) and improvements in the qualifications of teachers</td>
</tr>
<tr>
<td>Principal-preparation program practices: Emphasis on leadership for instructional improvement and achievement and reflection; included problem-based learning, alignment to standards, key content, individualized development, mentoring, cohort learning, performance assessments, and internships</td>
<td>Braun, Gable, and Kite, 2008</td>
<td>Tier III</td>
<td>Statistically significant correlations between preparation program practices and self-reported leadership practices and student English language arts achievement</td>
</tr>
</tbody>
</table>

and Improving Principal Preparation Programs (King, 2013) provides rubrics and indicators for programs to self-assess their preparation-program content, pedagogy, clinical practice, recruitment and selection, and graduate performance outcomes.

**Strategic Staff Management**

Strategic staff management may include activities to improve recruitment and selection processes, placement of principals in schools, and principal replacement. Recruitment and retention interventions may
include, for example, communication strategies to broaden the candidate pool or specialized processes and tools to screen and evaluate candidates (e.g., performance-based interview tasks).

Some researchers and policymakers have argued that replacing a principal is a necessary step to improving persistently low-performing schools, both to improve the quality of leadership and to create a disruption in dysfunctional processes that hinder school-wide reform (Hassel and Hassel, 2009; Le Floch et al., 2014). However, studies also have indicated that principal effectiveness increases with experience, suggesting that limiting turnover could improve outcomes (Clark, Martorell, and Rockoff, 2009). We identified only one tier III study examining the implications of principal turnover for student and other school-level outcomes. Table 8 summarizes the findings from the one tier III study: Changing principals does not correspond to achievement gains. This is consistent with findings from the broader literature relating principal turnover to student achievement, which finds either no relationship or a negative relationship between turnover and outcomes (Hochbein and Cunningham, 2013; Béteille, Kalogrides, and Loeb, 2012; Miller, 2013). Based on this review, principal replacement would not be considered evidence-based according to ESSA at this time.

### Professional Learning

Professional learning generally involves a variety of learning experiences for sitting school principals, such as professional development through workshops (single sessions or a series) and coaching or mentoring.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Study</th>
<th>ESSA Tier</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal Residency Network: Residency with mentor, individualized learning plan, action research project based on data analysis, participation in network events</td>
<td>Braun, Billups, and Gable, 2013</td>
<td>Tier IV</td>
<td>Prior research supports key components of the network</td>
</tr>
<tr>
<td>Exemplary preservice and in-service programs</td>
<td>Darling-Hammond et al., 2007</td>
<td>Tier IV</td>
<td>Presents conceptual framework for effective preservice and in-service, supported by evidence</td>
</tr>
<tr>
<td>Principal Pipeline Initiative: Sets leader standards; aligns districts' practices to recruit, train, and support new principals with standards, working in collaboration with local universities</td>
<td>Turnbull, Riley, and MacFarlane, 2013; Turnbull, Riley, et al., 2013</td>
<td>Tier IV</td>
<td>Prior research supports theory of action</td>
</tr>
</tbody>
</table>
These opportunities may be available throughout the principal’s career, although they often are most intensive early in his or her career or placement at a school. Principals have other learning experiences, such as attending conferences, which we do not include here because they are neither intensive enough to mobilize improvement nor discrete enough to evaluate.

As presented in Table 9, mixed outcomes were reported for two professional-development activities that had tier I or tier II evidence. Two studies showed positive effects on student achievement, and another showed greater staff stability in treatment schools but no effect on student achievement or instructional climate. Based on this review, the National Institute for School Leadership Executive Development Program would be considered evidence-based according to ESSA standards. McREL’s Balanced Leadership Program could be considered evidence-based because the one tier I study found reduced teacher turnover; however, the study also found no impact on student achievement (Jacob et al., 2014). There is also tier IV evidence supporting the Metropolitan ISD Principal Coaching Initiative, which provides district coaching for novice and experienced principals (Lee, 2010), and the Arkansas Leadership Academy’s Master Principal Program, which trains exemplary principals to be master principals (Peer, 2012), as well as Socratic Coaching (Lindle et al., 2015).

**Working Conditions**

Working conditions can include opportunities and incentives to improve teaching and learning. For this report, we focus on working conditions designed specifically to improve the effectiveness of school leaders, such as school autonomy and performance incentives or targeted support initiatives aimed to improve a principal’s efficiency. There are many other working conditions (e.g., school climate) that likely mediate or moderate leaders’ effectiveness but are not the focus here.

Principal-autonomy initiatives typically devolve decisions—such as hiring and removing teachers, budget, and school schedule—from district staff to school leaders. Autonomy initiatives focus on teaching and learning and building school capacity and may involve district offices to help support implementation (Honig and Rainey, 2012).

Two tier II studies showed mixed results of *principal autonomy* (see Table 10). Based on this review, principal autonomy might meet ESSA evidence standards, which currently only require one positive finding. There is substantial theory and qualitative evidence suggesting that school autonomy might be more effective if implemented in a way that overcomes known challenges (e.g., principals having little knowledge or experience to use their new flexibilities). Schools and their districts have struggled to implement autonomy (Hansen and Roza, 2005; Honig and Rainey, 2012). Yet there has been very little sustained research to help districts and schools overcome substantial barriers, such as costs, inefficiencies of decentralized authority, and union or legal constraints. Ultimately, theory and some empirical evidence suggest that

<table>
<thead>
<tr>
<th>Activity</th>
<th>Study</th>
<th>ESSA Tier</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal replacement: Removing the sitting principal and installing a new principal</td>
<td>Dhuey and Smith, 2014</td>
<td>Tier III</td>
<td>Installing a new principal correlates with achievement losses; study does not provide support for principal replacement</td>
</tr>
</tbody>
</table>
school-level autonomy can improve school functioning and student outcomes, but implementation challenges have consistently plagued efforts.

The effect of financial incentives for principal performance is not yet demonstrated through tiers I through III evidence, but there is tier IV evidence supporting this strategy (Hamilton et al., 2012).

<table>
<thead>
<tr>
<th>Activity</th>
<th>Study</th>
<th>ESSA Tier</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>McREL Balanced Leadership Program: Professional development, focus on 21 leadership responsibilities, case study approach, continuous feedback from facilitators, peer-to-peer interactions; ten two-day cohort-based workshops</td>
<td>Jacob et al., 2014</td>
<td>Tier I</td>
<td>No impact on student achievement or teacher-reported instructional climate; lower staff turnover in treatment schools</td>
</tr>
<tr>
<td>National Institute for School Leadership Executive Development Program: Professional development over 1 to 1.5 years, using group discussions, role-playing, video case studies, technology-assisted simulations, online learning, action learning projects; professional learning community</td>
<td>Nunnery et al., 2011</td>
<td>Tier II</td>
<td>Positive effects on reading and math achievement</td>
</tr>
<tr>
<td></td>
<td>Nunnery, Ross, and Yen, 2010</td>
<td>Tier II</td>
<td>Statistically significantly higher achievement gains in reading and math</td>
</tr>
<tr>
<td>Metropolitan ISD Principal Coaching Initiative: District coaching for novice and experienced principals, curriculum-based professional development, networking, visits to successful schools</td>
<td>Lee, 2010</td>
<td>Tier IV</td>
<td>Principals were more learner-centered following coaching; conceptual framework based on research</td>
</tr>
<tr>
<td>Coaching: Model for training districts staff to coach principals across rural districts on Socratic questioning and reflective listening</td>
<td>Lindle et al., 2015</td>
<td>Tier IV</td>
<td>Logic model based on research findings; pilot test found that training for coaches should include recognition of the emotions brought to the coaching relationship</td>
</tr>
<tr>
<td>Arkansas Leadership Academy’s Master Principal Program: Develop master principals; training on mission, leading change, knowledge of teaching and learning, collaborative relationships, and accountability systems; participants must qualify for each of three phases</td>
<td>Peer, 2012</td>
<td>Tier IV</td>
<td>Logic model based on research; positive evaluation findings, including improved leadership practices, school culture change, and improved achievement (self-reported)</td>
</tr>
</tbody>
</table>
Our review also found a well-specified theory of action based on evaluation in one document (Goldring et al., 2015), and a second document showed positive evaluation results for a targeted effort to support a principal’s time management through the School Administration Manager program (Turnbull, Haslam, et al., 2009).

**Comprehensive School Improvement Models**

Comprehensive school improvement models are multidimensional activities (e.g., changes in curriculum, instruction, staffing, management) focused on improving low-performing schools. Under ESSA, activities falling under this category can be supported by Title I School Improvement Funds, provided there is tiers I through III evidence to support the activities. Federally supported school improvement efforts have, over the past 14 or more years, embraced comprehensive approaches to school improvement. The Comprehensive School Reform Program, Title I under NCLB, and School Improvement Grants represent billions of dollars in funding for schools. Many of the models promoted by these programs involve school leadership components, such as replacing the principal. The School Improvement Grants, for example, required the use of one of four models, all of which directly or indirectly involved leadership change. Certain school
improvement models are also highly centered on school leadership. For this report, we included school improvement models in our review if school leadership was explicitly identified as one of a small number (five or fewer) of core components.

Our review of evidence uncovered two such school reform models—the KIPP model and the School Turnaround Specialist Model—with tier I, II, or III evidence. KIPP is a public charter school network that emphasizes leadership—including leadership autonomy and visionary leadership—at the heart of the model (KIPP, undated). Seven tier I or II studies found substantial and statistically significant improvements in student achievement (see Table 11); KIPP could be supported, under ESSA Title I, as evidence-based. The University of Virginia School Turnaround Specialist Model is a two-year initiative that emphasizes the central role of school leadership in driving school turnaround. One tier II study found positive effects of the program on student achievement after two years that persisted for two years after the intervention (Player and Katz, 2016).

Evidence on Potentially Effective School Leadership Actions
Our literature review also uncovered evidence about school leaders’ effective actions and characteristics. This research is not focused on interventions per se. For this reason, research on effective principal actions and characteristics was not included in our review of the evidence regarding interventions that could be supported under ESSA. However, this research base could help point states and districts toward activities or strategies that have the potential to improve the quality of school principals.

As discussed earlier, tier IV evidence might include studies that connect an intervention to a principal action that was demonstrated in prior research to influence teacher, school, and student outcomes. This section shares that prior research. For example, research identifies conditions that can be influenced by principals and are associated with student success: developing and communicating a vision; establishing a culture of high expectations for students and staff; monitoring and supporting instruction; evaluating teachers; hiring, developing, and retaining school staff; maintaining student discipline; managing the school budget; and engaging with the community (Bryk et al., 2010; Seashore Louis et al., 2010). Several meta-analyses identified leadership actions associated with improved student achievement, including supporting the development and use of curriculum, instruction, and assessments; building a shared culture of achievement; establishing goals and expectations; resourcing strategically; planning, coordinating, and evaluating teaching and curricula; promoting and participating in teacher learning and development; and cultivating an orderly and supportive environment (Copeland and Neeley, 2013; Leithwood et al., 2004; Marzano, Waters, and McNulty, 2005;
### Table 11. Tiers I–IV Evidence on the Effects of School Leadership–Focused School Improvement Models

<table>
<thead>
<tr>
<th>Activity</th>
<th>Study</th>
<th>ESSA Tier</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIPP: A public charter school network; “power to lead” is one of the five pillars, or operating principles (high expectations, choice and commitment, more time, power to lead, focus on results)</td>
<td>Angrist et al., 2012</td>
<td>Tier I</td>
<td>Substantial achievement gains, especially among limited-English-proficiency and special education students, and those with low baseline scores had achievement gains each year</td>
</tr>
<tr>
<td></td>
<td>Gleason et al., 2014</td>
<td>Tier II</td>
<td>Positive, statistically significant impacts on student achievement (math and reading), persisting over four years</td>
</tr>
<tr>
<td></td>
<td>Gallager, Sparrin, and Ross, 2005</td>
<td>Tier II</td>
<td>Positive impacts across for fifth- and sixth-grade English language arts, reading, and math</td>
</tr>
<tr>
<td></td>
<td>Tuttle, Teh, et al., 2010</td>
<td>Tier II</td>
<td>Significant positive impacts on reading and math test scores, when comparing student achievement in 22 KIPP middle schools with students in similar public middle schools</td>
</tr>
<tr>
<td></td>
<td>Woodworth et al., 2008</td>
<td>Tier II</td>
<td>Positive effects on math and English language arts tests in middle school</td>
</tr>
<tr>
<td></td>
<td>Tuttle, Gill, et al., 2013</td>
<td>Tiers I, II</td>
<td>Positive impacts across four academic subjects and in both low-stakes and high-stakes tests</td>
</tr>
<tr>
<td></td>
<td>Tuttle, Gleason, et al., 2015</td>
<td>Tiers I, II</td>
<td>Positive, statistically significant impacts on student achievement, especially in elementary and middle schools</td>
</tr>
<tr>
<td>University of Virginia School Turnaround Specialist Program: Involves a planning year with the district, school leadership–selection support, and executive development for school leaders and turnaround teams in residential programs and on-site coaching</td>
<td>Player and Katz, 2016</td>
<td>Tier II</td>
<td>On average, participating schools experienced statistically significant improvements in student achievement after completing the program</td>
</tr>
</tbody>
</table>

**NOTE:** These models are eligible for support under Title 1 School Improvement Funds.
There is qualitative and quantitative empirical evidence linking specific actions, competencies, or leadership styles of principals to student or other school outcomes. This research finds, for example, that principal training, experience as a teacher in the same school, experience as an assistant principal, and experience as a principal all are related to school proficiency growth (Bowers and White, 2014). Principals’ organizational-management skills relate to student outcomes (Grissom and Loeb, 2011). How principals spend their time might also matter. When principals spend time coaching and evaluating teachers, developing the education program, and focusing on organizational-management activities, school outcomes appear to be better (Grissom, Loeb, and Master, 2013; Horng, Klasik, and Loeb, 2010).

Research has also explored whether a specific combination of skills, knowledge, and characteristics can manifest in an overall style of leadership that is more effective than others (Heck and Hallinger, 2014; Robinson, Lloyd, and Rowe, 2008; James-Ward and Abuyen, 2015). Instructional leadership, which focuses on improving classroom instruction, may be three to four times more effective in improving academic and some engagement outcomes than transformational leadership, which relies primarily on a charismatic leader energizing staff (Leithwood and Jantzi, 2005; Robinson, Lloyd, and Rowe, 2008). Leadership in which staff share leadership roles appears to improve student achievement more than leadership in which the principal alone makes most school-level decisions (Seashore Louis et al., 2010, p. 21).

Although these studies do not constitute tier IV evidence in their own right, because they are not linked to an initiative or intervention, they could be leveraged by state and district officials or educators to support new leadership interventions with a theory of action grounded in this evidence.
CHAPTER 5. RECOMMENDATIONS FOR USING THIS REVIEW

School leadership matters for student and teacher outcomes. Further, activities designed to improve school leadership demonstrate positive impact on student and teacher outcomes. In this report, we have laid out the evidence indicating that school leadership can be a powerful driver of improved education outcomes and summarized the evidence on activities designed to improve the effectiveness of school leaders.

We set this review in the context of the types of school leadership activities supported by ESSA and the types of evidence considered sufficient to invest in leadership-improvement activities. We hope that we have laid out some of the areas in which further discussion and clarification can help states and districts that are implementing the school leadership provisions in ESSA. Further support in understanding the evidence tiers and examining existing evidence may provide states and districts with important tools to better select and implement activities likely to improve school leadership.

Recommendations for Administrators

For administrators at the state and local levels who are considering school leadership as a lever to promote school improvement, our review implies that ESSA provides increased opportunities for states and districts to use federal funding for initiatives that strive to improve the quality of school leaders. We note that, although our review did identify specific school leadership initiatives that can and do meet the evidence standards required by the legislation, we discourage educators from focusing only on the interventions listed in this report. There may be principal improvement interventions that do not appear in this review; perhaps they were too new to have a strong evidence base or their research was not publicly available. States and districts may opt to replicate a branded program for their own contexts, using the evidence from the latter to justify the use of the former. The inclusion of tier IV, which presents a substantial change from previous legislation on evidence requirements, offers administrators the opportunity to receive funding for implementing new and innovative school leadership interventions while building the evidence base for the interventions.

Recommendations for Educators

Educators may wish to seek guidance in adapting evidence-based interventions to local context. In view of outstanding ambiguities with regard to tier IV evidence requirements, we recommend that state and local decisionmakers interpret the definition of tier IV evidence conservatively and support initiatives that have a theory of action grounded solidly in research and that have the potential for more-rigorous validation.

Recommendations for Policymakers

This report aims to provide policymakers and thought leaders at every level food for thought on the current definitions of evidence tiers and ways in which they can be honed to better support the use of evidence for improving school leadership. More specifically, to support practitioners who would benefit from guidance
on evidence for school leadership–improvement activities, we recommend that policymakers address the following:

- Further clarify the types of evidence that qualify for tier IV, such as minimal requirements for a “well-specified logic model” and the quality of the research or evaluation that informs the logic model.
- Share information about school leader behaviors that are associated with positive outcomes, which will better guide the selection or development of interventions that aim to improve these behaviors.
- Provide technical assistance to states to determine the evidence on activities under consideration.

**Looking Ahead: Building the Evidence Base**

Our review identified categories of school leadership–improvement activities that had studies meeting ESSA evidence standards and, importantly, identified evidence at all levels of rigor (tiers I–IV) that showed that interventions can improve principal impact. By doing so, the review establishes that schools and districts have the opportunity to take advantage of ESSA funding to improve school leadership and support student achievement. Taking advantage of these funds depends strongly on a school or district’s particular context. With this in mind, we encourage educators to consider the broader evidence for interventions, rather than focusing only on the interventions listed in this report. Policymakers, administrators, and educators at the state and district levels who implement initiatives based on tier IV evidence will need to engage in ongoing evaluation that will contribute to the evidence base summarized here.
APPENDIX A. STATE EXEMPTIONS TO ESSA EVIDENCE REQUIREMENTS

ESSA allows states to waive evidence requirements in some cases. The cases relevant to school leadership–improvement activities are listed below.

  - . . . participates in evidence-based coursework, to the extent the State (in consultation with local educational agencies in the State) determines that such evidence is reasonably available, that is integrated with the clinical residency experience; . . .
  - . . . new teacher, principal, or other school leader induction and mentoring programs that are, to the extent the State determines that such evidence is reasonably available, evidence-based, and designed to—
  - (xxi) Supporting other activities identified by the State that are, to the extent the State determines that such evidence is reasonably available, evidence based and that meet the purpose of this title.
  - (E) providing high-quality, personalized professional development that is evidence-based, to the extent the State (in consultation with local educational agencies in the State) determines that such evidence is reasonably available, for teachers, instructional leadership teams, principals, or other school leaders, that is focused on improving teaching and student learning and achievement, including supporting efforts to train teachers, principals, or other school leaders to—
  - (P) carrying out other activities that are evidence based, to the extent the State (in consultation with local educational agencies in the State) determines that such evidence is reasonably available, and identified by the local educational agency that meet the purpose of this title.
APPENDIX B. EVIDENCE REQUIREMENTS ACROSS FEDERAL LEGISLATION AND POLICY

Since the early 2000s, federal education policy has promoted the development and use of rigorous impact evidence when selecting school improvement activities, to improve the likelihood that education investments will yield the hoped-for outcomes. Like NCLB, which was the prior version of the Elementary and Secondary Education Act (ESEA), ESSA includes a focus on encouraging education decisionmakers to use evidence to select programs. However, in response to concerns about earlier policies, ESSA’s statutory provisions may practically have the effect of providing more flexibility to states leaders and other policymakers to use more types of evidence and even—when the research base is weak—to waive evidence requirements. Because the requirements in ESSA and NCLB are very different, we conducted a careful comparison of the language in NCLB and ESSA to determine whether and how the evidence requirements for ESSA are more or less prescriptive than those in NCLB (see Table 12). This analysis may help policymakers and educators with a history of using federal resources for school leadership initiatives understand allowable interventions under ESSA vis-à-vis NCLB.

Table 12. NCLB and ESSA Evidence Requirements

<table>
<thead>
<tr>
<th>Evidence Requirements for Title I, Part A</th>
<th>No Child Left Behind</th>
<th>Every Student Succeeds Act</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sec. 1116. Academic Assessment and Local Educational Agency and School Improvement. (b) School Improvement (3) School Plan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“(A) REVISED PLAN.—After the resolution of a review under paragraph (2), each school identified under paragraph (1) for school improvement shall, not later than 3 months after being so identified, develop or revise a school plan, in consultation with parents, school staff, the local educational agency serving the school, and outside experts, for approval by such local educational agency. The school plan shall cover a 2-year period and— “(i) incorporate strategies based on scientifically based research that will strengthen the core academic subjects in the school and address the specific academic issues that caused the school to be identified for school improvement, and may include a strategy for the implementation of a comprehensive school reform model that includes each of the components described in part F”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“A schoolwide program shall include the following components: schoolwide reform strategies that use effective methods and instructional strategies that are based on scientifically based research.”</td>
<td></td>
<td>Title I Sec. 1111: “District must have comprehensive support and development plan for each school identified for school support and improvement, interventions named in plan must include at least one that is evidence based.” I U.S.C. § 1111 (d) p37-40, (d)(1)(B)(i) Same for targeted support and improvement (2)(B)(ii)</td>
</tr>
</tbody>
</table>
Defining Evidence Requirements

Both iterations of ESEA promote the use of evidence in decisionmaking. The term \textit{scientifically based research}, which appears 69 times in NCLB, is defined as having rigorous methods, analysis, data collection, design (experimental or quasi-experimental), and reporting and has been independently reviewed. The term \textit{evidence-based}, which appears 63 times in ESSA, also emphasizes design and implementation (which could be interpreted to include methods and data analysis). Both prioritize experimental studies and, secondarily, quasi-experimental studies as the strongest designs for judging the impact of interventions. There are also differences between the two laws, some trivial and some profound. For example, while ESSA requires statistically significant findings for three of the four evidence tiers, NCLB, as implemented by the Institute
of Education Sciences and the WWC (but not in the letter of the law), prioritizes statistically significant and substantively important outcomes. As another example, NCLB requires independent review, and ESSA does not. At a more basic level, ESSA allows tier III evidence (correlational analysis) and tier IV evidence (a rationale based on prior evidence that the intervention is likely to be effective) that do not meet the minimum requirements of NCLB.

Overall, there are components of NCLB legislation that have more-rigorous demands for the types of research used (e.g., peer review, no theoretical work without empirical investigation) than ESSA does. At the same time, ESSA lays out a priority structure for the types of research evidence that makes clear that the preferred evidence comes from the most-rigorous research designs.

Applying Evidence Requirements to Support Principal Improvement

Both NCLB and ESSA provide multiple opportunities for supporting principal improvement—ESSA more so than NCLB. The evidence requirements for these principal-support activities are not more stringent under ESSA than under NCLB. In fact, in many instances, ESSA provides more flexibility regarding evidence.

Under Title I, both iterations of ESEA provide resources for comprehensive school support and improvement; comprehensive school support and improvement can involve whole-school initiatives that focus on principals. NCLB specifies that all activities, strategies, interventions, or plans identified for the purposes of corrective action, targeted assistance, or school improvement rely on scientifically based evidence. For example, schools implementing school-wide reform strategies had to use “effective methods and instructional strategies that are based on scientifically based research.”

NCLB set the same expectation for schools implementing targeted assistance programs. Title I, Sec. 1116(b)(4)(C), establishes that any and all technical assistance given to schools, by either a local education agency or external provider, must be based on scientifically based research.

In ESSA, the language and expectations are similar. Comprehensive and targeted support and improvement plans for the lowest performing schools must include evidence-based interventions (Title I, Part A, Sec. 1111(d)(1)(B)(ii) and Title I, Part A, Sec. 1111(d)(2)(B)(ii)). There is some ambiguity in the exact requirements. Legal analysis suggests that the use of the term include means that an intervention that has one or more components demonstrated effective by eligible evidence (tiers I through III for Title I School Improvement Funds) would qualify for funding.

Title II may be somewhat more ambiguous regarding evidence requirements. Unlike for Title I school improvement activities, most initiatives funded under Title II can be supported by evidence from any of the four tiers. Several new components in ESSA Title II suggest greater flexibility about whether evidence is required. The 3 percent set aside for principal pipeline activities (ESSA 2015, Title II, Part A, U.S.C. Sec. 2101(c)(4), pp. 118–122) lists 20 potentially fundable types of activities. The last item in the list, “[s]upporting other activities identified by the State that are, to the extent the State determines that such evidence is reasonably available, evidence based,” could suggest but does not require that the preceding items be evidence-based, depending on interpretation of the law.

Finally, the two competitive grants under Title II that can support principal improvement activities give a competitive priority to grant applications based on tiers I–III evidence: “(e) PRIORITY.—In awarding grants under this section, the Secretary shall give priority to an eligible entity that will implement evidence-based activities, defined for the purpose of this subsection as activities meeting the requirements of section 8101(21)(A)(i)” (ESSA 2015, IIA4 U.S.C. Sec. 2242, pp. 147–148).
APPENDIX C. SUMMARY OF UPDATES TO EVIDENCE REPORT IN DECEMBER 2016 RELEASE

We began work on School Leadership Interventions Under the Every Student Succeeds Act: Evidence Review just a few months after ESSA became law. At that time, the language of the law stood alone. To understand the ESSA evidence tiers and apply them to research on school leadership interventions, RAND analyzed the ESSA and NCLB legislative language, as well as existing federal policy tools, such as the WWC standards. Based on the information available in early 2016, we developed an approach for the preliminary evidence review that was published in April 2016. The April 2016 version of our report cautioned, “We anticipate that federal or state departments of education might provide further suggestions or guidance to help apply the evidence tiers, because ESSA is silent on many points that have been important in past Institute of Education Science evidence standards. Our current review of school-leadership-improvement activities casts a broad net to include studies that meet ESSA’s evidence tiers, as specified in the legislation. Findings may be subject to change, depending on guidance or information provided by federal or state departments of education” (Herman, Gates, et al., 2016, p. 9).

To take advantage of new information and additional ESSA analysis, some changes were made to the report between the April 2016 release and the December 2016 updated release. This appendix describes the following changes.

- Evidence Guidance: The December 2016 updated report incorporates guidance provided by the U.S. Department of Education on the ESSA evidence tiers and Title II; incorporating this guidance led to a re-review of most of the studies in the original document and a few changes to study dispositions.
- Level of evidence reviewed: By design, the April 2016 report focuses on tiers I through III evidence. The December 2016 updated report also examines tier IV evidence.
- Additional research identified: Between April and December 2016, we expanded the literature search. Thus, additional studies were added to the review.

Evidence Guidance

On September 16, 2016, the U.S. Department of Education released the Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments (2016c). As anticipated, it provided clarity on many of the questions raised in the April version of the report, addressing all four evidence tiers. Table 13 provides a summary of the main points of the guidance as relates to this review (see last column).

We revised some evidence requirements based on this guidance and then re-reviewed studies that were categorized by any of our reviewers, in any stage of the process, as possible tier IV, as well as a large number of tiers I–III studies. This resulted in reclassification of a small number studies. Our decisions not to include in this report some of the studies discussed in the April report were made based on our analysis of the criteria in the Evidence Guidance and our own analysis of the studies and should not be interpreted as a judgment of the effectiveness of the leadership interventions discussed in those studies. As is the case for any of interventions, lack of strong evidence does not necessarily mean lack of effectiveness. This report assesses studies, not interventions, in accordance with ESSA’s evidence provisions.
Table 13. Clarifying ESSA Evidence Tier Ambiguities

<table>
<thead>
<tr>
<th>Sample size</th>
<th><strong>ESSA Ambiguity</strong></th>
<th><strong>Evidence Guidance</strong></th>
<th><strong>Changes to Report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Does not refer to sample size</td>
<td>Recommends large, multisite sample</td>
<td>No change; large sample sizes would be challenging for a program serving school leaders</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context</th>
<th><strong>ESSA Ambiguity</strong></th>
<th><strong>Evidence Guidance</strong></th>
<th><strong>Changes to Report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Does not indicate whether study context matters</td>
<td>Recommends that the sample and setting in the research overlap with that in the site</td>
<td>No change; information about the context depends on the end user</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flawed studies</th>
<th><strong>ESSA Ambiguity</strong></th>
<th><strong>Evidence Guidance</strong></th>
<th><strong>Changes to Report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Requires that experimental and quasi-experimental studies be well designed and well implemented to qualify as tier I or tier II evidence but is silent on how to consider randomized controlled trials and quasi-experimental studies that have some design or implementation flaws</td>
<td>Includes— but downgrades— somewhat-flawed randomized controlled trials&lt;sup&gt;a&lt;/sup&gt; Does not discuss flawed quasi-experimental designs</td>
<td>Downgraded flawed tier I to tier II and flawed tier II to tier III, if the study met the lower tier’s criteria, rather than rejecting studies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rationale</th>
<th><strong>ESSA Ambiguity</strong></th>
<th><strong>Evidence Guidance</strong></th>
<th><strong>Changes to Report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The lowest category of evidence (tier IV) under ESSA includes “a rationale based on high-quality research findings or positive evaluation” No clear definition of rationale No clear definition of high-quality research findings or positive evaluation</td>
<td>Defines a rationale as a “well-specified logic model” Does not define research findings</td>
<td>Established tier IV standards to require visual logic model consistent with EDGAR definition and required the document to cite evidence supporting logic model</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ongoing evaluation</th>
<th><strong>ESSA Ambiguity</strong></th>
<th><strong>Evidence Guidance</strong></th>
<th><strong>Changes to Report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Does not specify who should conduct the evaluation</td>
<td>Indicates that the evaluator could be an outside implementing site</td>
<td>No change; demonstrating ongoing evaluation is the role of the end user</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relevant outcomes</th>
<th><strong>ESSA Ambiguity</strong></th>
<th><strong>Evidence Guidance</strong></th>
<th><strong>Changes to Report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Does not specify eligible outcomes</td>
<td>Recommends focusing on student outcomes or outcomes associated with program goals</td>
<td>Based on additional clarification from the Department of Education, included student, teacher, and school leader outcomes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Important findings</th>
<th><strong>ESSA Ambiguity</strong></th>
<th><strong>Evidence Guidance</strong></th>
<th><strong>Changes to Report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Does not mention substantively important findings</td>
<td>Does not mention substantively important findings</td>
<td>No change</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Body of evidence</th>
<th><strong>ESSA Ambiguity</strong></th>
<th><strong>Evidence Guidance</strong></th>
<th><strong>Changes to Report</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Focuses on a single positive finding, rather than preponderance of evidence</td>
<td>Recommends that the favorable findings must not be countered by unfavorable findings</td>
<td>No change</td>
</tr>
</tbody>
</table>

<sup>a</sup> Somewhat flawed randomized controlled trials can be considered to meet WWC standards with reservations under some circumstances, putting them on par with well-designed and well-implemented quasi-experimental designs.
Level of Evidence Reviewed
Because of the quick turnaround of the first report, the project’s prioritization of more-rigorous evidence, and better existing methodological guidance for tiers I through III, the April report reviewed only tiers I through III evidence. The April report did offer a strategy for thinking about tier IV and noted the potential for tier IV evidence but did not systematically report findings based on tier IV evidence. With the longer review window and clarification in the Evidence Guidance, it was possible to expand the review to include tier IV, as presented in the December 2016 updated report.

Additional Research Identified
For the April report, RAND librarians searched the following sources for relevant articles:
• Databases:
  – Academic Search Premier
  – SocINDEX
  – ProQuest
  – Dissertation Abstracts
  – EJS E-Journals
  – Campbell Collaboration
  – Scopus
  – ERIC
  – WorldCat
  – Education Research Complete
• Websites:
  – National Association of Elementary School Principals
  – National Association of Secondary School Principals
  – Education Next
  – ASCD
  – Wallace Foundation
  – IES SEED and I3 grantees.

We aimed to conduct a broad search, using terms relevant to the sample (e.g., public education, public school), research (e.g., impact, RCT [randomized controlled trial]), and substantive focus (e.g., principal, administrator)—individually and together.

For the December updated report, we expanded the search to include articles released since April 2016 and to use intervention names, citations, and search terms that emerged in the initial search. Ultimately, we identified and prescreened more than 3,500 articles based on their titles, conducted a two-stage screening of more than 500 articles based on their abstracts, and fully reviewed 128 articles.
Notes

1 ESSA, signed into law on December 10, 2015, is the current iteration of the Elementary and Secondary Education Act.

2 ESSA refers to a state or local education agency or school activity as an “activity, strategy, or intervention” (Public Law No. 114-95, 2015). The What Works Clearinghouse (WWC) and other prior Department of Education documentation use a different general term, interventions, which includes “programs, policies, practices, and products.” We consider these terms to be equivalent. We use the terms activity and intervention interchangeably in this report, although we recognize that they might be interpreted as fundamentally different.

3 The reports cited in this paragraph were not reviewed against the ESSA evidence tiers.


6 “[T]he ESEA [Elementary and Secondary Education Act] considers those [local education agency] staff, such as the principals’ supervisors, who actively mentor and support principals and by doing so are themselves ‘responsible for the school’s daily instructional leadership and managerial operations,’ to also be eligible for Title II, Part A funded support. (ESEA section 8101(44))” (U.S. Department of Education, 2016b, p. 17). The Evidence Guidance also indicates that “[u]nder ESEA sections 2101(c)(4)(B)(vi) and 2103(b)(3)(B), Title II, Part A funds can be used to support those principal supervisors that actively and frequently take responsibility for helping principals with instructional leadership and the school’s managerial operations” (U.S. Department of Education, 2016b, p. 17).


8 Public Law No. 114-95, Every Student Succeeds Act, Title II, Part B, Subpart 1, Sec. 2211, December 10, 2015, p. 130.

9 Public Law No. 114-95, Every Student Succeeds Act, Title 2, Part B, Subpart 2, Sec. 2242, Subgrants to Eligible Entities in Support of Kindergarten Through Grade 12 Literacy, December 10, 2015, pp. 147–148.

10 Public Law No. 114-95, Every Student Succeeds Act, Title 2, Part B, Subpart 4, Sec. 2243, School Leader Recruitment and Support, pp. 148–150.

11 This simplified model does not illustrate the full complexity of the process for improving school leadership. For example, the model does not explore how leadership improvement might differ depending on the type of school (e.g., rural versus urban or high poverty versus low poverty). The simplified model also does not describe the many potential steps between changes in leadership and changes in instruction—such as teacher training on effective instruction, more-challenging curricula, or higher expectations.

12 Under Title II, states have flexibility in applying the evidence-based requirements for some specific ESSA allowable uses of funds, because of the following language: “[An activity] is evidence based, to the extent that the state . . . determines that such evidence is reasonably available.” Of the ESSA programs that relate to school leaders, the following may be exempted from the evidence requirements: coursework for residency programs, new-leader induction and mentoring programs, and any other relevant activities a state selects that are not explicitly listed as allowable uses under Public Law No. 114-95, Every Student Succeeds Act, Title II, Part A, Sec. 2002, Definitions, and Sec. 2101, Formula Grants to States, December 10, 2015.

Although we did not apply the size criterion from the Evidence Guidance, supplemental analysis of the findings shows that if we were to use a criterion of 50 schools or 350 students, only two findings would change.

ESSA focuses on statistically significant findings. The Evidence Guidance suggests that educators also consider the impact of the intervention on participants: “Stakeholders should also consider whether there is evidence that an intervention has substantially improved an important education outcome” (U.S. Department of Education, 2016c, p. 8). Such outcomes are often measured in effect sizes or metrics, such as “years of schooling gained.” The Evidence Guidance does not specify the size or metric for “substantially improved . . . outcome.” The Department of Education has said, “We believe that judgment should be made based on the context” (U.S. Department of Education, 2016d).

The definition of the logic model is from EDGAR, Code of Federal Regulations, Title 34, Subtitle A, Chapter 1, Part 77.1, Definitions That Apply to All Department Programs, as amended December 19, 2014, and is also cited in the Evidence Guidance.

This would include research that finds that the components of the intervention (e.g., weekly mentoring sessions) improve student outcomes (or other relevant outcomes), even in the absence of high-quality evidence proving the effectiveness of the intervention as a whole. In cases where interventions are new or are adapted to context, it may not be possible to have evidence of the impact of the full intervention. Knowing that the components are effective provides an indicator that the intervention overall might be effective—recognizing that this is far from strong evidence of proven impact.

For borderline cases, where there appeared to be a logic model that was not graphically displayed in the reviewed documents, we checked the developers’ website as well.

Public Law No. 114-95, Every Student Succeeds Act, Title VIII, Sec. 8101, Definitions, December 10, 2015, p. 289.

Education Department General Administrative Regulations (EDGAR), Code of Federal Regulations, Title 34, Subtitle A, Chapter 1, Part 77.1, Definitions That Apply to All Department Programs, as amended December 19, 2014.

See the introduction for background regarding the language of “activity, strategy, or intervention.”

Intervention descriptions were summarized from information within the reviewed studies, where available, and supplemented with information from developer websites, where available.


Meta-analyses tend to cast a wide net for studies, including some that have flaws (e.g., no controls for selection bias) and rely on variation across the set of studies to even out differences. If a large set of studies, each of which has some unique flaw, converges on a finding, we can presume that the finding is valid. Additional studies—some more rigorous and some less—point to many of the same leadership actions with varying degrees of emphasis.

Bibliography


Alliance to Reform Education Leadership, AREL State Policy Project, 2013. As of April 11, 2016:
http://www.bushcenter.org/alliance-reform-education-leadership/arel-state-policy-project


http://www.rand.org/pubs/monographs/MG885.html


http://ideas.repec.org/p/nbr/nberwo/17803.html


Braun, Donna, Robert K. Gable, and Stacey L. Kite, Relationship Among Essential Leadership Preparation Practices and Leader, School, and Student Outcomes in K–8 Schools, paper presented at the annual meeting of the Northeastern Educational Research Association, Rocky Hill, Conn., October 2008. As of November 9, 2016:
http://scholarsarchive.jwu.edu/k12_ed/8


http://sde.ok.gov/sde/sites/ok.gov.sde/files/documents/files/NATIONAL%20School%20Leader%20Eval%20Model%2020140204%5B5%5D.pdf


http://www0.gsb.columbia.edu/faculty/jrockoff/cmr_principals_calder_WP38.pdf


Educational Policy Institute, *Focus on Results: An Academic Impact Analysis of the Knowledge in Power Program (KIPP)*, Virginia Beach, Va.: Educational Policy Institute, 2005.


Institute of Education Sciences, “About IES: Connecting Research, Policy and Practice,” web page, undated. As of November 9, 2016:
http://ies.ed.gov/aboutus/


King, C., Quality Measures™ Principal Preparation Program Self-Assessment Toolkit: For Use in Developing, Assessing, and Improving Principal Preparation Programs, Waltham, Mass.: Education Development Center, 2013.

KIPP, “Five Pillars,” web page, undated. As of April 13, 2016:
http://www.kipp.org/our-approach/five-pillars


http://media.wiley.com/assets/61/06/lc_jb_appendix.pdf


http://gradworks.umi.com/34/18/3418101.html


Southern Regional Education Board, *Schools Can’t Wait: Accelerating the Redesign of University Principal Preparation Program*, Atlanta, Ga., 2006.


———, private communication with authors about Evidence Guidance, 2016d.


