Fidelity, Rigor, and Relevance: How SEAs are Approaching the ESSA Evidence Requirements

Lauren Yoshizawa[1] (iD)

Abstract

The Every Student Succeeds Act's evidence requirements mandate the use of research in the selection of school improvement interventions, with the aim of ensuring that schools and districts invest their efforts and funding more wisely. This study of eight states presents three different approaches to the evidence requirements: using lists of pre-sanctioned evidence-based interventions, training schools and districts to evaluate the research supporting potential interventions, and building local evidence of effectiveness. Through interviews with state administrators, I show how each approach relied on different understandings and prioritizations of research rigor and local relevance.

Keywords

evidence, policy adaptation, policy implementation, federal policy, state policies, educational policy

Over the last two decades, education policy has operated under the presumption that schools' improvement efforts will be strengthened by investing in practices and programs with evidence of effectiveness. Yet what counts as

[1]Colby College, Waterville, ME, USA

Corresponding Author:
Lauren Yoshizawa, Colby College, 4423 Mayflower Hill, Waterville, ME 04901, USA.
Email: lauren.yoshizawa@colby.edu

sufficient and meaningful evidence of effectiveness has been under continual negotiation. Federal evidence requirements have prioritized rigorous experimental research while studies show that practitioners often rely on other forms of evidence that are more immediately relevant such as student data, anecdotes, and research syntheses (Birkeland et al., 2005; Coburn & Talbert, 2006; Davidson et al., 2019; Tseng, 2012). The Every Student Succeeds Act (ESSA, 2015) includes the latest iteration of federal evidence requirements. By establishing multiple tiers of acceptable evidence, ESSA still privileges experimental evidence, yet also for the first time invites state and local education agencies (SEAs and LEAs) to use local evaluations and implementation evidence to fill gaps where such research does not yet exist. Therefore, as commentators have noted, this could be an opportunity for SEAs to reconceive the role of research as part of a more nuanced and inclusive sense of what evidence can be (Kane, 2017; Results for America, 2017; West, 2016).

Of course, it is by no means guaranteed that SEAs will implement ESSA's evidence requirements in this way. New policies are interpreted and re-shaped through policymakers' lenses of prior beliefs and practices (Coburn, 2004; Hamann & Lane, 2004; Spillane et al., 2002). Prior implementation studies suggest that SEAs may perceive different opportunities in ESSA and different approaches as more appropriate or commonsense than others, depending, for example, on their pre-existing understandings of research and its utility for decision-making.

Drawing on documentary and interview data from eight purposively selected SEAs, this article addresses the following two research questions:

1. How are SEAs responding to ESSA's evidence requirements?
2. What prior beliefs or understandings do SEA administrators draw on to explain their approaches?

I begin by briefly highlighting the key features of ESSA's evidence requirements and what is at stake in SEAs' implementation. Then I review the literature on different ways educational decision-makers use and interact with research in their decision-making, and develop a conceptual framework of SEA administrators as interpreters and mediators of federal policy. I then present findings on three approaches SEAs have adopted and surface patterns in administrators' perspectives that may help explain each approach. Specifically, I find that the starkest differences between SEA approaches were related to administrators' perceptions of the research base—as relevant or irrelevant to their context, and as outside of or within their control—and of the SEA's proper role in the school improvement process. Finally, I present implications for our understanding of the role of research in supporting school improvement efforts, as well as avenues for future research.

Research Evidence and the ESSA

Evidence requirements rely on the underlying theory that schools' improvement efforts will be more successful if decision-makers are able to identify effective interventions in which to invest their funding and efforts, and that better use of education research can help them to do so. The No Child Left Behind Act (NCLB) attached evidence requirements to several funding programs, pressing decision-makers to look for interventions with positive experimental effects. These were (and remain) few and far between, which undermined the real utility and impact of the requirement (Slavin, 2017; West, 2016). Researchers observed that practitioners sometimes consulted research to inform their decisions, as envisioned by NCLB, but that more often research had limited or indirect influences on decision-making (Coburn, Honig, et al., 2009; Coburn, Touré, et al., 2009; Farley-Ripple, 2012; Weiss et al., 2005).

The ESSA builds on NCLB's evidence requirements with a few key changes; the hope is that, to a degree that NCLB failed to, ESSA will prompt educators' understanding of, and effort to, find research-based solutions for their problems (Penuel, 2015; Slavin, 2017). First, advocates point out that ESSA's new definition of evidence better reflects the reality of available education research, making ESSA's evidence requirement more comprehensible and feasible. ESSA provides a four-tiered definition of evidence, ranging from "strong evidence" based on experimental studies to "promising evidence" from correlational studies. The final, fourth tier opens the door for interventions that do not yet have evidence, as long as decision-makers have a research-informed rationale and a plan for evaluating its implementation.

Second, advocates have also highlighted ESSA's fourth tier as a novel opportunity for SEAs to advance new processes of evaluation and evidence-building. For example, the policy advocacy group Results for America urges SEAs and LEAs to use ESSA as the impetus to "step in and play a more active role in generating evidence" for our shared research base (Results for America & Chiefs for Change, 2018, p. 10). Others agree that if ESSA's fourth tier is implemented well, it opens the potential for practitioners to try new interventions and fill gaps in our knowledge base simultaneously (Penuel & Farrell, 2016; West, 2016).

Finally, and importantly for this study, ESSA devolves more flexibility and responsibility to SEAs, LEAs, and schools. In particular, SEAs have new latitude to define and determine their own approaches to the evidence requirements in their states, which could result in a wide variety of implementations. At their best, the evidence requirements could push practitioners to make better decisions about school improvement. But the autonomy given to SEAs

to interpret and enact the evidence requirements also opens the possibility that they will have little effect on LEAs' decisions, for example, if research is used only to legitimize pre-existing choices, or even shift decision-making for the worse if, for example, the pressure to meet the requirements leads practitioners to try interventions that are a poor fit for their context (Weiss et al., 2008).

Thus, as we watch ESSA's implementation unfold across the states, we should be attentive to how SEAs choose to define compliance with the evidence requirements, how they structure LEAs' research use process, and how they perceive the purpose and potential of the fourth tier of evidence. This study aims to fill these gaps by using an inductive approach to identify patterns in how, and perhaps why, SEAs vary in their understandings of and responses to ESSA's evidence requirements. Its findings have implications not just for ESSA's implementation, but for the underlying theory of requiring the use of research.

Approaches to Research Use

Many studies have endeavored to conceptualize research utilization (Farley-Ripple et al., 2018; Weiss, 1977) and describe how practitioners typically use research (Coburn, Honig, et al., 2009; Farley-Ripple, 2012; Goertz et al., 2013). This literature has coalesced around the following framework of four common types of research use. I use this framework to understand and organize SEAs' responses to the ESSA evidence requirements—that is, how they envision LEAs and schools using research in their decision-making.

First, *instrumental use* is often policymakers' presumptive or intended model of research use. In instrumental use, decision-makers identify a problem and a gap in their knowledge, then seek and interpret research to generate possible solutions, and weigh the pros and cons of each (Coburn, 2010; Coburn, Honig, et al., 2009; Weiss, 1977). A recent survey of school and LEA leaders found that the majority say they use research frequently or all of the time to inform decisions about purchasing programs, adopting curricula, and designing professional development (Penuel, Briggs, et al., 2016). On the other hand, case studies have noted practitioners' disregard for the utility of research, preference for other forms of evidence and other concerns (e.g., stakeholder buy-in, budget constraints), and the challenging time and resource demands for instrumental use (e.g., Birkeland et al., 2005; Nelson et al., 2009; Slavin, 2019).

In *sanctioning* use,[1] the research base is still reflected in decisions, but not because decision-makers have engaged directly with the research or changed their thinking about the problem space. Instead, the process of seeking and

interpreting research is conducted by a separate body, and decision-makers are required to select from the resulting list of sanctioned, research-based options (Coburn, Honig, et al., 2009; Weiss et al., 2005). Multiple NCLB-era initiatives highlight the potential of sanctioning approaches for effecting research-based decisions: studies of the Safe and Drug Free Schools and Reading First programs show that LEAs ended up more often adopting interventions that were on the provided lists and dropping interventions that were not (Herlihy et al., 2009; Weiss et al., 2005). However, one challenge for sanctioning use is determining a credible body and process to develop the lists (Weiss et al., 2008).

In *conceptual* use, research influences decision-makers' understanding of an issue or field, but cannot necessarily be traced linearly to a resulting decision (Coburn, 2010; Coburn, Honig, et al., 2009; Coburn, Touré, et al., 2009; Farley-Ripple, 2012). Through conceptual use, research can meaningfully shift practice (Farrell & Coburn, 2016; Weiss, 1977), but because it tends to be a less conscious, more indirect process, it is difficult to cite or document. In contrast, in *symbolic* use, decision-makers cite research strategically in order to gain support and legitimacy, or reference research generically or selectively to support a decision after it has already been made (Coburn, Touré, et al., 2009; Farley-Ripple, 2012; Penuel, Briggs, et al., 2016). With limited time and resources, LEA and school leaders are likely to engage in symbolic use to satisfy evidence requirements (Coburn, 2010; Coburn, Touré, et al., 2009), but this does not indicate substantive engagement with research nor does it promise to change their existing practices in light of research.

The above framework of four types of practitioners' research *use* is well-established in the literature, but ESSA's fourth tier of evidence also makes relevant a separate literature on practitioners' participation in the *production* of evidence. From action research to lesson study to data-driven decision-making, advocates in this field emphasize the value for practitioners of systematically gathering data on their own practice; reflecting and making sense of evidence, often collaboratively; and formalizing or sharing knowledge of practice (Boudett et al., 2013; Lewis et al., 2004; Lytle & Cochran-Smith, 1992). In particular, the practitioner research field emphasizes the importance of doing research as a way of understanding one's own work (Lytle & Cochran-Smith, 1992), whereas the research use field has traditionally presumed that research is conducted elsewhere by someone else. One exception is research-practice partnerships, which aim to link together both fields through ongoing collaborations (Penuel & Farrell, 2016; Tseng et al., 2017). Both of these areas of work may reasonably be reflected in SEAs' responses to ESSA's evidence requirements, depending on how SEAs interpret and

integrate the fourth tier's permission to produce evidence with the expectations of the first three tiers.

Conceptual Framework

In order to understand how ESSA's evidence requirements will impact local decision-making, it is important to recognize the role SEA implementers play as interpreters and mediators of federal policy. Prior research has found that practitioners sometimes use old definitions to (mis)understand the language used in new reforms (Hill, 2001), and tend to layer reform ideas onto old ones rather than replacing them—even when that sometimes dilutes or contradicts the reform's intention (Cohen & Hill, 2008; Spillane et al., 2002). Over time, through joint work and continual negotiation, implementers tend to develop shared understandings with their colleagues. These context-specific understandings then shape what meaning people make out of new demands (Levinson et al., 2009). For example, implementers often make sense of new policy through deliberation with their peers and/or rely on trusted external experts and resources to inform their understanding (Hill, 2003; Spillane et al., 2002). In their study of two SEAs' implementation of NCLB's comprehensive school reform program, Hamann and Lane (2004) find that administrators' differing interpretations of the program were influenced by their norms about the role of the SEA and already-established problem diagnoses and strategies. Shared understandings about "what worked and what mattered" (p. 448) shaped how each SEA determined the most appropriate, strategic response.

Following this scholarly tradition, I attended to the ways SEA administrators relied on pre-existing, shared understandings to selectively consider and make sense of ESSA's evidence requirements, and to determine the most feasible and appropriate response. In particular, as SEA administrators interpreted the evidence requirements, they drew on prior understandings of research, evidence, and the role of the SEA and LEAs in school improvement efforts.

Understandings of Research and Evidence

SEA administrators made sense of ESSA's evidence requirements by drawing on their prior understandings of what "evidence" is, whether and why research design and ongoing evaluation (as presented in ESSA's tiers) matter, and what it looks like to make decisions "based" on evidence. The term "evidence" is increasingly used in reform discourse, sometimes meaning "data," sometimes "research," and sometimes both. Indeed, prior research indicates

that educational decision-makers use many forms of evidence beyond the research-based evidence described in ESSA, from stakeholder feedback to personal experience, student data, and research—and often use different sources and forms interchangeably (Goertz et al., 2013; Honig & Coburn, 2008; Nelson et al., 2009).

An agency's organizational routines around research shape administrators' understanding of when and how research is useful, and therefore what it looks like to arrive at a research-based decision. Recent research suggests that administrators do not find all forms of research equally useful, and they vary in how they weigh the value of research compared to other forms of evidence (Nelson et al., 2009; Penuel, Farrell, et al., 2016). One common concern for practitioners is finding research that they perceive to be relevant or generalizable to their local context (Birkeland et al., 2005; Nelson et al., 2009). Administrators also vary in the kinds of decisions for which they routinely consult research (Penuel, Briggs, et al., 2016).

SEAs also vary in their available research resources, from the centrality and size of their internal research offices to their networks of internal and external sources of research knowledge (Goertz et al., 2013). Not only do these offices and networks likely shape SEA routines for what research is available and what for, but they also may provide different recommendations and technical knowledge for responding to ESSA (Hill, 2003). The degree to which ESSA's hierarchy of research rigor makes sense and is salient to SEA administrators may therefore vary based on their particular understanding of research design, and their overall perception of the requirements may be informed by their knowledge of what makes research generalizable and relevant for LEAs' work.

Understandings of School Improvement and LEAs

SEA administrators' responses to the evidence requirements were also shaped by their understanding of what SEAs can and should do to support school improvement. One dimension of this is the degree to which states consolidate formal authorities at the SEA or devolve greater authority to districts (Zeehandelaar et al., 2015). States are also steeped in different political cultures around decentralization, which in turn influences the types of policy instruments (e.g., mandates versus capacity-building efforts) they most commonly use to influence school improvement (Louis et al., 2008).

Over time, SEA administrators also develop shared perceptions of their schools and districts, based on their judgments of past interactions and observations, which then affect what future actions they believe are appropriate and likely to succeed. These perceptions include how reliably others' actions

match what they say, how dependably they follow through, the degree to which they are open to and respectful of others' input, and their competence (or lack thereof) to fulfill their responsibilities (Adams & Miskell, 2016; Bryk & Schneider, 2002). Scholars have argued that positive perceptions in these areas support productive role relationships between different groups in the school system, such as teachers, principals, and administrators, but that negative perceptions can undermine progress (Adams & Miskell, 2016; Bryk & Schneider, 2002).

Methods

I conducted exploratory case studies at eight SEAs using a combination of documentary and interview data and an inductive analytic approach.

Setting and Participants

To understand how SEA decision-makers interpreted the evidence requirements, I sought to find administrators who were involved in shaping the SEA's ESSA plans and/or in supporting schools and LEAs in federal school improvement.[2] For the first participant in each SEA, I aimed to include a director- or executive director-level administrator involved in school improvement or federal programs, whose immediate work therefore will be impacted by the evidence requirements. These included individuals who manage and coordinate their SEA's system of school support and lead the implementation of federal school improvement. I also aimed to include an upper-level administrator to provide broader context about the SEA and insight into how the agency's ESSA plans fit into its larger strategic vision. Table 1 briefly describes the participants in each SEA.

Data Collection

I conducted a first round of interviews with 45 individuals at 26 SEAs. These interviews focused on attitudes and behaviors regarding research at the agency, the state's existing processes and priorities in school improvement, participants' understanding and interpretations of research evidence, and their perceptions of districts involved in school improvement and turnaround. These interviews included both survey-style and open-ended questions on most topics, adapting several items from prior studies of research use (Penuel, Briggs, et al., 2016) and trust (Adams & Miskell, 2016). Interviews also included both questions about participants' own work and beliefs and their perceptions of their colleagues at the SEA. I used the data from these initial

Table 1. Participants in Final Sample by SEA, Division, and Title at Time of Interviews (*N* = 19).

Participant role	SEA A	SEA B	SEA C	SEA D	SEA E	SEA F	SEA G	SEA H
Title								
Deputy, Associate, Assistant Superintendent/ Commissioner, or Chief Officer	1	1	0	1	1	2	0	0
Executive director	0	0	2	0	1	0	1	0
Director, manager	2	1	0	1	0	1	2	2
Primary area of work								
Research	1	0	0	0	0	1	1	0
School support, federal programs, and accountability	2	2	2	2	2	2	2	2
ESSA contact for U.S.D.O.E	1	1	1	0	1	1	1	0
Total participants	3	2	2	2	2	3	3	2

Note. In three SEAs, one participant from Round 1 interviews (*N* = 16) had retired or was unavailable for Round 2 (*N* = 15) and was replaced with a new participant. Table 1 reflects participants across both rounds of interviews.

interviews to select a purposive sample of eight SEAs that varied in their perception of their districts, history with local control, resources, and routines for research use.

I then conducted a second round of semi-structured interviews with the eight SEAs in the sample. These interviews focused on the state's vision for the role of evidence-based interventions (EBIs) in school improvement, the state's ESSA plan, and the decision-making process behind the plan. Among other things, participants were asked to explicate pertinent excerpts from the state plan and to rank a list of possible decision-making factors in order of their salience. Most interviews lasted 45 to 60 minutes; all were audio-recorded, transcribed, and cleaned of identifying information such as the names of individuals, districts, and states prior to analysis.

I supplemented my interviews by gathering relevant documents including, but not limited to, each state's ESSA Consolidated State Plan, webinars, websites, grant application templates and guidance materials, and internal decision memos. Throughout spring and summer 2018, I checked for new publicly-available documents and also sent requests to interview participants to share any new materials. As states were at different points of their implementation processes during data collection, the type and volume of documents included in this study (with the exception of the Consolidated State Plans) varied considerably across the states.

Analysis

I first analyzed my data state by state, starting with descriptive and pattern coding (Miles et al., 2014) of SEAs' responses to ESSA's evidence requirements and administrators' beliefs, understandings, and routines related to research and school improvement. Initial descriptive codes included summaries of participants' statements (e.g., *we should continue to do what research supports*), descriptions of participants' ways of talking (e.g., *referring to prior research findings*), and repeated phrases in participants' own words (e.g., *"it is not something we are worried about"*). Pattern coding involved looking across the descriptive codes and consolidating similar or related codes. I then developed case memos about each state, considering the degree to which codes were reflected across multiple participants and data sources.

In my analysis, I iterated between emergent code clusters and the literature to organize the eight SEAs by their implementation approach. I used code clusters to organize the main dimensions of contrast that emerged across SEAs. For example, one pattern that surfaced through comparative analysis was that some administrators tended to defer to ESSA's definition of "evidence-based," while others gave more flexible definitions of "evidence-based" practices. I returned to the literature to see where those codes map onto the existing types of research use. Through this process, I identified three SEAs that shared codes that aligned with instrumental use, while the remaining five shared codes that aligned with sanctioning use. I divided the latter group again when I noted that four SEAs additionally fit a pattern of codes related to evidence-building. I also relied on my conceptual framework to develop code clusters related to SEA administrators' pre-existing understandings about research and school improvement. For example, I grouped together codes reflecting SEAs' different routines for research use, or their different perceptions of LEAs.

Timing

Data collection for this study began in spring 2017 and concluded in fall 2018. During that time, states submitted their consolidated plans to the U.S. Department of Education for approval. All 50 states had their final plans approved by September 2018. For many states, the first cohort of schools for comprehensive support and improvement (with plans developed and reviewed under ESSA's new evidence-based requirements) were identified in fall 2018.

Therefore, the timeframe for this study includes SEAs' preliminary responses to ESSA's evidence requirements that informed the development of their submitted plan and then their early efforts to clarify the actual

processes, tools, and expectations for their upcoming cohorts of schools. Many participants acknowledged that their vision and plans were still tentative, with details yet to be determined and some questions still unanswered. At the time of data collection, some SEAs' materials were still in draft or pilot form, and some had not yet been developed. Therefore, these findings should not be interpreted as representing the final shape of implementation, but rather an early look at how SEAs have differently interpreted the ESSA evidence requirements and envisioned different ways to fulfill them.

Findings

I found that SEAs' implementation approaches differed in how they conceptualized LEAs' research use. That is, the tools and processes SEAs developed and the ways administrators explained them reflected different understandings of what forms of research LEAs should engage with, how LEAs will know what interventions qualify as "evidence-based," and finally how LEAs should incorporate that research and those EBIs into their decision-making. Below, I describe three general approaches SEAs have taken to ESSA's evidence requirements: from providing a simplified and straightforward set of tools for sanctioning use, to training LEAs in the skills for instrumental use, to balancing short- and long-term goals for evidence use and evidence building. For each approach, I also show how administrators relied on different perceptions of research and school improvement to make sense of the requirements.

Sanctioning Use: Focused on Implementation Fidelity

In SEA A, schools and LEAs in need of improvement will submit plans addressing a set of effective practices that have been vetted and organized in an online platform by an external organization, and prioritized by the SEA. This approach therefore resembles sanctioning use. Although many states subscribe to the same online platform, they use it quite differently. SEA A's particular use of the online platform reflected a sanctioning approach in two important ways: first, by interpreting all practices in the platform as evidence-based rather than maintaining ESSA's emphasis on *tiers* of evidence rigor, and second by pre-selecting a mandatory list of practices for LEAs and schools. These steps simplified the EBI selection process, which allowed SEA A to provide more targeted assistance and allowed LEAs to focus on implementation fidelity.

SEA A's sanctioning approach treated inclusion in the online platform as the simple definition of "evidence-based." The platform itself provides varied "strength of evidence ratings" for each of its practices; however, SEA A's 2018 training materials did not direct schools or LEAs to review or weigh these ratings (based on effect sizes) during their planning processes, nor did they mention ESSA's requirements or evidence tiers based on research design. Recognizing that not all practices in the platform may meet ESSA's new tiers, an administrator planned to speak with the platform developers to "make sure that what we have in that system [. . .] will in fact meet that test" (SEA47)—not to train LEAs to distinguish between practices within the platform. Thus it may not be that SEA A administrators did not *themselves* perceive and understand ESSA's new tiered definition of evidence, but that they did not see those tiers as important considerations for *LEAs*. Instead, for LEAs in SEA A, simply completing an improvement plan using the platform's practices will be considered sufficient to fulfill ESSA's evidence requirements. ESSA's newly defined tiers of evidence have potential implications for what *options* schools and LEAs will be offered through the platform, but not for their decision-making *process*.

SEA A further turned the online platform into a sanctioning tool by identifying 12 "key" practices that will be mandatory for all low-performing schools to address. To build the list of key practices, SEA school improvement staff identified strategies they found to be most high-leverage in their professional experience and verified that they matched practices in the online platform—for example, regular communication with parents and using a tiered instructional system. In this way SEA A used the evidence requirements to impose a particular set of SEA-sanctioned, evidence-based practices. The intention here was to provide better support for LEAs. One administrator explained, "In the past we would have gone to a low-performing district and said, 'tell us what we can do to help you.' Now we're flipping that. We're walking in with, 'Here are the evidence-based strategies that we know work. We also know [the SEA] has the support to help you get to where you need to be" (SEA41). In SEA A's 2018 training materials on "Selecting Your School's [Practices]," the first step was to "Select and Assess the 12 Key [Practices]"—in other words, not really a selection at all.

For SEA A, EBI selection is and should be simple, enabling LEAs and schools to focus their attention instead on implementation, with a particular emphasis on fidelity. For example, SEA A's planning guide for schools cautioned, "A school improvement plan may be well-developed, built on relevant data, loaded with evidence-based practices to improve learning [. . .] but the plan is only as good as the fidelity of implementation." Administrators hoped the online platform developers would provide more vignettes and research-based resources on "what it looks like when you're doing it right"

(SEA47). Thus, the research that SEA A administrators wanted LEAs to engage with were models and advice on *how* to do EBIs rather than *which* EBIs to do.

SEA A administrators had long relied on others' research to guide their work and decisions. These traditions supported their positive beliefs that external research could and had identified best practices that could and should be replicated in their LEAs. Yet while the SEA has regularly sought and used research instrumentally, administrators envisioned their LEAs engaging in sanctioning use instead. When SEA A administrators talked about school improvement, the problem they identified was LEAs' ability to follow through and implement interventions with fidelity. For example, one administrator shared that:

> It's not like teachers don't want to do certain things; it's that the school districts don't have the resources to purchase them or make good utilization of them. A part of that also does involve competence of the staff. You know, we hear and we see in some school districts that when they do get extra money, their first inclination is to buy some new program. And they don't have the wherewithal to implement it with fidelity. (SEA34)

By contrast, when the SEA takes over a district, "we [the SEA] go in, we see where the needs are, and we actually put something in place" (SEA41). In other words, administrators perceived that the SEA has the efficacy and "wherewithal" that LEAs sometimes lack. Additionally, SEA A had a tradition of highly centralized decision-making that supported their directness. Like five other SEAs in the sample, SEA A can and has taken over low-performing districts or schools; it also produces an optional list of approved or recommended textbooks. One administrator described how "many" of the initiatives forming SEA A's theory of action for improvement "come from an edict from the court system, come from a piece of legislation [. . .], come from a State Board of Education initiative that tells us this is important, [. . .] or could come from our state superintendent's list of priorities" (SEA41). Therefore, a sanctioning approach appeared to those decision-makers to be a commonsense way to efficiently fulfill the evidence requirements and help focus the SEA's hands-on support on replicating EBIs' promising results.

Instrumental Use: Focused on Rigorous Evaluation and Selection

SEAs B, C, and D shared a vision of helping LEAs to be "informed consumers" (SEA13) of research and thereby to make better decisions about what

interventions to pursue than they have in the past—clearly exemplifying instrumental use. Describing what LEAs and schools should do, administrators across these SEAs emphasized three common components: accessing a wide range of possible EBIs, knowing ESSA's tiers of evidence, and using a thoughtful rather than "box-checking" process to evaluate and select an EBI from those options.

SEAs B, C, and D gathered resources and tools for LEAs and schools, aiming to provide the most expansive set of EBI options with the least onerous search process. For example, administrators described developing "one-stop" websites with a "robust repository of materials and resources" (SEA33), or "streamlining the approach and pointing [schools] to the clearinghouses so it's not up to some beleaguered principal to go figure out where the evidence is" (SEA04). SEA C created a website with a growing list of over a dozen resources such as the What Works Clearinghouse, Evidence for ESSA, Results First, and Results for America. SEA D's Title I, Part A Handbook directed users to a dozen similar sites. Rather than narrow LEAs' attention to fewer, prioritized EBIs, these SEAs focused on providing many options.

Providing extensive options was seen as necessary for LEAs and schools to find an EBI appropriate for their needs. In SEA D's revised school improvement handbook, they advised, "It is important to review a variety of evidence-based practices to find the best strategy to match identified school priorities." Acknowledging the diversity of needs across their state, one SEA B administrator explained, "we are still using lists, but not quite so small or so definitive lists. In our past experience, whenever we put out a list and said, 'these are the five approved things,' [. . .] it turned into true concerns raised by districts about, 'well, but that doesn't work for me'" (SEA04). By contrast, "pointing them to clearinghouses that already have [ESSA-]aligned interventions, [. . .] we just thought that was a more nuanced approach and would help our districts be more able to find things that actually work for their context and their needs" (SEA04). SEA D similarly aimed to prevent the problem of "having LEAs say, 'well what choices do we have?'" (SEA37).

SEAs B, C, and D explicitly instructed LEAs and schools on ESSA's four different tiers of evidence rigor. One SEA B administrator articulated the vision for LEAs as follows: "When it's fully implemented, they'll know what evidence is, [that] there are levels of evidence, [and] where you can find evidence" (SEA04). Thus, implementation is not about *generally* being more evidence-based, but about *specifically* applying ESSA's tiered definition of evidence rigor to selecting interventions. In fact, the very specificity and clarity of ESSA's requirement was what these SEAs found most promising. A SEA C administrator explained that "the way they've laid out levels of evidence and what's expected with those [. . .] I think that's just lifted [evidence]

up beyond just a catchphrase" (SEA13). These SEAs shared an understanding that one of the problems in school improvement had been schools and LEAs making decisions that either were not based in evidence (e.g., "[teachers] are resorting to looking on Pinterest for good ideas or going to Teachers Pay Teachers [. . .], but they're not going about it with intentional purpose" [SEA33]) or were based in anecdotal or unreliable evidence (e.g., "I mean, they didn't quote or cite the What Works Clearinghouse [. . .] mostly what they were quoting was whoever the publisher said the evidence came from" [SEA37]).

Therefore while the *search* for EBIs was meant to be easy, administrators in these SEAs hoped the *evaluation* and *selection* processes would take serious consideration and time—and they produced tools and templates to structure those processes. For example, SEA B hoped to create guidance in the form of questions to help LEAs use the clearinghouses and "negotiate the world of evidence," such as "How do you weigh out different pieces? What if there's two studies in there that say something conflicting?" (SEA04). The aim was that LEAs will "not just [pick] a research-based strategy because, you know, 'these are [Robert] Marzano strategies and they're all good, so we're just going to pick one'; I think it's a more thoughtful, more thorough way of trying to ascertain something that will actually work for them and move the needle" (SEA19). SEA C, which adopted the guides for identifying EBIs developed by the Florida Center for Reading Research, emphasized "walk[ing] through a process where [. . .] it lets you as a group kind of talk through several different interventions, what does the research say, what do the outcomes look like, and arrive at a consensus about what might work for your particular contexts" (SEA13).

Administrators emphasized the importance of deferring to local control in this approach, and tended to describe their work with LEAs as tentative and restrained. For example, SEA D produces a list of recommended textbooks, but administrators noted that "we don't want to micromanage that" (SEA33). Similarly, one SEA C administrator explained that "[we] understand that the district at the end of the day is the one that has to implement it and we need as positive a relationship with them as we can. So for us to go in and mandate XYZ, that doesn't sit well politically in the state, doesn't sit well with superintendents. And so if we can draw them to the well and get them to drink, we stand a whole lot better of a chance at getting the work done" (SEA13).

At the same time, administrators in SEAs B, C, and D also highlighted how external researchers and research-based syntheses and tools were useful in their own work. Although SEAs C and D did not have their own research office or staff, and SEA B was in the middle of "basically offshore[ing] some of that [research] work that we've been trying to do in-house" (SEA04), they

all were able to point to multiple external research partners or organizations whom they were able to access easily and often. Most often, the kind of help they wanted and the questions they asked took the form of research-based training, tools, and literature reviews—all products that could quickly and directly be applied to their work. For example, one administrator in SEA C spoke about "relying heavily" on research clearinghouses, using the comprehensive center "pretty extensively," and "reach[ing] out" to other external organizations "to pull from other states what might be working in other places or what's going on" (SEA13). A SEA D administrator recounted how, when they wanted to revise their improvement system, they reached out to their comprehensive center who "actually conducted a lit review on our behalf [. . .] and then we took that information and developed our school improvement needs assessment" (SEA33). And "then we had that research base that [the comprehensive center] had pulled together for us to support what we were proposing" (SEA33).

In their talk about research, these administrators shared a perception that knowledge originates out in the research community and then is transferred into the SEA. Therefore with decisions firmly made at the *local* level and the evidence to inform better decisions primarily developed *externally* by research and intermediary organizations, the appropriate SEA role was neither to make decisions nor to create evidence, but instead to make a process that connects local decision-making with external research.

Sanctioning Use and Evidence-Building: Focused on Local Relevance

SEAs E, F, G, and H embraced the promise of evidence use more than ESSA's specific evidence requirements. Concerned that holding tightly to ESSA's top three tiers of evidence could force LEAs away from good, context- and culture-specific practices, their implementation approaches aimed to comply with ESSA's requirements evidence without allowing that compliance to compromise their broader vision of LEA evidence use. In order to satisfy ESSA's evidence requirements in the short term, these SEAs largely adopted elements of a sanctioning approach, but, unlike SEA A, were enacting sanctioning use in ways that intentionally created space for local judgment around specific action steps. Meanwhile, their long-term ambition was to cultivate evidence-based habits of mind and evidence-building capacities in both the SEAs and LEAs in order to support continuous improvement and to generate new evidence on their locally-relevant practices—embracing ESSA's flexible fourth tier of evidence. This emphasis on local evidence-*building* as a valuable and

necessary supplement to doing EBIs was what most clearly distinguished these four SEAs from the others.

Short-Term Sanctioning Use

SEAs E, F, and H identified a set of practices that they were confident could meet the evidence requirements, but were quick to add that those practices were intentionally broad, leaving room for LEAs to choose more specific, context-appropriate action steps—even if those steps were not necessarily evidence-based.[3] In all three cases, SEA administrators felt it was important to ensure that their framework of sanctioned EBIs satisfied ESSA's high expectations. For example, SEA E sent their school improvement strategies to a policy research organization to analyze and "give us feedback on what is statistically sound or evidence-based" (SEA23). SEA F administrators sought to "understan[d] the evidence requirements well enough to know what of our existing research already met them [. . .] and we were able to demonstrate convincingly that we were already doing the expectation" (SEA20). But once the evidence requirement is fulfilled at this general practice level, administrators wanted to give LEAs more flexibility on the specifics:

> We are promoting our evidence-based practices and then the LEA [. . .] is making decisions on how to use their funds to implement those practices in a way that perhaps has not been proven to be successful, just because there is a dearth of research on it. (SEA25)

Similarly, a SEA F administrator explained, "if a school [. . .] wanted to use some money for family engagement strategies to have a monthly potluck [. . .] that might not be research-based—it is going to be really hard to find studies that say potlucks lead to school turnaround—but we are not going to say to that school, 'no, you can't do that,' if it is a solid component of their family engagement strategy" (SEA46).

Long-Term Evidence Building

SEAs E, F, G, and H were investing heavily in the theory that schools and LEAs will improve most when they are encouraged to build evidence on their implementation of local practices—both to inform their own continuous improvement processes and to establish a more relevant and rigorous knowledge base on what works. That is, administrators saw an intrinsic value in having schools and LEAs go through a process of gathering evidence on their practice, as well as an instrumental value in filling perceived gaps in the

research base. Therefore, SEAs E, F, G, and H put their energy and resources into supporting action research networks, local pilots, and other evidence-building strategies using ESSA's fourth tier of evidence.

SEA G's plan was to gradually build an ESSA-aligned, state-specific EBI clearinghouse that includes local and SEA research. One administrator described the SEA's vision thus: "we are putting some processes in place to evaluate those things with our research team so that when we do not have a level of evidence, we can generate one" (SEA39). In order to achieve this, SEA G planned to facilitate school- and LEA-level action research networks and rapid evaluations "all with an eye toward getting those strategies into that Level 3, if you will, Level 2, evidence strategies" (SEA48). Their clearinghouse website told practitioners to "buil[d] up the evidence base," because "when you evaluate the strategies being used in your districts and schools, you are contributing to the broader understanding of how evidence-based strategies work."

Other SEA administrators emphasized the process of developing LEAs' positive habits and dispositions for evidence-building over ESSA's specific evidence requirements. For example, SEA F produced guidance materials on how to conduct a program evaluation, but administrators acknowledged "not every study they do or every analysis they do is going to meet the ESSA evidence requirements, but it is going to be better than nothing" (SEA20). Starting with evidence-building was purposeful for SEA F. As one administrator explained, "researching our own work [is] what gets people interested [. . .] you start there, then you build the appetite for, 'oh, well, you know, our research is showing that this isn't terribly effective—is there other research that shows that a different strategy might work better?'" (SEA20). In other words, SEA F was hoping that making research relevant first would build the dispositions and curiosity necessary to seek and use rigorous research later. Similarly, SEA H planned to "require that LEAs are starting to gather the evidence that they need to establish the effectiveness of their interventions," but its priority was "that we are carving out a space where we can be responsive to local expertise and knowledge of need, even if that local expertise has not yet developed the capacity to document an evidence base as required most rigorously by ESSA" (SEA25). The logic was to eventually build a rigorous evidence base that will codify local expertise and knowledge, beginning first with just the habit of gathering evidence, and developing the capacity for rigor over time.

Both in creating flexibility for LEAs to translate the SEA's EBIs to their own contexts and in encouraging evidence-building, administrators in SEAs E, F, G, and H emphasized not only local control but the value of local expertise. As one administrator in SEA G explained, prior state-centric approaches

had ill-suited their large, diverse state, "because everyone has a different context, different need, different human capital structure, etc., so we really have to back away from that" (SEA38). A SEA H administrator noted, "we have limited capacity here at the state, and so I am not certain that the local folks at the LEA level are not the most knowledgeable experts on making those decisions; I would probably just as a default defer to them" (SEA25). An administrator in SEA F similarly shared an "understanding that the folks that are leading the efforts at the local district level really have a better understanding than the state does about what their current needs are" (SEA12). And an administrator in SEA F explained that this means the SEA's role is primarily about "enabling schools to get the good information that they need and helping them with critical friend support" (SEA06).

On the other hand, these administrators shared a certain skepticism of the utility and relevance of external research. They raised more caveats about when and to what degree external research was applicable to their decisions, often pointing to instances when research failed them. For example, an administrator in SEA F explained that the external research base trails behind their needs:

> We do like to use as good research as we can [. . .] but sometimes we are so much on the edge of change, things that have not been done before. So we have to, with our eyes wide open, accept that we may not have perfect data to move forward. [But] for us to wait until there are these research studies that are well-designed and all of that would leave a lot of schools behind. (SEA06)

Other administrators saw the research base as rarely applicable to their contexts. For example, an administrator in SEA E recounted their experience with the research-backed School Improvement Grant models, where "a lot of those turnaround principles were really about big city schools and it is not practical in a rural state like [ours . . .] so we got a lot of pushback [. . .] about 'they don't know the environment or the culture that we are in'" (SEA23). A SEA H administrator shared a similar story with Success For All: "if a school or a district [here] adopts that program, by the nature of their size, they have to fine-tune it and implement it a little bit different than it's been implemented back in Baltimore, right? And so, that is, a concern [. . .] because these small schools, multi-level classrooms are just a different beast that has often not been researched explicitly" (SEA25).

By contrast, administrators spoke readily of the benefits of generating and using their *own* research in the SEA to support continuous improvement on locally-appropriate or innovative strategies. In SEA E, the research office was small and their processes were ad hoc, yet its administrators described

Table 2. Summary of SEAs' EBI Implementation Approaches.

Implementation approach	What research do LEAs engage with?	What does "evidence-based" mean?	How do LEAs decide on an EBI?
Sanctioning	Implementation-oriented, research-based tools	Vetted by external organization	Choose from SEA's list
Instrumental	Research studies and clearinghouses	Meeting ESSA's evidence tiers of rigor	Follow guided procedure
Sanctioning and evidence-building	Conduct their own research/ evaluations	Aligned with research, corroborated with local evidence	Use good judgment regarding local appropriateness

regularly getting help with research-based answers to support the state's decision-making. They recounted times their research office developed literature reviews and "just things for the Board to look at when they were trying to transform how they wanted to think about professional learning" (SEA23), and conducted a mixed methods study of how LEAs were implementing new standards (SEA12). Every year, staff from SEA F's research office "go to the program staff and [. . .] say, essentially, 'what is it we need to learn this year to help you do your job better?'" (SEA20). By conducting their own research, SEA F staff were able to ask different questions for which they did not see answers externally in existing research.

All together, administrators in SEAs E, F, G, and H tended to describe their LEAs as doing good work but the research community as lagging and lacking relevant evidence. Thus a loose sanctioning approach combined with evidence-building made sense as the way to help LEAs reap the most benefits of research-based decision-making.

Discussion

Summary of Main Findings

How SEAs implement ESSA's new evidence requirements will reshape the policy context within which LEAs use research (or not) to make consequential decisions about school improvement. In this exploratory study, I found that SEAs held different conceptions of what it means to fulfill these requirements (summarized in Table 2). In SEA A, administrators adopted a straightforward understanding of EBIs: a trusted source has determined what

practices are evidence-based, and the SEA will require that LEAs implement a pre-selected list of practices. SEAs B, C, and D honed in instead on ESSA's multi-tiered definition of "evidence-based" and will train LEAs to find, evaluate, and select EBIs accordingly. For these SEAs, interventions were no longer simply evidence-based or not evidence-based, but were supported by bodies of research with different levels of rigor, and it will be up to LEAs to weigh those strengths in a structured process of instrumental use. Finally, SEAs E, F, G, and H conveyed perhaps the most nuanced sense of EBIs. Though strongly invested in the *spirit* of evidence-based decision-making, administrators in these SEAs were more skeptical of the *letter* of ESSA's requirements based on what research is available and what LEAs are ready for. Therefore, they saw a sanctioning approach as a way to technically comply with the requirements while building capacity and original evidence.

This study also noted the ways that SEA administrators relied on different understandings and routines regarding research and their relationships with LEAs, which shaped how they justified the SEA's approaches as appropriate and feasible. First, I found that SEA administrators were accustomed to different patterns of state- or local-centric decision-making in school improvement. SEA A administrators were more likely to describe the SEA as the primary decision-maker and direction-setter. This orientation informed their sanctioning approach: LEAs will not directly engage with research to select EBIs because those selections will be made for them. The other SEAs' administrators instead perceived their LEAs as the primary leaders and decision-makers for school improvement. Second, I noted that administrators in SEAs B, C, and D described external research as providing the expertise LEAs need for their decisions, whereas administrators in SEAs E, F, G, and H cited meaningful lessons from their SEA's internal research and contrasted those with the limited relevance of external research. These different understandings justified an instrumental approach for some SEAs and an evidence-building approach for others.

Implications

This study's findings should make us question the underlying theory of evidence requirements. Evidence requirements presume that the education research base contains information that is relevant and useful to the problems and decisions involved in school improvement. ESSA's evidence tiers steer practitioners toward more rigorous studies with experimental designs, reflecting the belief that greater research rigor should give practitioners greater confidence in the findings. The administrators in SEAs A, B, C, and D seemed to agree: research quality matters, is more efficiently achieved by external,

full-time researchers, and external research findings can be immediately applied. Yet half of the administrators in the study indicated the *opposite*— that they would put the greatest confidence in locally-generated evidence, even when that involved temporarily putting aside ESSA's expectations for rigorous design. There were many intertwined reasons for this perception: administrators vouched for the local judgment and expertise of their district and school leaders; they perceived the current research base to have limited generalizability for particular demographics and contexts; and for some administrators, internal research routines had become part of the SEA's own learning process. What is interesting is that, when asked directly, the SEA administrators in this study uniformly agreed or strongly agreed that research is trustworthy and can help make better decisions about school improvement. The variable understandings of research, researchers, and the relationship between relevance and rigor are worth exploring further.

Another question for the theory of evidence requirements arises from the way SEA A adapted the mandate to their existing understandings about school improvement and the role of the SEA. Administrators in SEA A saw improvement initiatives and best practices being determined either by state actors or by organizations creating research-based tools; districts were not seen as key decision-makers and therefore the evidence requirements were seen to have little relevance to their work. Weiss (1980) noted that officials struggled to describe instances of research use in part because they struggled to identify as decision-*makers*. With decisions coming from above them, made with many others' input, and/or made incrementally, they rarely saw themselves owning a specific decision or consciously using research for it; rather, research concepts diffused into their working knowledge in less perceptible ways. Indeed, later research corroborates her sense that conceptual use is quite common, if under-acknowledged (Coburn, 2010; Farrell & Coburn, 2016). SEA administrators adapt policy mandates to fit the way work is done in their context; if the evidence requirements are a poor fit for the way school improvement decisions are actually made, then, as in SEA A, they may have little influence.

Finally, the approach in SEAs E, F, G, and H raises the question whether engaging in research use *most* substantively requires, at least for a while, *not* doing the evidence requirements. Their combination of short-term sanctioning use and long-term evidence building is novel and promising in the way it positions practitioners as both consumers and producers of research. The SEAs in this group created sanctioning lists that more closely resembled the kinds of research-based frameworks and broad priorities that characterize conceptual use (e.g., a culture of high expectations, instructional leadership), than a menu of proven-effective programs. On one hand, taking a

more conceptual approach and pushing practitioners to use a research-based framework might lead to more substantive changes in thinking. But on the other hand, by considering ESSA's requirements satisfied at this broad level, and allowing practitioners to decide their specific actions without regard for their evidence base, these SEAs are side-stepping ESSA's intended use. Similarly, by taking up ESSA's fourth tier, these SEAs reflect commentators' high hopes, attempting to generate new evidence on context-specific practices, and help practitioners learn from and be persuaded by evidence on their own and peers' practice (Kane, 2017; West, 2016). At the same time, to engage in evidence-building, the SEAs have also temporarily put aside ESSA's expectations for rigor. As one SEA F administrator put it, they are "less interested" in meeting ESSA's evidence requirements than in "our aspiration is that the whole state gets better at using evidence" (SEA20). These SEAs suggest that *requirement* of research use may get in the way of effective, substantive research use.

Limitations and Directions for Future Research

There are three major limitations to what this exploratory study can tell us. First, the sample of SEAs is small and not representative. It should not be assumed that the three approaches are an exhaustive representation of all approaches across the country, nor that their frequencies are reflective of trends in the broader population of states. The small sample of decision-makers in the study also do not necessarily represent the understandings and opinions of their SEAs as a whole. This study focused specifically on school improvement work; however, other SEA administrators may be informed by different perceptions and take different approaches to evidence use in other domains of work. Finally, this study's implications are limited by its timeframe. In fall 2018, when data collection concluded, many SEAs were still very early in their implementation, and some of the findings are based only on administrators' hopes and projections. Therefore, additional research could address whether and why SEAs' approaches change over time as they transition from idea to reality.

In addition, as states begin to implement their approaches, future research should explore how each approach plays out at the LEA level. We will need to look at LEAs' experiences to see whether and how these approaches differ in learnability and feasibility, and whether and how they shape the decisions LEAs make. For example, SEA E's approach banks on the theory that building evidence on local practice will foster LEAs' curiosity about external research evidence—but how does that manifest in practice? SEAs B, C, and D plan to train LEAs about ESSA's tiers of evidence and how to evaluate

research accordingly—how will LEAs fit these new procedures into their existing decision-making practices?

Over the past decade, researchers and practitioners alike have innovated and studied different ways of using research to inform school improvement decisions. This study suggests that ESSA's evidence requirements, as divergently implemented across the states, pose a new and important opportunity to see whether and how local decision-making practice responds to different visions of research use.

Declaration of Conflicting Interests

Funding

ORCID iD

Lauren Yoshizawa https://orcid.org/0000-0002-4407-3696

Notes

1. Weiss first named this phenomenon "imposed use" (Birkeland et al., 2005; Weiss et al., 2005, 2008) to describe the growing policy trend of evidence requirements. However, given that the "imposition" or requirement of research use could be fulfilled in multiple ways, I adopt Coburn, Honig, et al.'s (2009) term "sanctioning" to describe the role that research plays in the decision-making process.
2. ESSA includes evidence requirements outside of school improvement, such as in the use of Title II and Title IV funds for professional development and safe and healthy school programs. However, these other uses are "allowable" rather than "required," and states' draft ESSA plans suggested that they were focused on responding to the requirements in school improvement.
3. SEA D's evidence-based practices are similarly broad, but this particular intention was not explicitly articulated by SEA D administrators or documents.

References

Adams, C. M., & Miskell, R. C. (2016). Teacher trust in district administration. *Educational Administration Quarterly*, *52*(4), 675–706.

Birkeland, S., Murphy-Graham, E., & Weiss, C. H. (2005). Good reasons for ignoring good evaluation: The case of the drug abuse resistance education (DARE) program. *Evaluation and Program Planning*, *28*(3), 247–256.

Boudett, K. P., City, E. A., & Murnane, R. J. (2013). *Data wise: A step-by-step guide to using assessment results to improve teaching and learning*: ERIC.

Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. Russell Sage Foundation.

Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, *77*(3), 211–244.

Coburn, C. E. (2010). The partnership for district change: Challenges of evidence use in a major urban district. In C. E. Coburn & M. K. Stein (Eds.), *Research and practice in education* (pp. 167–182). Rowman & Littlefield.

Coburn, C. E., Honig, M. I., & Stein, M. K. (2009). What's the evidence on districts' use of evidence. In J. D. Bransford, D. J. Stipek, N. J. Vye, L. M. Gomez, & D. Lam (Eds.), *The role of research in educational improvement* (pp. 67–87). Harvard Education Press.

Coburn, C. E., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education*, *112*(4), 469–495.

Coburn, C. E., Touré, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *Teachers College Record*, *111*(4), 1115–1161.

Cohen, D. K., & Hill, H. C. (2008). *Learning policy: When state education reform works*. Yale University Press.

Davidson, K. L., Farrell, C. C., & Penuel, W. R. (2019). *What types of research do education leaders find useful, and for what purposes* [Paper presented]? Annual Conference of the Society for Research on Educational Effectiveness, Washington, DC.

Every Student Succeeds Act, Pub. L. No. 114-95 (2015).

Farley-Ripple, E. (2012). Research use in school district central office decision making: A case study. *Educational Management Administration & Leadership*, *40*(6), 786–806.

Farley-Ripple, E., May, H., Karpyn, A., Tilley, K., & McDonough, K. (2018). Rethinking connections between research and practice in education: A conceptual framework. *Educational Researcher*, *47*(4), 235–245.

Farrell, C. C., & Coburn, C. E. (2016). What is the conceptual use of research, and why is it important? William T. Grant Foundation. http://wtgrantfoundation.org/conceptual-use-research-important

Goertz, M. E., Barnes, C., Massell, D., Fink, R., & Francis, A. T. (2013). *State education agencies' acquisition and use of research knowledge in school improvement strategies* (CPRE Research Report# RR-77). Consortium for Policy Research in Education.

Hamann, E. T., & Lane, B. (2004). The roles of state departments of education as policy intermediaries: Two cases. *Educational Policy*, *18*(3), 426–455.

Herlihy, C., Kemple, J., Bloom, H., Zhu, P., & Berlin, G. (2009). Understanding reading first: What we know, what we don't, and what's next. *MDRC Policy Brief*.

Hill, H. C. (2001). Policy is not enough: Language and the interpretation of state standards. *American Educational Research Journal*, *38*(2), 289–318.

Hill, H. C. (2003). Understanding implementation: Street-level bureaucrats' resources for reform. *Journal of Public Administration Research and Theory*, *13*(3), 265–282.

Honig, M. I., & Coburn, C. E. (2008). Evidence-based decision making in school district central offices: Toward a policy and research agenda. *Educational Policy*, *22*(4), 578–608.

Kane, T. J. (2017). Making evidence locally. *Education Next*, *17*(2), 52–58.

Levinson, B. A. U., Sutton, M., & Winstead, T. (2009). Education policy as a practice of power: Theoretical tools, ethnographic methods, democratic options. *Educational Policy*, *23*(6), 767–795.

Lewis, C., Perry, R., & Hurd, J. (2004). A deeper look at lesson study. *Educational Leadership*, *61*(5), 18–22.

Louis, K. S., Thomas, E., Gordon, M. F., & Febey, K. S. (2008). State leadership for school improvement: An analysis of three states. *Educational Administration Quarterly*, *44*(4), 562–592.

Lytle, S., & Cochran-Smith, M. (1992). Teacher research as a way of knowing. *Harvard Educational Review*, *62*(4), 447–475.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.

Nelson, S. R., Leffler, J. C., & Hansen, B. A. (2009). *Toward a research agenda for understanding and improving the use of research evidence*. Northwest Regional Educational Laboratory. http://www.nwrel.org/researchuse/report.pdf

Penuel, W. R. (2015). *What will research use look like under the Every Student Succeeds Act?* National Center for Research in Policy and Practice. http://ncrpp.org/blog/2015/what-will-research-use-look-like-under-the-every-student-succeeds-act

Penuel, W. R., Briggs, D. C., Davidson, K. L., Herlihy, C., Sherer, D., Hill, H. C., Farrell, C. C., & Allen, A.-R. (2016). *Findings from a national study on research use among school and district leaders*. National Center for Research in Policy and Practice.

Penuel, W. R., & Farrell, C. C. (2016). Research-practice partnerships and ESSA: A learning agenda for the coming decade. In Quintero, E. (Ed.), *Teaching in context: The social side of education reform* (pp. 181–200). Harvard Education Press.

Penuel, W. R., Farrell, C. C., Allen, A.-R., Toyama, Y., & Coburn, C. E. (2016). What research district leaders find useful. *Educational Policy*, *32*(4), 540–568.

Results for America. (2017). *Leverage points: Thirteen opportunities for state education agencies to use their ESSA state plans to build and use evidence to improve student outcomes*. Author. https://results4america.org/tools/leverage-points-thirteen-opportunities-state-education-agencies-use-essa-state-plans-build-use-evidence-improve-student-outcomes/

Results for America, & Chiefs for Change. (2018). *Evidence-building opportunities under ESSA: How states can and should generate evidence to drive better outcomes for students*. Author https://results4america.org/tools/evidence-building-opportunities-essa-states-can-generate-evidence-drive-better-outcomes-students/

Slavin, R. E. (2017, December 6). Evidence and the ESSA. *Huffington Post*. https://www.huffpost.com/entry/evidence-and-the-essa_b_8750480?

Slavin, R. E. (2019, August 29). Superman and statistics [Blog post]. https://robert-slavinsblog.wordpress.com/2019/08/29/superman-and-statistics/

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational research*, *72*(3), 387–431.

Tseng, V. (2012). The uses of research in policy and practice. *Social Policy Report*, *26*(2), 3–16.

Tseng, V., Easton, J. Q., & Supplee, L. H. (2017). Research-practice partnerships: Building two-way streets of engagement. *Social Policy Report*, *30*(4), 3–15.

Weiss, C. H. (1977). Research for policy's sake: The enlightenment function of social research. *Policy Analysis*, *3*(4), 531–545.

Weiss, C. H. (1980). Knowledge creep and decision accretion. *Science Communication*, *1*(3), 381–404.

Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence: How evaluations affect DARE. *American Journal of Evaluation*, *26*(1), 12–30.

Weiss, C. H., Murphy-Graham, E., Petrosino, A., & Gandhi, A. G. (2008). The fairy godmother—and her warts: Making the dream of evidence-based policy come true. *American Journal of Evaluation*, *29*(1), 29–47.

West, M. R. (2016). *From evidence-based programs to an evidence-based system: Opportunities under the Every Student Succeeds Act*. The Brookings Institution. https://www.brookings.edu/research/from-evidence-based-programs-to-an-evidence-based-system-opportunities-under-the-every-student-succeeds-act/

Zeehandelaar, D., Griffith, D., Smith, J., Thier, M., Anderson, R., Pitts, C., & Gasparian, H. (2015). *Schools of thought: A taxonomy of American education governance*. Thomas B. Fordham Institute.

Author Biography

Lauren Yoshizawa is an assistant professor of education at Colby College. Her research focuses on policy implementation and understanding when and how educational organizations change.