"What Counts" as Research? Comparing Policy Guidelines to the Evidence Education Leaders Report as Useful

Caitlin C. Farrell
William R. Penuel iD
Kristen Davidson
University of Colorado Boulder

Despite calls for evidence-based decision making, the field has a limited understanding of how educational leaders actually engage research. This study draws on a nationally representative sample of 368 district and school leaders who named pieces of research that were useful to their work. Educational leaders found frameworks and practical guidance in the form of books to be most useful. They report turning to research across different domains of leadership practice, including supporting their own professional learning, guiding instructional activities for others, and monitoring and supporting implementation. While a small portion of sources named would qualify for the top three "tiers of evidence" of the Every Student Succeeds Act, those sources named as useful for program selection more frequently met these criteria. Together, these findings offer a broader portrait of research use, one rooted in leaders' engagement with research as a part of their multifaceted and complex practice.

Keywords: *evidence use, research use, Every Student Succeeds Act, educational leaders, school leaders, district leaders, leadership practice*

FEDERAL and state policies in education reflect a commitment to the idea that research evidence has an important role to play in supporting educational improvement and transformation. The current reauthorization of the federal Elementary and Secondary Education Act (ESEA), called the Every Student Succeeds Act (ESSA), explicitly mandates that state, district, and school leaders use research evidence when they select interventions supported by federal funds. For the first time in ESEA history, ESSA defines "tiers" of evidence, based on the research methods used, that "count" when choosing an evidence-based program or intervention.[1] Policies such as ESSA focus squarely on one kind of decision in which education leaders engage: selecting among programs and approaches that have been developed and studied elsewhere. Such policies infer that the most useful research for education leaders will be individual studies that offer causal evidence of program impact using experimental or quasi-experimental study designs (Gordon & Conaway, 2020; Haskins & Margolis, 2015).

However, this view represents only a narrow slice of when and why leaders engage with research in their work (Mills et al., 2020; National Research Council, 2012; Penuel et al., 2018). It does not account for the wide range of activities in education leaders' practice, such as supporting everyday instruction, developing curriculum, engaging in and providing professional learning, cultivating school climates and communities, making systems-level decisions, and more (Grissom et al., 2021; Huguet et al., 2019; Rorrer et al.,

2008). It therefore misses the broader range of research that education leaders find useful to their work, beyond those that measure the effectiveness of programs.

There is a critical need to systematically examine the predominant premises related to evidence use in education policies. Rather than assuming that education leaders' decisions are primarily related to selecting programs, we need to better understand the broad range of purposes for which research is useful for leaders as a part of their regular work practice. Rather than advance the idea that the most useful research studies for leaders are impact and effectiveness studies, we need to learn more about what types of research educators *actually* find useful in their work. Calls for "evidence-based policy making" are likely to fall short if we do not understand the kinds of research that leaders readily use in the context of their complex daily work.

To offer a systematic look at leaders' use of evidence in their decision making, this article presents results from a survey of research use administered to a nationally representative sample of school and district leaders from mid- and large-size school districts across the United States (Penuel et al., 2016). We focus on school and district leaders because it is at the local level that educational improvement and transformation efforts take root (Gamson & Hodge, 2016; Spillane et al., 2018). Using open-ended survey responses, we identified the sources and types of evidence that leaders specifically named as useful to their work across their practice. We found that, rather than individual

effectiveness studies, leaders most often pointed to syntheses of research that offered frameworks to apply more broadly across various aspects of their work. By comparing the research methods represented in the sources named by leaders with those specified in the ESSA tiers of evidence, we found that most leaders turned to sources of evidence that met an ESSA tier when selecting an intervention. However, leaders much more often named sources that were useful to other aspects of leadership practice. These sources often did not meet the ESSA tiers but instead drew from a range of research methods to support a broad evidence base for a given practice.

We argue that if policies are to be responsive to—and supportive of—the practice of district and school leaders, then they need to go beyond advancing evidence use for decision making narrowly defined to providing evidence that supports the full complement of leaders' work (Ming & Goldenberg, 2021). Beginning with this richer rendering of leadership practice as a policy premise suggests establishing practice-centered policy regimes that support the use of more plural forms of evidence and a practice-forward understanding of educational improvement.

The Need to Focus on What Education Leaders Find Useful in Their Practice

Over the past two decades, connecting research to educational improvement has been a key aim of several federal policies and initiatives. The No Child Left Behind Act (2002), the Education Sciences Reform Act (2002), and education initiatives funded under the American Recovery and Reinvestment Act (2009) all included several provisions related to research (Tseng & Coburn, 2019). Many have argued that research evidence is likely to yield the best information to help educators decide how to allocate scarce resources to select programs or interventions for improvement efforts (e.g., Whitehurst, 2003). All include similar definitions of the type of research that should be useful to educational decision makers, namely, research that examines the impacts of policies, programs, and practices using random assignment study designs (Haskins & Baron, 2011; Hess & Petrilli, 2006).

In December 2015, Congress passed the ESSA, the long-awaited reauthorization of the ESEA. ESSA represents a significant devolution of decision-making authority from the federal government to state and local agencies. With this new authority comes the explicit expectation that local policy makers rely on "evidence-based" information. Indeed, ESSA includes the term "evidence" or "evidence-based" more than 80 times throughout the legislation. This language is most notable in the regulations for the large formula and competitive grant programs. For the largest category of Title I funds, for instance, leaders are to develop improvement plans for low-performing schools in which they select "evidence-based" interventions, programs, or activities for adoption. These programs must have demonstrated impact through previous studies that meet one of three "tiers of evidence" as defined by ESSA. Each tier specifies the types of research methods used to justify claims made about program impact (Section 8002; Civic Impulse, 2017):[2]

- *Strong evidence* has at least one well-designed and implemented experimental study, meaning a randomized controlled trial, that shows a positive impact;
- *Moderate evidence* has at least one well-designed and implemented quasi-experimental study, for example, a regression discontinuity analysis, that shows a positive impact;
- *Promising evidence* has at least one well-designed and implemented correlational study that controls for selection bias that shows a positive impact.

Overall, ESSA presents normative views of educational improvement as anchored in the adoption of programs and interventions and leadership practice as focused on making decisions about resource adoption. In this view, leaders are solitary actors engaged in principally cognitive acts of reviewing evidence and drawing conclusions from causal impact studies to make decisions over programs.

However, the conception of leadership on which ESSA and other contemporary policies are based is not particularly well-founded in contemporary theories of leadership. Leadership is more aptly understood as practice unfolding across interactions among individuals and groups who are situated in and interact with the contexts of their schools, school systems, and communities (Diamond, 2013; Diamond & Spillane, 2016; Spillane, 2006; Spillane et al., 2004). A distributed perspective on leadership draws attention to "how leadership actually gets done on the ground, what people actually do together (and with what resources), how they do it, and why they do it" (Diamond & Spillane, 2016, p. 147). This literature teaches us to center the daily activities of leadership practice to understand how leaders engage with research across different contexts, and for different purposes.

The view of evidence use in ESSA runs counter to the body of empirical research that investigates educators' *actual* engagements with research in their day-to-day work. From this growing literature base, we know that the daily activities of local leadership practice are likely far broader than those addressed in ESSA. Leaders participate in a range of activities: They design or adopt initiatives to improve teaching and learning, create structures and processes to implement these initiatives at sites or at scale, and provide management, oversight, and leadership in schools or system-wide (Cobb et al., 2018; Coburn, Toure, & Yamashita, 2009; Rorrer et al., 2008; Wong et al., 2020). Within these spaces, leaders rarely, if ever, decide or act on research alone.

Leaders navigate a multitude of considerations in their practice, including the needs of different stakeholder groups, logistics and policy constraints, and beliefs, principles, and values, all intertwined with race, class, and language discourse (Huguet et al., 2021); their efforts are fundamentally shaped by organizational and institutional contexts (e.g., Turner, 2015).

Furthermore, when policy makers encourage education leaders to use research to select a program or intervention, they implicitly invoke a theory of action in which evidence from research findings directly shapes decisions related to policy or practice, particularly those related to program adoption (Sharkey & Murnane, 2006; Weiss & Bucuvalas, 1980). However, studies of research use indicate that evidence use is not a single process, serving a single purpose (Coburn, Honig, & Stein, 2009; Farley-Ripple, 2012; Huguet et al., 2019; Wong et al., 2020). For example, conceptual uses of research do not inform one specific decision directly. Instead, research can influence what district leaders prioritize and focus on as they do their work which, in turn, influences a variety of policy actions and problem-solving decisions across the school system (Farrell & Coburn, 2016). Research can inform leaders' thinking about various problems, focusing attention on issues that were not in their immediate sights or cultivating ideas for improving current programs and policies (Cobb et al., 2018; Hubbard, 2010; Weiss & Bucuvalas, 1980). Leaders also use research symbolically, where findings are invoked to persuade others toward a held position or justify decisions already made (Asen et al., 2011). Recently, Coburn et al. (2020) surfaced a new form of research use, latent use, which occurred as district leaders embedded research in artifacts, which then guided the work of leaders in other routines in substantive ways.

Furthermore, education leaders may have different conceptions of "what counts" as evidence than the narrow focus on effectiveness and impact studies prescribed in current policies (Mills et al., 2020). Corcoran et al.'s (2001) study of one district's use of research found that "the distinctions between empirical research, theories, and simple advocacy were not well understood" by educators (p. 80), while Coburn and Talbert (2006) found that even within a single school district, leaders' conceptions of research evidence can vary widely. Often, pressures to be "evidence-based" for academic performance can be confused with policy measures to use research evidence. In Finnigan et al.'s (2013) examination of more than 100 school and district leaders' definitions of "research evidence," many equated research evidence with test score data. These differing definitions are troublesome if education leaders consider chosen policies and programs to be "research-based" when the evidence that grounds them is, in fact, quite weak (Dynarski, 2010; Lubienski et al., 2014; Mills et al., 2020; Welner et al., 2010).

More recent efforts of education leaders' engagements with research within their ongoing practice shed additional light on these questions. In an in-depth case study of leaders in three districts, district leaders used research as federal policies intended, to select programs, but they also reported drew on research to support their own professional learning, guide their instructional leadership activities, and monitor and support implementation of district-adopted programs and practices (Penuel et al., 2018). Beyond the impact studies required by ESSA and other policies, these leaders found a variety of types of research useful to their work, especially when research was extended to offer practical frameworks and guidance. Findings from this study suggested that leaders may turn to different kinds of evidence for different activities. When selecting a program, an efficacy study was helpful, but when designing professional learning, a source that provided new ideas, language, or a framework was seen as more useful. Yet, it is not clear if these patterns generalize to a wider sample of education leaders nationally.

Finally, it is conceivable that education leaders in different roles might turn to different types of evidence in efforts that span multiple domains of practice. Leaders in different role groups—such as those in different departments in a large school district central office—often have unique responsibilities that may lead to engaging with research in a variety of ways (Coburn & Talbert, 2006). Leaders in a research, assessment, and evaluation office may be called upon to locate and share research findings with other departments and schools, or might conduct research studies and evaluations of their own programs. Leaders working primarily on federal and state policy compliance might encounter research most frequently as part of their Title I responsibilities. District leaders working in curriculum and instruction might develop and offer professional learning opportunities that share instructional strategies with school-based educators. Yet, some studies do not show differences in research use by role (e.g., Mills et al., 2020). In contrast, in earlier work, we found differences by role groups in leaders' reports of the extent to which they used research in conceptual and symbolic ways (Penuel et al., 2017). Because of the mixed evidence in the literature base, an investigation of useful research in leaders' practices should consider how position in the school system may make a difference.

Overall, we need a better understanding of the purposes for which education leaders use evidence, and the types of research that they find useful in doing so. Rather than prescribing "what counts" as research evidence in decision making, we need to understand "what's useful" for education leaders across the full range of activities in their daily practice so that policy can better support leaders in accessing, recognizing, and contributing to high-quality evidence. We also need to compare these results with the expectations

of recent federal policies to know more about the potential gap between "what counts" and "what's useful." To this end, we focus here on what Tseng (2012) calls the "demand side" of research use through a nationally representative study on the purposes for which education leaders turned to research, the kinds of research evidence they found useful, and the extent to which the sources of research they used align with policy expectations.

Research Methods

Our study examines the following questions:

1. What types of evidence characterize the sources of research that leaders name as useful?
2. For what activities of daily practice do leaders claim that research is useful?
3. How do these types of evidence and activities compare with expectations in federal policy guidelines, namely, those set forth in the Every Student Succeeds Act (ESSA)? How do they vary by professional role?

We examined these questions through the analysis of an open-ended item on a nationally representative survey of school and central office leaders from mid- and large-size school districts in the United States.

Population

Our target population for survey respondents was school and district leaders from mid- and large-size U.S. urban districts who were likely to be involved in K–8 instructional decision making. We chose K–8 because there is more research available on effective programs and interventions at these grade levels and because more variety exists in the curricular materials, assessments, and other instructional programs districts may implement. We focused on the local level because principals and central office leaders make most programmatic decisions. We aimed for balanced participation across the following role groups: Deputy, associate and network superintendents and school supervisors; curriculum and instruction department leaders; special education department leaders; accountability, assessment, and research department leaders; federal programs directors; leaders crossing multiple departments; and elementary, middle and K–8 school principals.[3] We focused on the 1,000 largest school districts, serving more than 9,000 students each according to NCES (National Center for Education Statistics) Common Core data, as smaller districts may not staff some of the positions in our sampling frame.

Sample

We identified a set of 41,000 school and central office leaders in the 1,000 largest school districts across the above role groups using a data set purchased from MDR, a private education marketing firm that develops databases of educator and institution information. To address overrepresentation of school principals, we pruned the target population to 14,276 by taking a random sample of 10 principals within each school district.[4]

We randomly sampled leaders in the target population within two strata of interest: role group and district enrollment. We aimed to achieve 100 responses for each role group, with half of each above and below the median enrollment of 17,860 students across sampled districts. We established two stratified random samples, our primary field test sample and a reserve sample, each containing 168 potential respondents by role or 84 for each role by size stratum.[5] The reservoir sample was created in case we were unable to achieve our target of a 60% response rate by stratum. We pulled additional cases from the reservoir sample for groups with lower-than-anticipated response rates or due to issues with contact information. In the process of securing current contact information, a potential respondent's role had changed, they were replaced, when possible, by whomever had taken over the target role.

After multiple rounds of outreach, the full survey sample consists of 733 individuals from 485 unique school districts across 423 cities and 45 states, with an overall response rate of 51.5%, varying from a low of 33% and 37% for deputy, associate and network superintendents and principals in larger school districts, to 66% and 71% for assessment and special education leaders. The sample of 485 districts appears to be representative of the larger population of 904 districts found in the MDR list with respect to district size.

Within the 733 respondents, a total of 359 leaders fully completed the open-ended item that is the focus of this study, in which they were asked to name a piece of research that was useful to their work, and for what purpose they used it.[6] While complete responses to this item represented only half (50%) of the 733 respondents to the entire survey, the proportion of respondents who completed the open-ended item by role category paralleled the response rate for the entire survey. Because some respondents named the same piece, leaders named a total of 262 unique sources. Table 1 below compares sampling frame population sizes, the numbers of randomly sampled respondents who were invited to respond to the survey, the number of people per role group that completed the full survey, and the number of people per role group that completed the focal survey item described below and gave a response to it that allowed researchers to identify a specific piece of research.

TABLE 1
Respondents by Role Group

Role group	Sampling frame	Field test sample	Field test responders	Responders who named identifiable sources
Deputy/associate/network superintendents	1.304	202	90	38
Curriculum and instruction	1,941	239	115	59
Special education	742	167	102	45
Assessment	745	152	91	40
School principals	7,123	325	138	69
Federal programs	1,138	144	89	42
Multirole	1,283	203	108	66
Total	14,276	1,422	733	359

Survey Instrument

The full survey focused on how school and district leaders use research to inform their decision making, including their attitudes toward research, the ways they used it (i.e., instrumentally, conceptually, and symbolically, as described above), their efforts to acquire it, and the culture of research use in their organizations. Each page of the survey included a definition of research as "an activity in which people employ systematic, empirical methods to answer a specific question." In addition to broad patterns of research use, we wanted to better understand what types of evidence leaders were referring to in reporting their own use. We included an open-ended item that asked respondents to name a particular piece of research that they found useful in their work, providing as much information as possible so that we could locate it ourselves (i.e., title, author, year published, and topic), as well as why they found it useful. Data on survey piloting and validation activities for the survey, which involved extensive cognitive piloting of the survey instrument with leaders in different role groups are described elsewhere (Penuel et al., 2016; Penuel et al., 2017).

Data Analysis

A team of three researchers began analysis by locating each source online and either downloading a copy or, for books, accessing them through the university and local libraries. Research team members, led by the third author, first coded 20 sources together to develop codes and understandings (see Table 2 for full code list). Two research assistants then proceeded to code sources individually, reviewing coding together with the lead researcher weekly. Once all three researchers agreed on 70% of codes applied, two research assistants continued to code individually, noting their level of confidence about their coding. The team together reconciled all coding marked as "unsure" on a weekly basis, with the team lead reviewing and making the final decision on all coding marked as "somewhat confident" as well as a random sample of coding marked as "very

TABLE 2
Code List

Codes and subcodes	Codes and subcodes
Format	Methods
Book	Experimental
Research/policy report	Quasi-experimental
Journal article	Correlational
Practitioner-oriented magazine article	Case study
Book chapter	Mixed methods
Online media	Synthesis/review
Dissertation	Theoretical analysis
Other	Not specified
	Not applicable
Why useful	Group comparisons
Selecting programs/interventions	Random assignment
Designing programs/policies	Controlled for differences
Supporting and monitoring implementation	Did not control for differences
Supporting leaders' own learning	Did not compare groups
Providing instructional leadership for others	Not specified
Not specified	Not applicable
ESSA Tier	
Strong	
Moderate	
Promising	
None	

Note. ESSA = Every Student Succeeds Act.

confident." The team lead then took one final pass to review all final coding determinations of all sources.

First, we coded the format of the source named (i.e., book, research or policy report, journal article, or other, including tools/programs, media, book chapters, and dissertations). Next, we investigated the types of evidence that characterized each source. The team first coded each source as presenting an *original analysis*, an *evidence-informed* framework, or *other*. We determined that a source presented an *original analysis* if it included details on research methods, whether these

appeared in a "methods" section of an article or an appendix of a book, or represented systematic theoretical analyses, such as those conducted in legal and other humanities-based fields.[7] We counted syntheses of research as original analyses only if they detailed their methods of selection criteria for included studies and their process of synthesizing included studies. Those sources that did not represent original analyses were coded as *evidence-informed* if they cited other research to offer a framework, often in an introduction to a book or article. We coded as *other* the few sources that were not based in research. These included pieces that made claims about ideas or strategies based only in anecdotal evidence (i.e., stories or examples) and those that offered strategies without referencing evidence to justify them, such as implementation guides or workbooks.

For sources that presented original analyses, we coded for the *type of research methods* used (i.e., experimental, quasi-experimental, correlational, systematic synthesis, case study, ethnography, theoretical, and other). We then determined whether these original analyses would meet ESSA's *strong, moderate*, or *promising tiers* of evidence. Sources that used experimental, quasi-experimental, or correlational methods and adequately controlled for group differences qualified for the strong, moderate, or promising tier, respectively. In addition, because the ESSA tiers allow leaders to gather evidence from multiple studies to satisfy large sample requirements, we counted systematic syntheses as meeting the criteria for a tier if they only included studies that met the criteria for that tier. For example, syntheses that only included randomized controlled (experimental) studies we coded as meeting the "strong" tier. On the other hand, syntheses that included a variety of quantitative and qualitative methods were coded as "nontier qualifying syntheses."

Next, we coded each response to the activities of daily leadership practice with which the research was associated. Here, we used the categories identified in an earlier analysis of pilot survey data for the same open-ended item (Penuel et al., 2018), including *selecting programs/interventions*, *designing programs/policies*, *supporting and monitoring implementation*, *supporting leaders' own learning*, and *providing instructional leadership for others*. For example, we coded the response, "It provided the basis for our selection of an early education literacy program" as *selecting programs/interventions* while we coded the response, "Changed the paradigm for supporting low-income students in college access work" as *supporting leaders' own learning*. We conducted counts on these codes to summarize the types of evidence reflected in the sources named by education leaders and compared findings by the practices for which the leaders claimed the research was useful. Through cross-tabulations, we then investigated differences statistically (using χ^2 tests) in these results by respondents' professional role group since we anticipated role group might be associated with differences in the research that leaders found useful.

Limitations

There are limitations to the survey data that preclude us from concluding about how leaders engage with their named sources in situ. The process of evidence use is an interactive one (Contandriopoulos et al., 2010), and other methods such as observations can provide much greater detail on the role for research in distributed leadership practice over time (e.g., Asen & Gent, 2018; Huguet et al., 2017). For instance, we do not know how leaders went about applying the ideas within their contexts considering other constraints, forms of information, or values (Huguet et al., 2021). Second, our goal is to compare practitioners' named resources against policy standards. While we recognize the importance of high-quality, relevant research and methodologies (Ming & Goldenburg, 2021), it is beyond the scope of this article to make normative claims about the quality of the research, "depth" of use (Farley-Ripple et al., 2018), or the quality of research use (Monash Q Project, 2020).

Findings

Below, we explore the nature of the evidence that leaders named as useful across multiple domains of their leadership practice.

Books With Broad Topics Were Most Popular Resource for Leaders

Responses to the open-ended item that asked leaders to name a particular piece of research they found useful to their work revealed patterns that run contrary to policy assumptions. Contrary to the common belief that education leaders prefer short articles or reports due to time constraints, 57% of the 359 respondents named books, which represented 44% of the 262 unique sources named (see Figure 1). The next most common formats of sources named were research or policy reports (17% of respondents; 22% of unique sources) and journal articles (13% of respondents; 18% of unique sources). Remaining sources included a mix of practitioner-oriented magazine articles, book chapters, online media, and dissertations.

Sources that leaders named as useful focused on a variety of topics that leaders are responsible for, ranging from teaching, learning, and assessment practices to school- and school system-level organizational management and policies. Eighteen percent of sources focused on a particular disciplinary content area, such as literacy or mathematics, while 28% of sources focused on a particular student subgroup, such as students in Special Education or emerging bilingual learners. However, the great majority more broadly addressed topics that had implications across academic disciplines or students' identity groups.

Some sources and authors were notably popular. In fact, sources authored by three well-known names—John Hattie
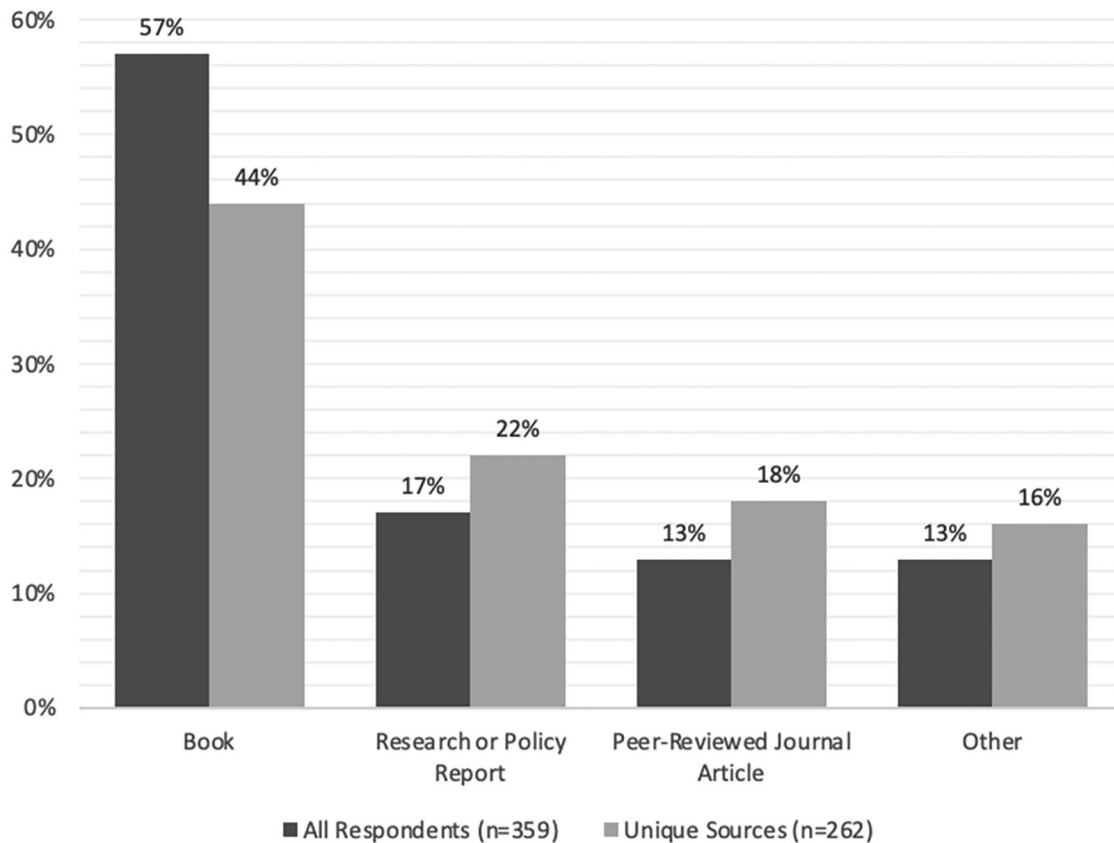
FIGURE 1. *Formats of the sources education leaders named as useful.*

(46 mentions), Robert Marzano (33 mentions), and Richard DuFour and colleagues (14 mentions)—made up one-quarter (25%) of all responses. These authors have a strong presence in education leaders' professional associations, where leaders claimed they most frequently accessed research in another part of the survey (Penuel et al., 2016).

Evidence-Informed Frameworks Were Prevalent, Followed by Original Analysis

We determined whether each of the 262 unique sources named represented an *original analysis*, an *evidence-informed* framework, or *other* with regard to the type of evidence. As Figure 2 shows, less than half (47%) of sources presented original analyses, including less than one-quarter (22%) of books, and two thirds (68%) of all other formats.

The remaining half (53%) of sources did not present original analyses. Most of these represented evidence-informed frameworks, which accounted for over one third (37%) of all sources, including over half (59%) of books and one fifth (20%) of all other formats. Examples of evidence-informed frameworks include *A Framework for K–12 Science Education* (National Research Council, 2012), Black and Wiliam's (1998) paper from *Phi Delta Kappan* on formative

assessment and learning, and Edmonds' (1986) framework for characterizing effective schools. We did not evaluate the quality of the evidence on which these frameworks were based. Most provided a summary of the relevant research (e.g., in the first chapter of a book or in a summary section before a report) that would require education leaders to be familiar with the scholarship cited to gauge the quality of the evidence claimed as the basis for a particular framework. The remaining 17% of sources were classified as "not research-based." These either only pointed to anecdotal examples to make their claims, which would not qualify as research under the definition given in the survey, or did not state any reasons for their claims (e.g., they offered how-to strategies with a justification for the strategies' effectiveness).

Few Pieces of Research With Original Analysis Met ESSA Tiers

For research coded as original analyses, we then determined the research methods used, and whether these methods met the criteria for ESSA's strong, moderate, or promising tiers. Of the original analyses (47% of sources named), just under one fifth (18%) qualified for an ESSA
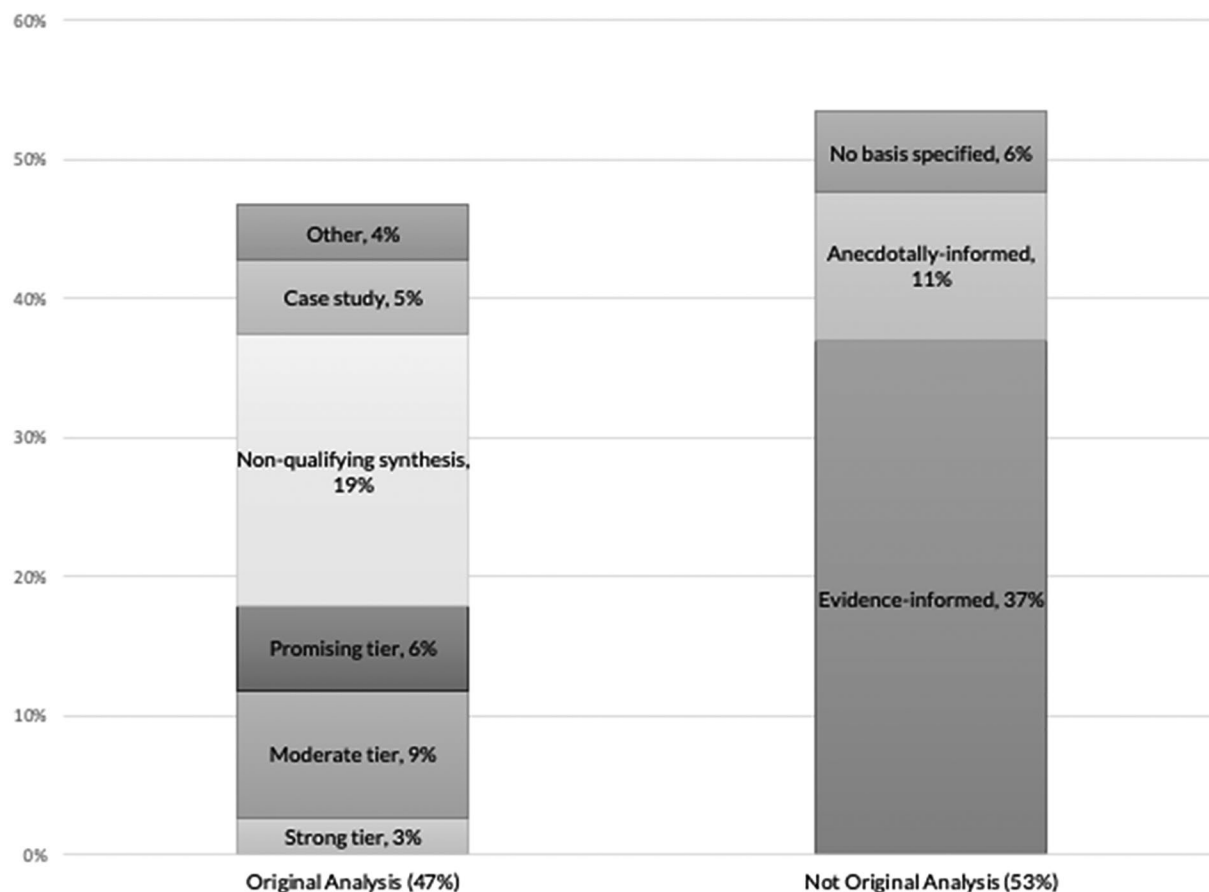
FIGURE 2. *Basis of evidence for the sources education leaders named as useful.*
Note. n = 262 sources.

tier (including research syntheses for which all included studies met a tier). Another one fifth (19%) of original analyses reflected "not tier-qualifying" syntheses that did not meet the criteria an ESSA tier, another 5% used case study methods, and the remaining 4% used other methods, such as mixed (quantitative and qualitative) methods, theoretical analyses, and a variety of methods reflected in edited book volumes. In other words, few of the studies that leaders found useful met the criteria in policy for what kinds of evidence should be used to inform decision making.

Leaders Reported Research Was Useful Across Range of Activities of Leadership Practice

Along with the piece of research, we asked leaders to explain, "Why was it useful?" Among the 359 codeable responses to the open-ended item, responses from 278 participants provided enough information to identify reasons for the usefulness of the research. The remaining respondents gave an answer that only provided a description of the piece (e.g., "Study of formative assessment practices") rather than a reason, merely stated "yes" or "very useful," or

did not complete this part of the item so were excluded from the analysis presented here.

Consistent with the multifaceted nature of leadership practice described earlier, district leaders discussed how research was useful across multiple dimensions of their work (see Figure 3). About one third (30%) of respondents' reasons related to supporting leaders' learning by developing their own knowledge. One assessment coordinator noted that the book *The Teaching Gap* ". . . expanded my perspective on the challenges teachers face when attempting to improve instructional practice" (Stigler & Hiebert, 1999). Next, in 28% of responses, leaders named research that helped them as they designed policies, programs, and initiatives. A Special Education director reported that a book on gradual release of responsibility was ". . . useful because it provided the foundation for the district's instructional framework" (Fisher & Frey, 2013). In 23% of responses, leaders reported drawing on research to provide instructional leadership for others in central offices or schools. A bilingual education administrator described how their district had drawn on an edited volume on language development to support professional development (California Department of Education, 2010):
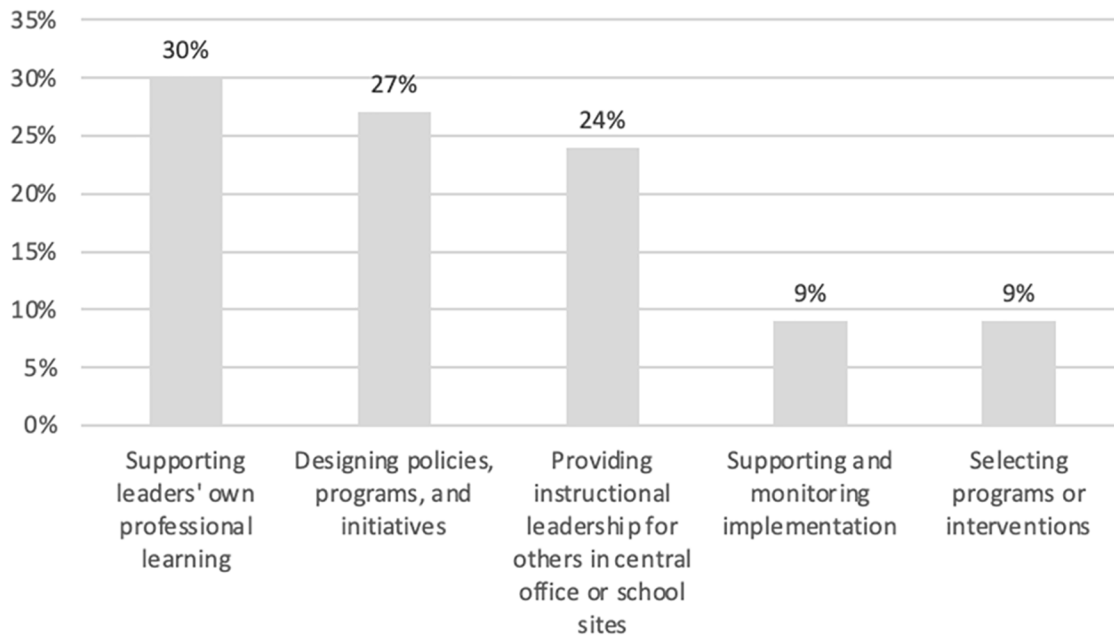
FIGURE 3. *Purposes for which sources named by education leaders were useful.*
Note. n = 278 respondents.

> We developed our professional learning for our leadership teams using these research-based approaches and guidelines for instruction by William Saunders and Claude Goldenberg [in chapter 1]. Leaders read the research, and then we applied the learning. Leaders gained deep insight and understanding of instruction and programs for English Learners, and how to implement a complex process systematically and cohesively within the school day.

In 11% of responses, leaders described how a piece of research helped them support and monitor implementation. For example, an assistant superintendent explained that a book on instructional rounds gave him and his colleagues

> . . . direction in how to best observe classrooms and ensure we're noting information that can improve instruction, not just looking at whether or not teachers are complying to a prescriptive set of actions that are "checked off" of an observation tool. (City et al., 2009)

In contrast to policy guidelines, when leaders named a source of research evidence that was useful to their work, only 9% named one they had used to select programs or interventions. One chief academic officer noted that a book on Reading Recovery ". . . led our district to adopt this program for early intervention, reallocate funding and increase staffing" (Clay, 1993).

Finally, we considered the nature of the evidence named across these different areas of practice. Figure 4 summarizes the evidence basis compared with the purpose for which leaders stated the piece they named was useful. Although only 25 (9%) of 278 respondents reporting engaging their piece of research to select an intervention or program, 76%

of those who did named a piece that reflected an original analysis, with 40% naming a piece that qualified for an ESSA tier. For example, a special education leader indicated that an experimental study by Powell et al. (2015) that focused on an intervention to support prealgebraic thinking informed a decision to adopt the intervention in the district. Similarly, an assessment leader in another district used an experimental study of a curriculum-based measurement intervention (Fuchs et al., 1990) to guide selection of screening instruments for their Response-to-Intervention program. Both these studies met ESSA's highest evidence tiers.

In contrast, about half (39 to 54%) of those indicating research was useful for some purpose other than selecting curriculum or interventions named sources that presented original analyses, with about one quarter or less (8% to 27%) qualifying for an ESSA tier. For example, an English/Language Arts coordinator indicated they used L. Resnick's (2010) paper outlining principles for a "thinking curriculum" helped shape their district's curriculum writing and implementation. Similarly, an assistant superintendent of curriculum and instruction used a framework for rigor and relevance (Daggett, 2008) to develop a framework for intervening with ninth-grade students. This framework did not meet any evidence tier for ESSA.

Given the differences in the types of evidence leaders named for different purposes, and the range of practices in which leaders engage in different professional roles, we investigated whether the basis of evidence reflected in sources named differed by leaders' roles (Figure 5). Our
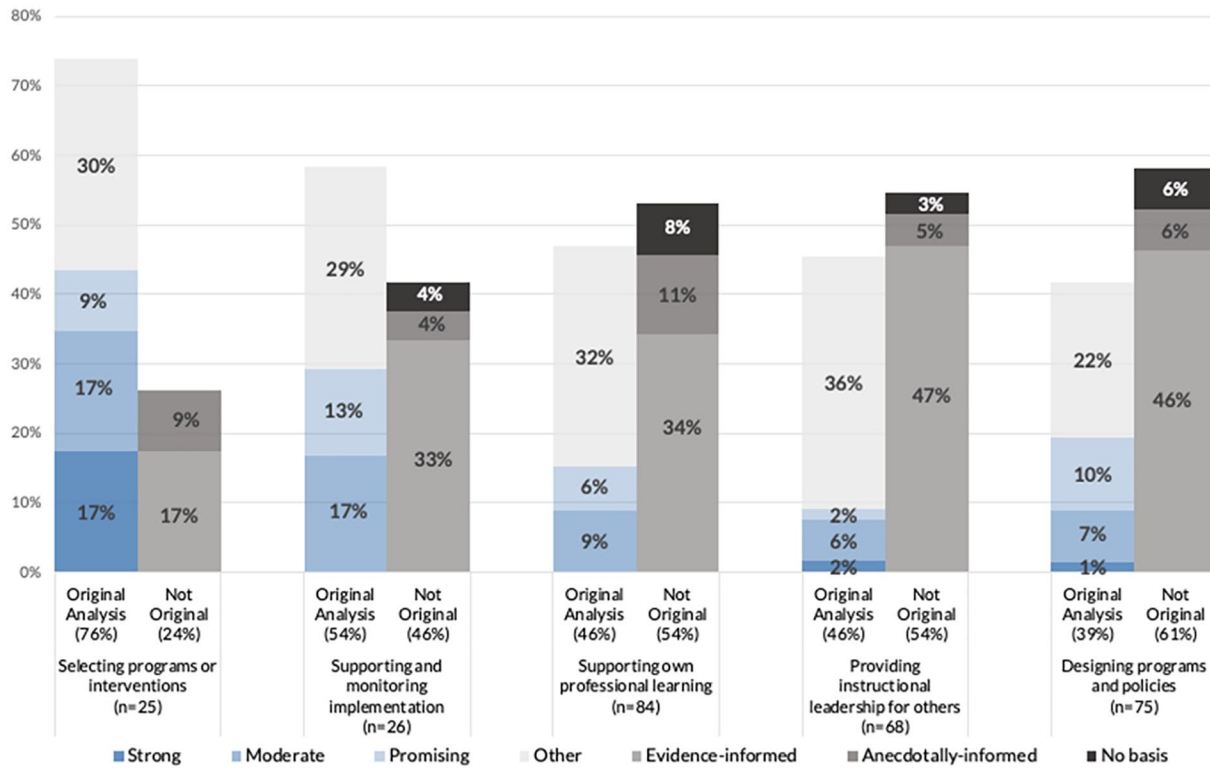
FIGURE 4. *Basis of evidence reflected in sources named by education leaders according to activity.*
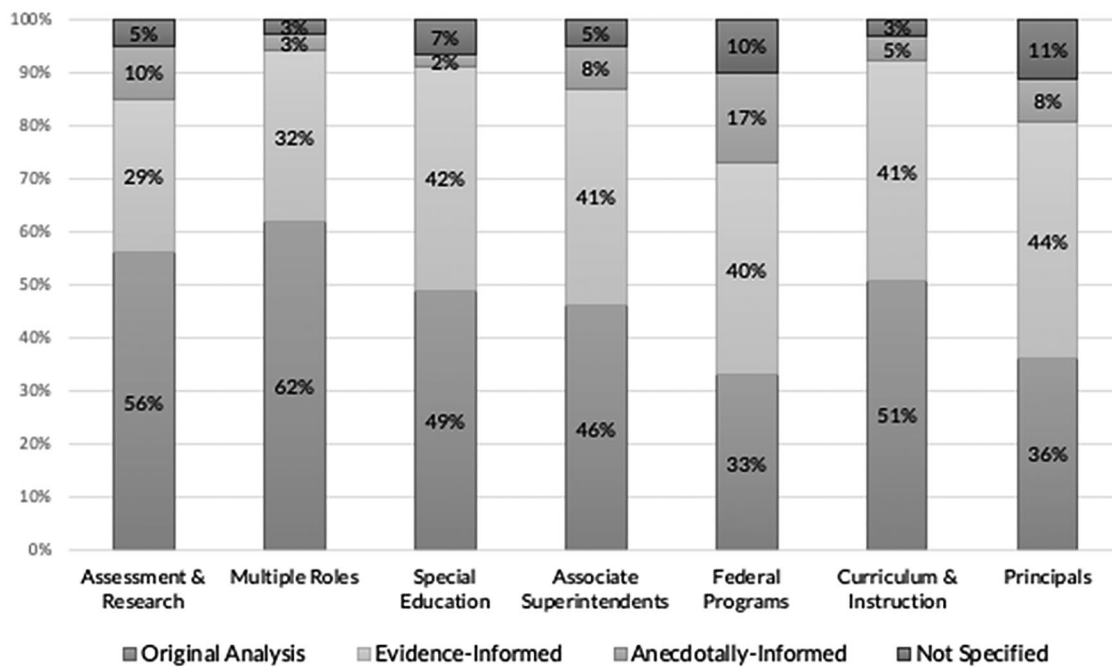*Note.* n = 278 respondents.

FIGURE 5. *Basis of evidence by professional role.*
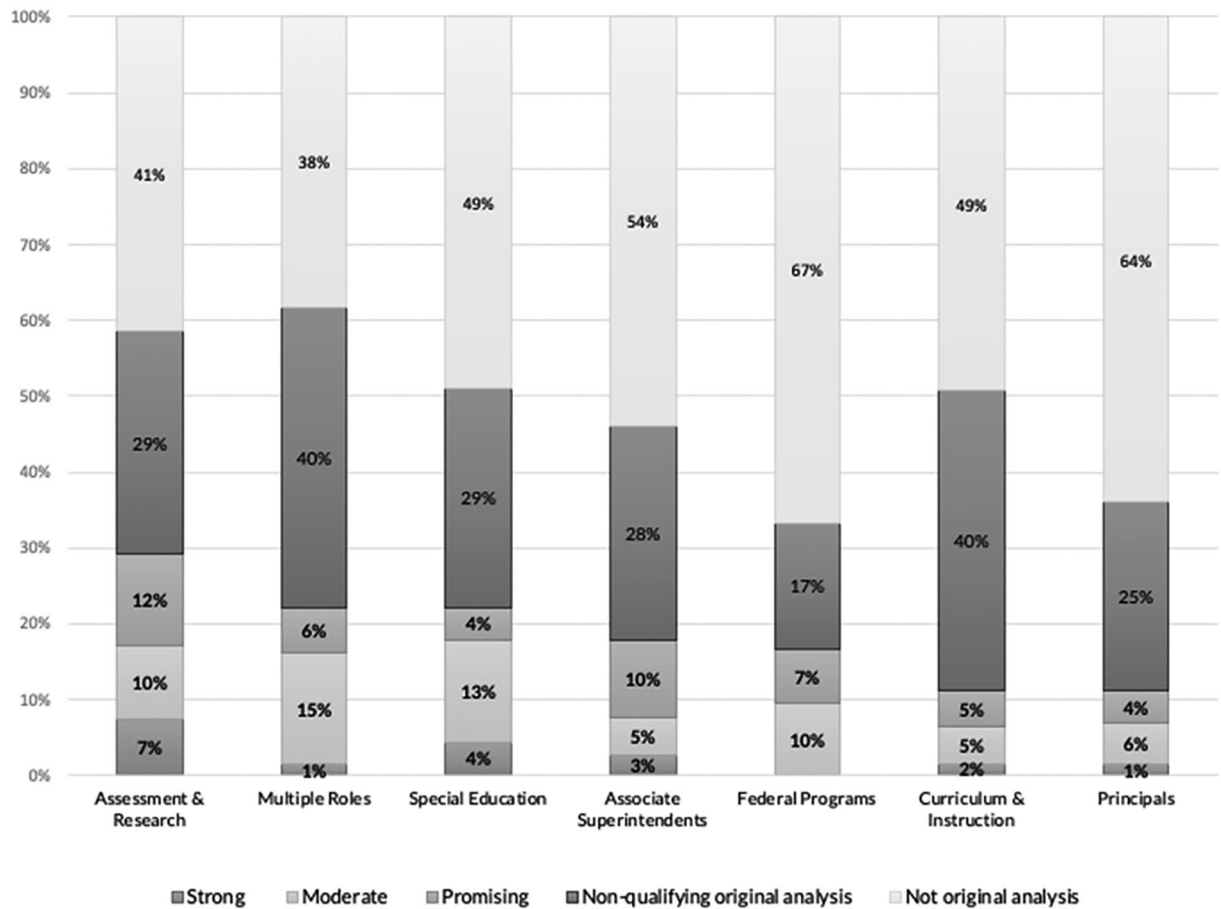*Note.* n = 359 respondents.

FIGURE 6. *Basis of evidence and ESSA tier by professional role.*
Note. $n = 359$ respondents. ESSA = Every Student Succeeds Act.

analysis did not reveal any significant differences related to the types of evidence used, $\chi^2(24, n = 359) = 19.409$, $p = .73$. This indicates that despite some roles being likely more linked to experience with research (e.g., assessment leaders) or more obligated to turn to research for evidence (e.g., federal programs leaders), these leaders' use of evidence was similar to those with less direct responsibility for using research in their practice. Similarly, professional roles did not matter significantly for whether evidence met ESSA tiers (Figure 6), nor were there significant differences between role and leadership practices named (Figure 7).

Discussion and Implications

A fundamental premise of ESSA and other contemporary policies is that evidence-based, externally developed resources are a primary lever for educational improvement. The adoption and use of evidence-based programs and interventions is expected to drive improvement in districts, schools, and classrooms. This focus casts the primary role of school and district leaders as people who make decisions about what programs to adopt, scale, or eliminate based on

evidence of impact. These premises have policy makers focused keenly on the development of such things as evidence standards, evidence-based resources, infrastructure to publicize these resources and their evidence base, and the incentives and sanctions needed to guide education leaders' decision making accordingly.

Our analysis of a nationally representative survey of district and school leaders makes clear how narrow these conceptions of leadership practice and the possible roles for research are. While selecting programs is an important aspect, it is one among many different activities in regular leadership practice. Other domains include designing policies, programs, and initiatives, providing instructional leadership in central office and schools, supporting and monitoring implementation, designing professional learning for others, and advancing their own professional learning. Within these activities, leaders seek sources of evidence beyond those privileged in contemporary education policy, that is, peer-reviewed studies of the impacts of programs and policies. Leaders relied on a wide variety of evidence-informed sources—including many books—that supported them in their complex, multifaceted leadership practice. The
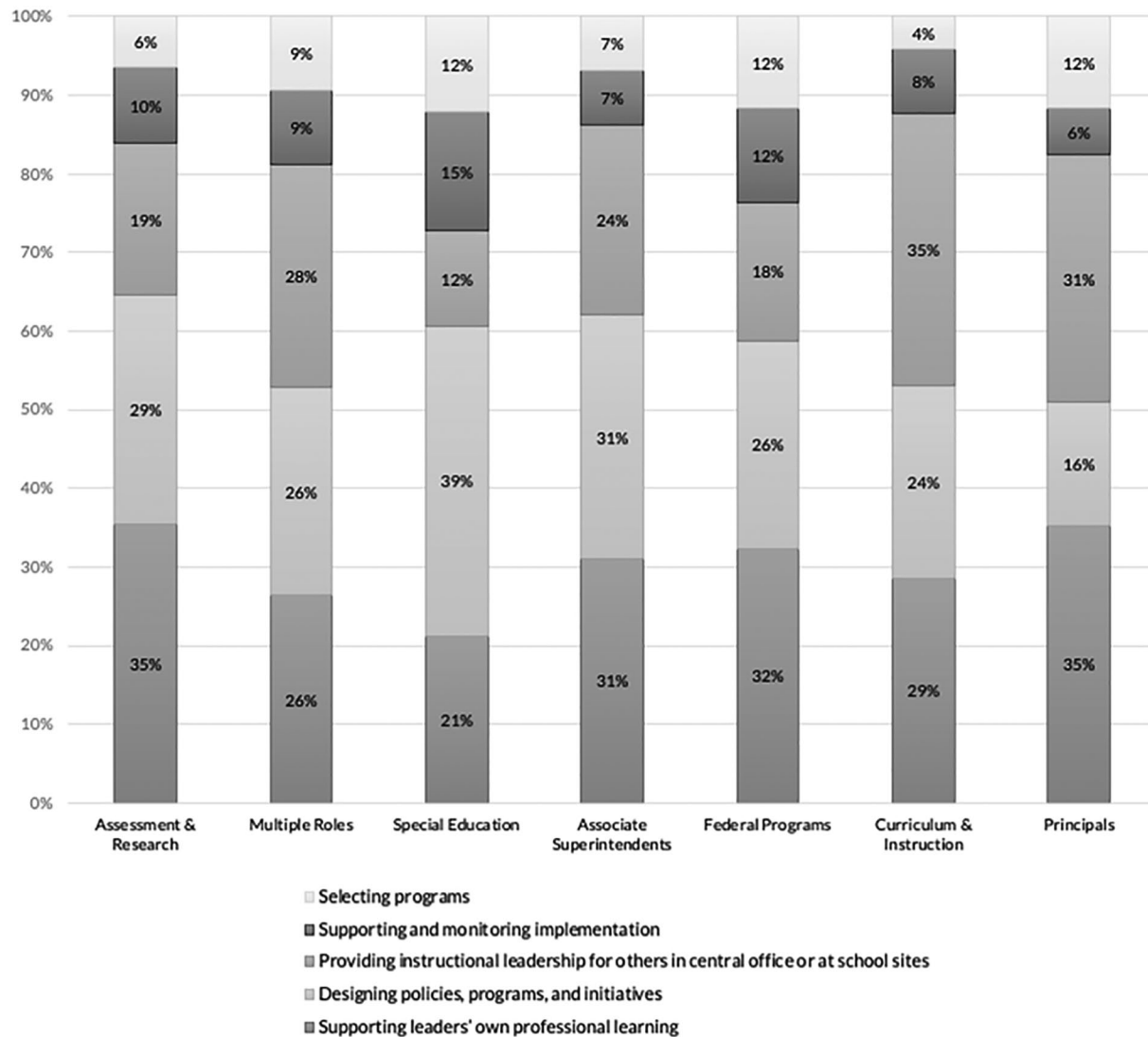
11

FIGURE 7. *Activities for why research is useful by professional role.*
Note. n = 278 respondents.

most popular sources directly addressed the kind of work leaders did (although a leader's role was not related to evidence they cited as most useful to them as one might expect.)

Some may find these findings surprising or even discouraging. Little research named by leaders fell into ESSA's top tier of evidence, suggesting that the quality of evidence as understood by many in the policy community may figure very little in education leaders' efforts. Roles in which we might hope or expect leaders to rely on evidence based in causal research designs—such as leaders in federal programs offices where many decisions to purchase programs or interventions are made—are not ones where leaders are more likely to turn to such evidence. The variety and types of research cited underscore what others have found, namely, that leaders likely have very different conceptions about what research *is* from researchers (Coburn & Talbert, 2006;

Corcoran et al., 2001; Finnigan et al., 2013; Mills et al., 2020).

However, this perspective is incomplete and ignores some possible strengths in how leaders engage with resources that count as research for them. For one, many leaders did name specific pieces of research, and it was relatively easy for a team of researchers to identify from partial information reliably the piece of research named. In addition, the large number of books named suggests that time to read is not an inherent limitation to research use, as is often implied in solutions to the problem of evidence use that focus on writing shorter reports. Significantly, leaders engaged with research to support a wide variety of activities central to leadership practice, from designing professional development to monitoring implementation. Finally, regardless of role, a primary motivation for many leaders to read research

is that they view being up to date about research as a part of their professional identity. These findings suggest an expansive portrait of how research evidence can inform leadership practice through multiple pathways.

Our findings have implications for researchers and research production. We did not find that "shorter is better," in that many books appeared as sources of research that leaders found useful. Thus, there is little evidence to suggest that creating shorter research briefs for leaders to read would lead to greater research use by leaders. Instead, if books were a major source of influence, preparing researchers to develop guides that synthesize research into usable frameworks for action might be more valuable. Second, we need to encourage and incentivize researchers to post and share their research within a wider range of venues, including professional associations where leaders get their research (Penuel et al., 2016). We likely also need to support more substantive and ongoing engagement with researchers themselves, through partnerships and critical engagement, if the engagement with research is to accurately reflect what conclusions can be viably supported by research (Farrell et al., 2021).

More broadly, though, the findings point to the need for an evidence infrastructure supported by educational policies that center the practice of educational improvement and transformation broadly conceptualized. Policies should better reflect the complex, multifaceted nature of leadership practice and the kinds of research that will be most useful within those activities. For example, in our survey, we found educational leaders regularly turned to research during the design and leading professional development for others. Yet, it is not clear that a randomized controlled trial of an existing program or practice would be a sufficient guide for design activities. Even for researchers, what is most useful for guiding design is a coherent theory of how to support learning that is informed by a mixture of empirical findings, theoretical frameworks, and an understanding of where gaps exist in the knowledge base (Sandoval, 2014). Moreover, design in public settings needs to consider the varied—and often conflicting—needs, concerns, purposes, and constraints of stakeholders (Bang & Vossoughi, 2016). Evidence of the latter can be generated to inform design, but it would not take the form of a randomized controlled trial. While we agree that ESSA's normative emphasis on relying on high-quality research is appropriate, the narrow conception of quality as consisting only of causal impact studies for programs offers inadequate guidance for leadership practice broadly. Rather than solely focusing on impact studies of programs, policies might encourage leaders to turn to a methodologically broader range of sound research to inform their thinking or to support a strategic direction in their districts.

Furthermore, it is possible to imagine making use of ESSA's call for more evidence-based policy making to support some of these broader purposes of evidence use and ways of gathering evidence to inform practice. As we have described elsewhere (Penuel & Farrell, 2017), ESSA's provisions in some places call for the *generation* of research evidence in the context of designing or adapting evidence-based practices (i.e., Tier IV; see Conaway, 2018). The generation of research evidence presents an opportunity for long-term research–practice partnerships to work together to coproduce high-quality, relevant evidence that informs not just studies of impact but also studies of implementation (Farrell et al., 2021; Penuel et al., 2016). In generating research together, moreover, we have the potential to produce a broader culture in our school systems in which evidence serves the aims of equity by including the voices and aims of communities that have been historically marginalized or harmed in research (Doucet, 2019; Kirkland, 2019; Tseng et al., 2018).

Finally, our findings suggest directions for future research. We provide insights in the nature of the evidence with which educational leaders *actually* engage, versus that prescribed by policy or envisioned by researchers. Several authors cited by school and district leaders aim to distill research findings in a practitioner-friendly form in ways that simplify complex research findings and make them actionable. The popularity of research syntheses (e.g., Hattie, 2009) likely adds value to leaders who have responsibility for developing strategies to address a wide range of educational problems they face over the course of their work. Alignment of topic and role suggests education leaders are finding pieces useful that relate to what it is they do, even if the type of evidence included is not linked to role. These preferences for pieces that are actionable, relevant, and connected to daily practice may, at times, stand in contrast to characteristics of research valued by the policy and research communities, particularly as related to issues of study design, internal and external validity, and conclusiveness of findings (Farley-Ripple et al., 2018). Indeed, there is certainly evidence from our study to suggest that not all that counts as useful research would be judged as high quality by even relatively broad standards of research. While recent work has started to conceptualize ways to better aligning research produced and valued by researchers with those preferred by practitioners (e.g., Farley-Ripple et al., 2018; Farrell et al., in press; Gutiérrez & Penuel, 2014; Ming & Goldenburg, 2021), additional investigation is needed on how to develop and support high-quality research that is useful for daily leadership practice.

Furthermore, these findings raise questions about the "marketplace of evidence" in the educational ecosystem, that is, the networks and practices of organizations and individuals that shape and translate research evidence toward different ends (e.g., Jabbar et al., 2014; Scott et al., 2014). Indeed, we suspect that the popularity of some of the resources that leaders cited does not arise solely from the qualities of their products. Certain authors and curriculum

developers work in collaboration with commercial publishers that assist in promoting their work. These publishers adeptly connect to professional associations and practitioner-facing journals as media through which to publicize, disseminate, and profit from research. In this way, publishers and professional associations are supplying the market for evidence that can inform the full complement of leadership practices in ways that current evidence policies alone do not. We would benefit from better understanding of successful "translators" of research who not only package their research differently but also engage with educators in multiple and varied ways and venues (e.g., at conferences and as consultants to districts), as compared with traditional researchers. The influence of publishers and professional associations in mediating access to and interpretation of evidence is an important area of future study.

ORCID iD

William R. Penuel https://orcid.org/0000-0001-7096-6669

Open Practices

The data and analysis files for this article can be found at https://www.openicpsr.org/openicpsr/project/155961/version/V1/view

Notes

1. Several outlets offer resources to assist leaders in selecting programs or interventions that meet ESSA guidelines. For example, see guides published by RAND (Herman et al., 2017), the Institute of Education Science's What Works Clearinghouse (https://ies.ed.gov/ncee/wwc/essa), or the "Evidence for ESSA" website (www.evidenceforessa.org).

2. ESSA also includes a fourth tier of evidence that "demonstrates a rationale" that allows for practices that have a well-defined logic model or theory of action, are supported by research, and have some effort underway to assess its effectiveness. We excluded this tier in our analysis because it was excluded as evidence for interventions under Title I, Section 1003 (School Improvement) which require strong, moderate, or promising evidence to support them.

3. For more information on role groups, see Penuel et al. (2016).

4. The principals excluded from the target population were included in a pool of candidates (that also included central office leaders) who received a pilot test survey prior to the field test.

5. All stratified random sampling was done using the function strata in the R package sampling.

6. An additional 56 respondents responded to the item but gave too few details to accurately discern the source.

7. Examples of books we counted as *original analyses* included Bryk et al. (2009) and Hattie (2009). A book we counted as a *theoretical analysis* was Delgado and Stefancic (2001). Examples of books that offered *evidence-informed* frameworks included Dweck (2007) and Heath and Heath (2010). Examples of sources we coded as *other* (not based in systematic research) included Bambrick-Santoyo (2010) and Fullan (2011). See the references list for the citation of pieces noted here.

References

Asen, R., & Gent, W. (2018). Reconsidering symbolic use: A situational model of the use of research evidence in polarised legislative hearings. *Evidence & Policy: A Journal of Research, Debate and Practice*, *15*(4), 525–541. https://doi.org/10.1332/174426418X15378681033440

Asen, R., Gurke, D., Solomon, R., Conners, P., & Gumm, E. (2011). "The research says": Definitions and uses of a key policy term in federal law and local school board deliberations. *Argumentation and Advocacy*, *47*(4), 195–213. https://doi.org/10.1080/00028533.2011.11821747

Bambrick-Santoyo, P. (2010). *Driven by data: A practical guide to improve instruction*. Jossey-Bass.

Bang, M., & Vossoughi, S. (2016). Participatory design research and educational justice: Studying learning and relations within social change making. *Cognition and Instruction*, *34*(3), 173–193. https://doi.org/10.1080/07370008.2016.1181879

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *92*(1), 81–90. https://doi.org/10.1177/003172171009200119

Bryk, A., Sebring, P., Allensworth, E., Luppescu, S., & Easton, J. (2009). *Organizing schools for improvement: Lessons from Chicago*. The University of Chicago Press. https://doi.org/10.7208/chicago/9780226078014.001.0001

California Department of Education. (2010). *Improving education for English learners: Research-based approaches*. Hippocrene Books.

City, E., Elmore, R., Fiarman, S., & Teitel, L. (2009). *Instructional rounds in education: A network approach to improving teaching and learning*. Harvard Education Press.

Civic Impulse. (2017). *S. 1177—114th Congress: Every Student Succeeds Act*. https://www.govtrack.us/congress/bills/114/s1177

Clay, M. (1993). *Reading recovery*. Heinemann.

Cobb, P. A., Jackson, K., Henrick, E. C., Smith, T. M., & The MIST Team. (2018). *Systems for instructional improvement: Creating coherence from the classroom to the district office*. Harvard Education Press.

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., Spillane, J. P., Bohannon, A. X., Allen, A.-R., Ceperich, R., Beneke, A., & Wong, L.-S. (2020, December). *The role of organizational routines in research use in four large urban school districts* (Technical Report No. 5). National

Center for Research in Policy and Practice. https://files.eric.ed.gov/fulltext/ED612257.pdf

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. https://doi.org/10.1086/505056

Coburn, C. E., Toure, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *Teachers College Record*, *111*(4), 1115–1161. https://doi.org/10.1177/016146810911100403

Conaway, C. (2018). Tier 4 evidence: ESSA's hidden gem. *Phi Delta Kappan*, *99*(8), 80. https://kappanonline.org/conaway-tier-4-evidence-essas-hidden-gem/

Contandriopoulos, D., Lemire, M., Denis, J.-L., & Tremblay, E. (2010). Knowledge exchange processes in organizations and policy arenas: A narrative systematic review of the literature. *Milbank Quarterly*, *88*(4), 444–483. https://doi.org/10.1111/j.1468-0009.2010.00608.x

Corcoran, T. B., Fuhrman, S., & Belcher, C. L. (2001). The district role in instructional improvement. *Phi Delta Kappan*, *83*(1), 78–84. https://doi.org/10.1177/003172170108300116

Daggett, W. (2008). *Achieving academic excellence through rigor and relevance*. International Center for Leadership in Education. https://www.cbsd.org/cms/lib/PA01916442/Centricity/Domain/2713/Achieving%20Academic%20Excellence.pdf

Delgado, R., & Stefancic, J. (2001). *Critical race theory: An introduction*. New York University Press. https://doi.org/10.1093/acref/9780195301731.013.51089

Diamond, J. B. (2013). Distributed leadership: Examining issues of race, power, and inequality. In L. C. Tillman & J. J. Scheurich (Eds.), *Handbook of research on educational leadership for equity and diversity* (pp. 83–104). Routledge. https://doi.org/10.4324/9780203076934.ch5

Diamond, J. B., & Spillane, J. R. (2016). School leadership and management from a distributed perspective: A 2016 retrospective and prospective. *Management in Education*, *30*(4), 147–154. https://doi.org/10.1177/0892020616665938

Doucet, F. (2019). *Centering the margins: (Re)defining useful research evidence through critical perspectives*. William T. Grant Foundation. http://wtgrantfoundation.org/library/uploads/2019/12/Fabienne-Doucet-2019-WTG-Digest.pdf

Dweck, C. (2007). *Mindset: The new psychology of success*. Ballantine Books.

Dynarski, M. (2010). Connecting education research to practitioners—and practitioners to education research. *Phi Delta Kappan*, *92*(1), 61–65. https://doi.org/10.1177/003172171009200111

Edmonds, R. (1986). The characteristics of effective schools. In U. Neisser (Ed.), *The school achievement of minority children* (pp. 93–104). Routledge. https://doi.org/10.4324/9781315060187

Farley-Ripple, E. N. (2012). Research use in school district central office decision making: A case study. *Educational Management Administration & Leadership*, *40*(6), 1–21. https://doi.org/10.1177/1741143212456912

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. https://doi.org/10.3102/0013189X18761042

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

Farrell, C. C., Penuel, W. R., Anderson, E., Coburn, C. E., Brown, S., & Bohannon, A. (in press). Learning at the boundaries of research and practice: A conceptual framework for understanding research-practice partnerships. *Educational Researcher*.

Farrell, C. C., Penuel, W. R., Coburn, C. E., Daniels, J., & Steup, L. (2021). *Research-practice partnerships today: The state of the field*. William T. Grant Foundation. http://wtgrantfoundation.org/library/uploads/2021/07/RPP_State-of-the-Field_2021.pdf

Finnigan, K. S., Daly, A. J., & Che, J. (2013). Systemwide reform in districts under pressure: The role of social networks in defining, acquiring, using, and diffusing research evidence. *Journal of Educational Administration*, *51*(4), 476–497. https://doi.org/10.1108/09578231311325668

Fisher, D., & Frey, N. (2013). *Better learning through structured teaching: A framework for the gradual release of responsibility*. Association for Supervision and Curriculum Development.

Fuchs, L., Fuchs, D., Hamlett, C., & Stecker, P. (1990). The role of skills analysis in curriculum-based measurement in math. *School Psychology Review*, *19*(1), 6–22. https://doi.org/10.1080/02796015.1990.12087335

Fullan, M. (2011). *The six secrets of change: What the best leaders do to help their organizations survive and thrive*. Jossey-Bass.

Gamson, D. A., & Hodge, E. M. (2016). Education research and the shifting landscape of the American school district, 1816 to 2016. *Review of Research in Education*, *40*(1), 216–249. https://doi.org/10.3102/0091732X16670323

Gordon, N., & Conaway, C. (2020). *Common-sense evidence: The education leader's guide to using data and research*. Harvard Education Press.

Grissom, J. A., Egalite, A. J., & Lindsay, C. A. 2021. *How principals affect students and schools: A systematic synthesis of two decades of research*. The Wallace Foundation.

Gutiérrez, K. D., & Penuel, W. R. (2014). Relevance to practice as a criterion for rigor. *Educational Researcher*, *43*(19), 19–23. https://doi.org/10.3102/0013189X13520289

Haskins, R., & Baron, J. (2011). *Building the connection between policy and evidence: The Obama evidence-based initiatives*. NESTA. http://coalition4evidence.org/wp-content/uploads/uploads-dupes-safety/Haskins-Baron-paper-on-fed-evid-based-initiatives-2011.pdf

Haskins, R., & Margolis, G. (2015). *Show me the evidence: Obama's fight for rigor and results in social policy*. Brookings Institution Press.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Heath, C., & Heath, D. (2010). *Switch: How to change things when change is hard*. Crown Business.

Herman, R., Gates, S. M., Arifkhanova, A., Barrett, M., Bega, A., Chavez-Herrerias, E. R., Han, E., Harris, M. A., Migacheva, K., Ross, R., Leschitz, J. T., & Wrabel, S. L. (2017). *School leadership interventions under the Every Student Succeeds Act: Evidence review: Updated and Expanded*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR1550-3.html

Hess, F. M., & Petrilli, M. J. (2006). *No child left behind*. Peter Lang.

Hubbard, L. (2010). Research to practice: The case of Boston public schools, education matters and the Boston plan for excellence. In C. E. Coburn & M. K. Stein (Eds.), *Research and practice in education: Building alliances, bridging the divide* (pp. 55–72). Rowman & Littlefield.

Huguet, A., Allen, A.-R., Coburn, C. E., Farrell, C. C., Kim, D. H., & Penuel, W. R. (2017). Locating data use in the microprocesses of district-level deliberations. *Nordic Journal of Studies in Educational Policy*, *3*(1), 21–28. https://doi.org/10.1080/20020317.2017.1314743

Huguet, A., Coburn, C. E., Farrell, C. C., Kim, D. H., & Allen, A. R. (2021). Constraints, values, and information: How leaders in one district justify their positions during instructional decision making. *American Educational Research Journal*, *58*(4), 710–747. https://doi.org/10.3102/0002831221993824

Huguet, A., Wong, L., Harrison, C., Coburn, C., & Spillane, J. (2019). Research use in schools: A framework for understanding research use in school-level decision making. In M. Connolly, D. H. Eddy-Spicer, & C. James (Eds.), *The SAGE Handbook of school organization* (pp. 618–631). Sage. https://doi.org/10.4135/9781526465542.n36

Jabbar, H., La Londe, P. G., Debray, E., Scott, J., & Lubienski, C. (2014). How policymakers define "evidence": The politics of research use in New Orleans. *Policy Futures in Education*, *12*(8), 1013–1027. https://doi.org/10.2304/pfie.2014.12.8.1013

Kirkland, D. E. (2019). *No small matters: Reimagining the use of research evidence from a racial justice perspective*. William T. Grant Foundation.

Lubienski, C., Scott, J., & DeBray, E. (2014). The politics of research production, promotion, and utilization in educational policy. *Educational Policy*, *28*(2), 131–144. https://doi.org/10.1177/0895904813515329

Mills, K. J., Lawlor, J. A., Neal, J. W., Neal, Z. P., & McAlindon, K. (2020). What is research? Educators' conceptions and alignment with U.S. federal policies. *Evidence & Policy*, *16*(3), 337–358. https://doi.org/10.1332/174426419X15468576296175

Ming, N. C., & Goldenberg, L. B. (2021). Research worth using: (Re)framing research evidence quality for educational policy-making and practice. *Review of Research in Education*, *45*(1), 129–169. https://doi.org/10.3102/0091732X21990620

Monash Q Project. (2020). *Quality use of research evidence framework report*. Monash University. https://www.monash.edu/education/research/projects/qproject/publications/quality-use-of-research-evidence-framework-qure-report

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*.

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 survey of research use among school and district leaders* (Technical Report No. 1). National Center for Research in Policy and Practice. https://files.eric.ed.gov/fulltext/ED599966.pdf

Penuel, W. R., Briggs, D. C., Davidson, K. L., Herlihy, C., Sherer, D., Hill, H. C., Farrell, C. C., & Allen, A.-R. (2017). How school and district leaders access, perceive, and use research. *AERA Open*, *3*(2). https://doi.org/10.1177/2332858417705370

Penuel, W. R., & Farrell, C. C. (2017). Research-practice partnerships and ESSA: A learning agenda for the coming decade. In E. Quintero (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. (2018). What research district leaders find useful. *Educational Policy*, *32*(4), 540–568. https://doi.org/10.1177/0895904816673580

Powell, S., Fuchs, L., Cirino, P., Fuchs, D., Compton, D., & Changas, P. (2015). Effects of a multitier support system on calculation, word problem, and prealgebraic performance among at-risk learners. *Exceptional Children*, *81*(4), 443–470. https://doi.org/10.1177/0014402914563702

Resnick, L. (2010). Nested learning systems for the thinking curriculum. *Educational Researcher*, *39*(3), 183–197. https://doi.org/10.3102/0013189X10364671

Rorrer, A. K., Skrla, L., & Scheurich, J. J. (2008). Districts as institutional actors in educational reform. *Educational Administration Quarterly*, *44*(3), 307–357. https://doi.org/10.1177/0013161X08318962

Sandoval, W. A. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*(1), 18–36. https://doi.org/10.1080/10508406.2013.778204

Scott, J., Lubienski, C., DeBray, E., & Jabbar, H. (2014). The intermediary function in evidence production, promotion, and mobilization: The case of educational incentives. In K. Finnigan & A. Daly (Eds.), *Using research evidence in education: From the schoolhouse door to Capitol Hill* (pp. 69–89). Springer International.

Sharkey, N. S., & Murnane, R. J. (2006). Tough choices in designing a formative assessment system. *American Journal of Education*, *112*(4), 572–588. https://doi.org/10.1086/505060

Spillane, J. P. (2006). *Distributed leadership*. Jossey-Bass.

Spillane, J. P., Halverson, R., & Diamond, J. (2004). Towards a theory of school leadership practice: Implications of a distributed perspective. *Journal of Curriculum Studies*, *36*(1), 3–34. https://doi.org/10.1080/0022027032000106726

Spillane, J. P., Hopkins, M., & Sweet, T. M. (2018). School district educational infrastructure and change at scale: Teacher peer interactions and their beliefs about mathematics instruction. *American Educational Research Journal*, *55*(3), 532–571. https://doi.org/10.3102/0002831217743928

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. Free Press.

Tseng, V. (2012). *The uses of research in policy and practice*. Society for Research in Child Development.

Tseng, V., & Coburn, C. (2019). Using evidence in the US. In A. Boaz, H. Davies, A. Fraser, & S. Nutley (Eds.), *What works now* (pp. 351–368). Policy Press.

Tseng, V., Fleishman, S., & Quintero, E. (2018). Democratizing evidence in education. In B. Bevan & W. R. Penuel (Eds.), *Connecting research and practice for educational improvement: Ethical and equitable approaches* (pp. 3–16). Routledge. https://doi.org/10.4324/9781315268309-1

Turner, E. O. (2015). Districts' responses to demographic change: Making sense of race, class, and immigration in political and organizational context. *American Educational Research Journal*, *52*(1), 4–39. https://doi.org/10.3102/0002831214561469

Weiss, C. H., & Bucuvalas, M. J. (1980). *Social science research and decision-making*. Columbia University Press.

Welner, K. G., Hinchey, P. H., Molnar, A., & Weitzman, D. (2010). *Think tank research quality: Lessons for policy makers, the media, and the public*. Information Age.

Whitehurst, G. J. (2003, April 21–25). *The Institute of Education Sciences: New wine, new bottles* [Paper presentation]. American Educational Research Association annual meeting, Chicago, IL, United States. https://files.eric.ed.gov/fulltext/ED478983.pdf

Wong, L. S., Coburn, C. E., & Kamel, A. (2020). How central office leaders influence school leaders' decision making: Unpacking power dynamics in two school-based decision-making systems. *Peabody Journal of Education*, *95*(4), 392–407. https://doi.org/10.1080/0161956X.2020.1800175

Authors

CAITLIN C. FARRELL is an associate research professor at the University of Colorado Boulder. Drawing on organizational theory, she specializes in research on the role of evidence in policy and practice.

WILLIAM R. PENUEL is a professor of learning sciences and human development in the School of Education and Institute of Cognitive Science at the University of Colorado Boulder. His research examines the relationship between research and practice and the design and implementation of innovations in science education in schools, districts, and community ecosystems.

KRISTEN DAVIDSON is a research associate at the School of Education at the University of Colorado Boulder. Her work focuses on building the field's capacity to carry out effective research–practice partnerships to improve equitable learning opportunities for all youth.