

What Research Do State Education Agencies Really Need? The Promise and Limitations of State Longitudinal Data Systems

Carrie Conaway

Massachusetts Department of Elementary and Secondary Education

Venessa Keesler

Michigan Department of Education

Nathaniel Schwartz

Tennessee Department of Education

State longitudinal data systems (SLDSs) have created more opportunities than ever before for rigorous research to influence education policy decisions. As state practitioners who play central roles in building and using our states' longitudinal data systems, we are excited about their promise for supporting policymaking and research. Yet, we also recognize that the data in SLDSs will not answer many of our most pressing research questions, nor will the presence of these systems create the meaningful collaboration between researchers and practitioners that we feel is needed to inform our states' policy landscapes. The barriers to the kinds of research we need are mostly unrelated to the promises of SLDSs. We outline the challenges we have experienced in developing research agendas, building our internal capacity for research, and working with external partners, and we identify the research questions we need to answer that are not easily addressed with SLDS data.

Keywords: *state longitudinal data systems, state education agency/agencies, education research, policy research, Massachusetts, Michigan, Tennessee*

As state practitioners who play central roles in building and using our states' longitudinal data systems, we are excited about their promise for supporting policymaking and research. Through our work in Massachusetts, Michigan, and Tennessee, we are intimately familiar with the data problems that these systems are intended to solve. We deal daily with concerns over siloed data and struggle to link disparate data sets from multiple sources. We debate the weight we should give to student assessment results versus longer term student outcomes, for which we are less likely to have available data. And we wrestle

with ways to appropriately transform the data into usable information for state policymakers, district leaders, and the public in a timely fashion. State longitudinal data systems (SLDSs) offer important advances in each of these areas, and we welcome the enthusiasm with which researchers are embracing the possibilities of these tools.

Yet, we are simultaneously skeptical that SLDSs will meaningfully alter the research-practitioner landscape to drive more and better research-based decision making within our organizations. Indeed, when we consider the types of

research that we would like to see and the types of relationships we would like to have with our research partners, the advent of SLDSs does not feel like a game-changer. We can easily imagine a scenario where states invest considerable sums of money to build longitudinal data systems and yet see very little change in the type or quality of research taking place, since the barriers to the kinds of research we as state practitioners need are mostly unrelated to the promises of SLDSs. We focus this piece both on where we see potential in the construction of longitudinal data systems and what more is needed to strengthen education research to support state decision making.

The Purpose and Use of SLDSs: A State Perspective

Thanks to the influx of more than half a billion federal dollars over the last 8 years, nearly every state has invested substantial new resources in building data systems to track key information on their students and educators over time. According to the Institute of Education Sciences (IES, 2011), the purpose of this effort is “to enable State educational agencies to design, develop, and implement statewide, longitudinal data systems to efficiently and accurately manage, analyze, disaggregate, and use individual student data” (p. 3).

Before the introduction of these systems, the reason districts submitted data to the state was to meet federal and state reporting requirements. This was originally accomplished through aggregate collections, with most states transitioning to individual student-level data collections with a unique statewide student identifier sometime in the early to mid-2000s. Some states went on to build systems to collect demographic and work assignment data on educators, as well as information on which courses students were enrolled in and which educators taught those courses, so that students and teachers could be linked. By the time the State Fiscal Stabilization Fund (SFSF) was created as part of the American Reinvestment and Recovery Act in 2009, building more robust data systems was feasible and common enough that the U.S. Congress made working toward a complete statewide longitudinal data system a requirement of accepting SFSF dollars. And by

2011 (the most recent year for which data are available), 44 states had advanced to the point of having a statewide teacher identifier with a teacher–student link, which facilitates sophisticated data analyses at both the student and teacher levels (Data Quality Campaign, 2014).

Under the older data collection approaches, data were reported to the state and were then transformed and sent to the federal government or state legislatures. Not much went back to districts in a format that could help them make educational decisions about their students. Reports were slow to come, often published months after the data were submitted. Their content reflected the needs of the government agencies that requested them rather than of the districts which provided the data. They did not include individual student-level data, let alone real-time access to data on the actual students in an educator’s classroom that day. Furthermore, it was challenging to connect the various data sets to each other and over time because the collections and reporting were cross-sectional. Measuring changes in data within a year or over time took substantial effort and manual linking, leading to underutilization of these data for anything beyond limited reporting.

What gets state education agencies (SEAs) excited about SLDSs, and the focus for the bulk of their investment and development efforts, is the opportunity to provide much more useful, timely information back to district personnel and the public. All three of our states have published basic aggregated reports on student assessment, enrollment, and so forth for years, but through our SLDSs, we have been able to expand to more sophisticated and individualized reporting and analysis. Massachusetts has developed an Early Warning Indicator System (EWIS) to predict the likelihood that each student in Grades 1 to 12 will reach an expected academic goal: proficient or advanced on the Grade 3 English language arts (ELA) state assessment, proficient or advanced on Grade 6 ELA and mathematics, passing all Grade 9 courses, and on-time high school graduation. EWIS information is timely, with reports available prior to the beginning of each school year, and it is highly valued by districts. The EWIS reports are among the most heavily used in Massachusetts’s SLDS reporting tool, with over 10,000 views just between February and June

2013. In Michigan, the SLDS and the associated portal, MiSchoolData.org, provide a number of trend reports to educators and the public, including data and information on postsecondary enrollment and course taking. Similarly, in Tennessee, the state has used Race to the Top funding to begin building a system known as MeasureTN with interactive, longitudinal reports that will go far beyond the traditional school report cards required by No Child Left Behind (NCLB). In the new system, users at all levels—state, district, and the general public—will be able to sort and filter graphs and tables incorporating data that follow students from K–12 through postsecondary education and career and to see the changes immediately reflected in user-friendly data displays. These types of analyses and reports are simply not possible without an SLDS, and the information they yield is crucial for districts to plan effectively for meeting their students' needs and for the public to have easy access to key school and district performance data.

Most states have focused their SLDS programs on making data available to districts and the public in real time, in a format that provides insight into critical district issues. Facilitating research is certainly an additional intended purpose of this investment in SLDSs; indeed, it is listed as a requirement in every associated federal grant program. But from a state perspective, research is a by-product of the SLDS, not its *raison d'être*.

Using SLDSs for Research

The availability of longitudinal student data has generated an outpouring of research that takes advantage of the newly linked data. The national push toward using measures of student growth for educator evaluation and for accountability decisions, for instance, would be impossible without longitudinal data to quantify students' progress over time on state assessments. Similarly, these data have facilitated greater use of experimental and quasi-experimental research designs. In Michigan, longitudinal data are being used to evaluate the impact of the Michigan Merit Curriculum (Center for Local, State, and Urban Policy, Gerald R. Ford School of Public Policy, University of Michigan, 2010), and of the Persistently Lowest Achieving schools designation. Tennessee recently

randomized a set of dual credit offerings across high schools and is examining the effects of these courses on a range of outcomes, including college remediation, entrance, and persistence. The state is also in the midst of a statewide randomized control trial around the state's voluntary pre-K program that follows students across multiple years and settings. Researchers using Massachusetts data have explored questions such as the impact of performance labeling on the likelihood of high school graduation (Papay, Murnane, & Willett, 2010); the impact of a state-sponsored scholarship for high-performing high school students on college enrollment and completion (Cohodes & Goodman, 2013); and the impact of charter school attendance on student achievement on state assessments (Abdulkadiroglu, Angrist, Dynarski, Kane, & Pathak, 2011), all using research designs that can credibly establish causality. These research questions are answered well by a longitudinal data system, because changes happened at a point in time to an easily defined group and are experienced as population shifts over time. Much can be learned from these types of analyses that is relevant for policymaking.

Although individual impact studies have their place, these studies by themselves are not particularly responsive to the way business gets done around our departments. Rarely do SEAs have the time or resources to start a research project from scratch to get an answer in time for a policy decision. Most often, for a combination of reasons—timing, funding, politics, and sometimes prior research findings—our SEAs commit to a broad initiative or theory of action and then create a set of policies and programs to support that goal. At that point, we want to know not only whether the initiative is working but also how it might work better, what we can learn from the variation across sites, whether there are specific points of leverage that would increase the policy's influence, and how we can take advantage of these instances within the relatively short period needed to establish political buy-in.

This does not mean that rigorous evaluation of the impact of an overall initiative has no place, but it suggests that such evaluations need to be combined with a series of studies at all points along what the IES and the National Science Foundation (NSF) have described as the research pipeline—from foundational and design-focused

to effectiveness and scale-up studies—that can extend the knowledge base beyond whether a program works or not (IES and NSF, 2013). To address this challenge, our three states have pursued a multi-faceted approach: designing a strategic and proactive research agenda tied to our agencies’ policy priorities, increasing state internal capacity for research, and building managed portfolios of partnerships with external researchers.

Research Agendas in State Agencies

SEAs have historically afforded little attention to research or research support (Massell, Goertz, & Barnes, 2012). As a result, studies have tended to be defined by external researchers who approach states with requests for data access. When each of us arrived in our positions, the existing research partnerships had been built through this process. The projects that our states were involved in felt disconnected and scatter-shot, not because of any data-related issues but because they each existed in isolation. When seen from afar, in concert with other research that was taking place across the country, these studies might be viewed as building a greater knowledge base. But within our agencies, the projects tended to exist on the sidelines: nice to take part in, but separate from the central work of program development and improvement.

Changing that dynamic has required each of us to take an active role in defining our state’s research agenda. A research agenda for an SEA is fundamentally different from that of an individual researcher or even a think tank or research shop. As one of us has written elsewhere, “In the public sphere, nothing dooms research faster than irrelevance” (Conaway, 2013). For an SEA, data and information need to be relevant to the current policy agenda and available when necessary for a policy decision. Thus, an SEA’s research agenda, if it is to be useful, must be tied closely to the strategic policy priorities of the organization and must be able to be expanded and refined on a regular and unexpected basis.

In our states, we all have research agendas that connect our research priorities to our state’s educational priorities, but we all must adapt when new priorities emerge. For instance, Massachusetts’s research priorities are one and the same as the agency’s strategic objectives, and they are defined

broadly, for example, strengthening curriculum and instruction and improving educator effectiveness (Massachusetts Department of Elementary and Secondary Education, 2014). The state has defined a number of specific research projects to advance the agency’s knowledge in these priority areas. But it also had to shift quickly to accommodate new questions when the state decided to implement a major initiative, affecting more than 25,000 educators statewide, to strengthen core academic teachers’ ability to teach English language learners. This initiative was too important to simply ignore, and evaluation work had to be designed on short notice and on a shoestring to ensure that the agency gathered the knowledge it needed to inform program implementation and measure outcomes. Michigan and Tennessee both follow similar processes to define and modify their research agendas to reflect the agency’s current and emerging priority policy areas. In all three states, the research agenda becomes a public document used to prioritize internal work and to manage external requests.

We must then make choices about how to enact this agenda. Figure 1 shows how we conceptualize the relationships between internal capacity and external partners in meeting our research needs. Internal resources tend to be better positioned for rapid-response, targeted, descriptive analyses. External partners are better positioned for investigating long-term questions or research agendas, particularly those focused on causal analyses, and for situations where an independent voice is needed. The figure highlights our key argument: Many of our research needs are currently met neither by our available internal and external resources nor by the information in SLDSs. We expand on these points below.

State Capacity for Research

To enact our research agendas, we have first focused attention on building systems, skills, and knowledge within our own agencies to facilitate research. A major part of this work is translating data from SLDSs into research-ready data files, for longitudinal data collected cross-sectionally does not equate to a usable longitudinal data set. Hundreds of decisions must be made: how to reconcile differing values over time (Johnny was

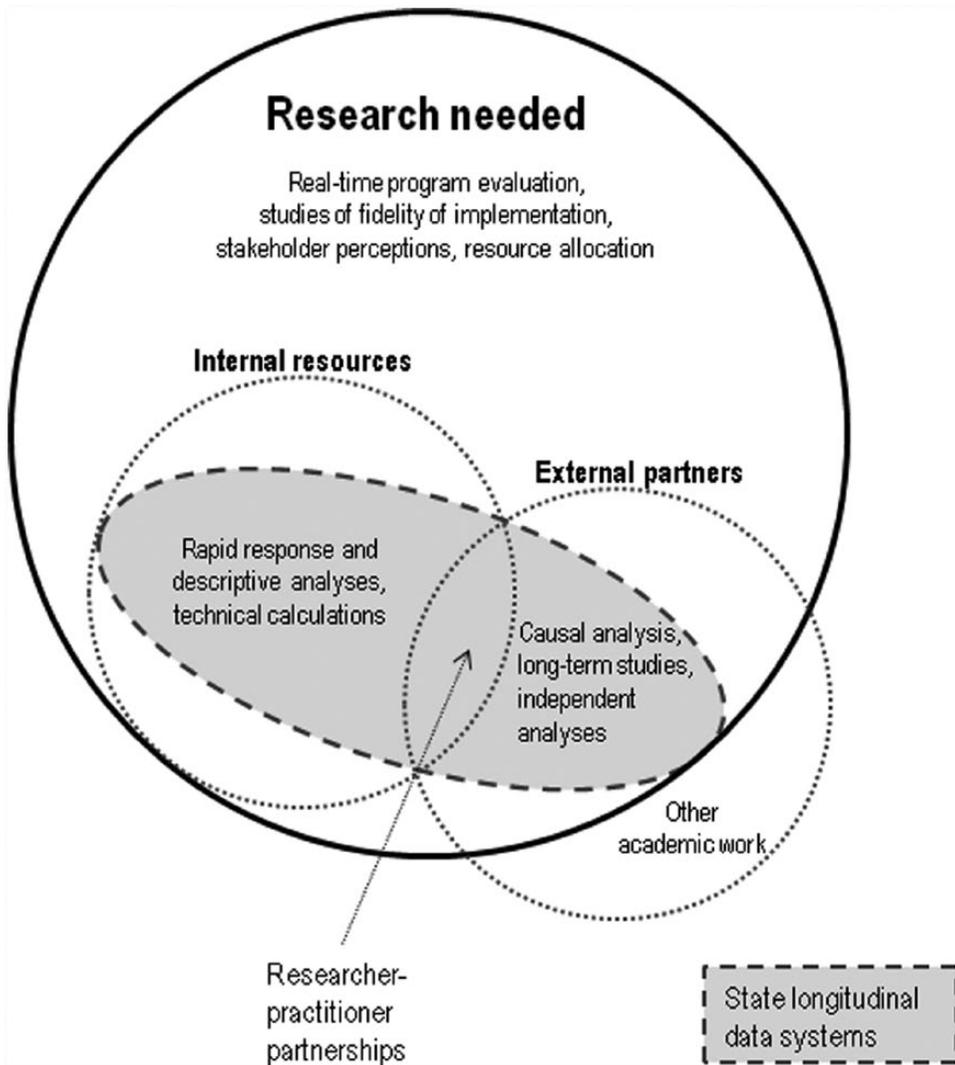


FIGURE 1. *Education policy research from the state perspective*

coded as Black, then Multi-race); where to attribute a student who has changed districts (by headcount to the district attended on a particular date, a proportional accounting for time in each location, or other methods), and what to do about changes in administrative data collection rules and codes over time, just to name a few. In some cases, it makes sense for the state to standardize the data provided to researchers, whereas in others, it makes more sense to allow individual researchers to make their own choices. We cannot just build one data set and provide it to everyone. We have also built systems to prioritize which research projects get access to which types

of data. No SEA has unlimited resources, so if researchers are studying a topic outside the state's priorities, they may be directed to publicly available or standard research files (vs. a customized data pull) or may be told that we cannot work with them at this time.

Furthermore, we have concentrated on increasing the skills of the staff within the agencies to understand research and work with researchers. The individuals who are building state data systems and doing the painstaking work involved in extracting, transferring, and loading data; building data structures; and maintaining documentation are rarely researchers. They are skilled

technicians, but they do not necessarily have the research skills or training necessary to determine how to use the longitudinal data they manage for answering research questions. Many SEAs also employ analysts who have some research skills, but the intense operational demands of accountability determinations, public reporting, and so forth mean that research becomes the type of high importance, low urgency work that never gets the attention it deserves.

One response to this has been participation in the Strategic Data Project (SDP), an initiative run by the Harvard Graduate School of Education to recruit, train, and support cohorts of data fellows placed in states and districts to increase capacity and the ability to use data diagnostically. Our states have all partnered with SDP and have had fellows placed in our agencies. Participating in this program has helped us demonstrate the value to the agency of having internal staff with strong research skills. Other approaches are to find funding for additional full-time analyst positions and to bring in student help through internships, credit programs with local universities, and graduate research associateships. We have all also focused effort on increasing the skill of our existing staff. Massachusetts, for example, has offered a series of trainings to its analysts on topics related to causal analysis, including randomized controlled trials, regression discontinuity design, instrumental variables, difference in differences, and propensity score matching. By placing staff with research skills within our agencies and by building our staff's statistical analysis skills, we have been able to do more research work internally, under our own direction and on the time frame needed for policymaking.

This linked set of efforts has paid off almost immediately in each of our agencies in terms of what we have been able to produce and what our colleagues have come to expect from internal research. In Tennessee, each division within the Department of Education works with the Office of Research and Policy to build out a calendar of major upcoming decisions and rapid turnaround studies that might inform these decisions. These mostly descriptive analyses in turn have affected everything from the way the state selected teacher coaches to lead Common Core trainings to the design for launching a pilot to increase Advanced Placement pass rates. Several of the analyses

have then spun off into public white papers that have the potential to shift the statewide conversation about data and research and also demonstrate the state's commitment to the process of data-driven decision making (e.g., Office of Research and Policy, 2013, 2014a, 2014b).

External Research Partnerships

Even after working to strengthen our internal research capacity, a gap still remains for each of us between our agency's capacity and our research needs. To bridge this gap, we have each sought out external partnerships to help us enact our research agendas. External partners can provide important rigor in the approach to evaluation, as well as political independence that allows them to state clearly what works and what does not. They are also helpful when we have a need for in-depth qualitative analysis and data collection, survey research, or other forms of data gathering. These partnerships represent both a critical piece of our research strategy and a central frustration for each of us, and it is our experience with these partnerships that creates our skepticism about the ability of the SLDS alone to fundamentally alter the landscape of state-level education research.

All three of us have seen great benefits from partnering with external researchers. In Michigan, for instance, the state formed the Michigan Consortium for Educational Research (MCER), a multi-year, IES-funded partnership with the University of Michigan and Michigan State University designed to study the impact of the Michigan Merit Curriculum and the Michigan Promise Scholarship. Researchers at the University of Michigan and Michigan State University separately reached out to the Michigan Department of Education (MDE) for support on a grant application. Officials at MDE requested that the two universities partner together, to have a single Michigan-based proposal that met statewide needs. Principal investigators (PIs) from both the universities and the state were selected and remain as PIs on the grant. The team has worked together on the main research questions identified in the grant, other analyses the state needs, data quality and processing issues, and procedures and policies for providing data to an ever-growing team of graduate students and

professors. MCER has also assisted the state in gathering and analyzing postsecondary transition data, helping Michigan answer the question of what happens to their students after they graduate from high school. It is currently engaged in drawing a representative sample of high schools and gathering transcript data from those high schools. This will allow for better estimation of the implementation of the Michigan Merit Curriculum and how course taking has or has not changed over time; it also gives the state a representative sample of high schools that the state hopes to use for other purposes. MCER has also created the most comprehensive longitudinally linked research-ready file structure, one that the state is using to understand how to build these files for other researchers.

Yet even when partnerships are as successful and productive as MCER, they still create challenges for state agencies. One issue is finding common ground between the personal research agendas of external researchers and a state's research priorities. Academic institutions privilege researchers who have built a continuous stream of research in a specific topic and, particularly among quantitative researchers, have used causal analysis techniques to answer their research questions. But states' research needs tend to shift quickly and often can be satisfied with descriptive analysis. Given the priorities of the academic enterprise, it is difficult to find researchers who are willing to study what the state needs, versus what the researcher wants to study.

Another difficulty is that research proposals, when they arrive on our desks often tend to feel like opportunistic requests that hope to take advantage of a particular data set that the state collects or a program that fits within the researcher's agenda. Applicants rarely speak to us about how the project will build out our understanding of a particular reform strategy in our state, but rather about the way the project will add to the far more general knowledge base of researchers across the country. As a result, our states have had to develop a new orientation toward research management that involves moving beyond individual research projects to a coherent research portfolio including multiple studies on key initiatives. Our goal is to create a local knowledge base that directly addresses the design and implementation of our highest priority initiatives.

To make this concrete, one of the major areas of interest in Tennessee lies in using teacher evaluation as a tool for instructional improvement. In 2011–2012, the state implemented a statewide system of teacher evaluation that included multiple measures of teacher practice and student achievement. Tennessee's decision to move in this direction was linked to prior research such as Taylor and Tyler's work in Cincinnati and the Measures of Effective Teaching project sponsored by the Bill & Melinda Gates Foundation (e.g., Kane & Staiger, 2012; Taylor & Tyler, 2012), but much of the operational work felt like uncharted territory, and the initiative also called for a system that would evolve over time through research-driven improvement.

Over the past year, Tennessee has worked to build a set of partnerships that it hopes will act in concert to provide powerful direction for its work. In addition to a substantial set of internal analyses on the implementation of teacher evaluation, the state has enlisted John Tyler, Eric Taylor, and John Papay at Brown and Stanford to put together a series of randomized experiments around teacher collaboration and mentoring based on evaluation results. It has joined with RAND to plan a study that looks qualitatively at districts using alternate teacher observation systems and compares the kinds of support these systems provide to different teacher subgroups. And the state is conducting multiple streams of research in coordination with partners at the Tennessee Consortium on Research, Evaluation, and Development (TNCRED) at Vanderbilt University to understand the ways that evaluation is currently being used in schools and districts across the state. TNCRED both administers an annual survey to all teachers and administrators around teacher evaluation and feedback and is putting together a series of qualitative case studies describing the different methods that districts are using to integrate teacher evaluation with professional development initiatives.

These research projects have considerable potential, and they are each likely to advance the field in their own right. Yet, if they proceed as individual studies, the individual researchers are likely to make different decisions about data and analysis that make it difficult to look across studies and set broader policy direction. This in turn promotes confusion among policymakers who

are forced to try to make sense of a series of separate studies rather than a linked series of recommendations. As a result, such studies are likely to offer practitioners a single nugget of information while changing little about broader agency strategy. Tennessee is currently engaged in an attempt with Vanderbilt to shift this paradigm by creating a research consortium built around synthesizing individual studies into coherent research strands. The aim is to create a hub for state-level research that can break the trend toward partnerships that result only in one-shot studies.

This speaks to a central point about the communication of research findings. Study results are only useful to the extent that they can be placed within a broader contextual framework and provide an integrated set of recommendations to directly inform state and district policies. We look for research partners who are willing to think of their research as a developing story, where the researcher's role is partly to help those within the state agency make sense of the results and come up with next steps. The leaders of the Chicago Consortium for School Research (CCSR) have referred to this issue as one of "building knowledge of core problems across time and across studies":

Ultimately, research reports take time to read and absorb, even with painstaking efforts to make them accessible. The window of time in which people pay attention to the findings of a research study can be very short. Yet it requires considerable time for educators to grapple with the importance of the findings, their potential implications, and what these implications mean for their day-to-day work. CCSR seeks to extend the time that the results of research are considered, first by building coherence across studies and second by developing indicators that keep those ideas on the agenda. (Roderick, Easton, & Sebring, 2009, p. 10)

The format and timing through which findings are shared also matter considerably. In the independent evaluations it commissions, Massachusetts has moved away from the traditional annual tome: a single massive document that includes all the findings from a year's worth of research and that typically comes well after the year has ended. Too often these types of reports did not have the influence they could have if the findings had come in a more accessible, timely format. Instead the state now requires

researchers to turn around preliminary findings as quickly after data collection as possible: a few weeks after a survey is conducted, perhaps a bit longer for qualitative work. This rapidly injects the findings into the agency's discourse and allows program staff to use them immediately to improve the quality of program implementation. The researchers often produce a more detailed final report later, but even then, the deliverable is now much more likely to take the form of a summary intended for public consumption. Similarly, Massachusetts now requires independent academic researchers to produce a two-page summary of findings for all papers using state data. This brief is then distributed to the state's superintendents and principals via the commissioner's weekly email to the field and is made available on the state's public Website, expanding the work's reach and influence.

A final challenge of external partnerships is the funding landscape, both at the federal and state levels. Most states (ours included) do not have a dedicated research budget with which we can hire researchers. This means we need to partner with external researchers and then support their search for other funding sources. Many of the IES research grant programs ask researchers to find state or local agencies to partner with on questions of interest or prioritize applications with support from these agencies. For many researchers, this amounts to trying to retrofit the state's needs into their own, so that they can continue their pre-existing research agenda. Our states are trying to avoid this approach, instead recruiting researchers within our priority policy areas, defining our projects and research needs with a good deal of specificity, and supporting researchers' applications for funding. But, particularly with university-based partners, it can still feel as if the partnership is happening to us rather than with us. Researchers can decline to partner; they can decline to submit for funding. They can indicate they want to partner but then be hampered by the other demands on their time and tenure clocks and not provide results in a timely fashion. We must continuously negotiate with researchers to develop partnerships that truly address our needs.

Another potential source of funding for research is the foundation community. However, philanthropic organizations do not generally wish

to build infrastructure within SEAs and often require or desire a match of funds from the organization. Foundations are also agenda driven: If the research questions of the foundation match the SEA's research and information needs, then the foundation may make a good partner to support the work, but when this is not the case, SEAs are again left without a nimble funding source. There are certainly many examples of foundations investing in research agendas—the Measures of Effective Teaching project is just one—but that is not the same as an SEA identifying research that relates to priority policy areas and crafting their own responsive research agendas.

We applaud the IES for recognizing this reality with its recent turn toward funding continuous improvement research. We appreciate the opportunity to build relationships with researchers on key areas of interest that may develop into bigger evaluations or studies. But the largest amount of money in IES competitions is still reserved for large-scale impact evaluations, not the more formative, iterative work that states also desire and that we cannot conduct without external support. One suggestion to IES and other funding agencies would be to make block grants available to states, which states can then use to identify and form funded research partnerships directly connected to the needs of the state. The onus would then be on the state to convince IES of the appropriateness of the research and remove the difficulty of an individual researcher serving two masters—their funding source and the state.

Additional Limitations

Several other limitations of research with SLDSs are, in our view, not sufficiently well recognized in the research community. First, state data systems were designed for data reporting, not social science research. It is common to encounter variables identified in the research literature as key indicators that are not available in administrative data sets, as we typically limit collections only to the information required for federal or state reporting. Few states, for instance, collect data on a student's mother's education level, yet scores of studies indicate that this is one of the most important inputs into student performance (cf. Duncan, Brooks-Gunn, &

Klebanov, 1994). Other key indicators, such as student motivation or resilience, are rarely captured even in social science research databases. Although some of these variables could potentially be collected from districts, it is complicated and expensive to add data elements to district data collections, and states often experience significant pushback from districts on new data collection requirements. We cannot fundamentally change our systems to accommodate researchers' individual agendas.

Second, we cannot understate the importance of the privacy and data security concerns about the use of individual-level student data in the conduct of this work. These are administrative data collected by the state for federal and state reporting and accountability purposes, not for research. There are deep tensions in the public right now over what we are collecting, how we are collecting it, and how it is used. Violations of this trust—perceived or real—will be detrimental to our ability to continue to collect, report, and analyze these data. The Family Education Rights and Privacy Act (FERPA; the federal law governing student data confidentiality) does include a researcher exemption to allow for data to be used to evaluate educational programs. But this does not mean that every research request for student-level data can, must, or should be granted. And the FERPA regulations around what constitutes personally identifiable information mean that we cannot simply remove students' names from our data set and hand over the files to researchers. States must be extraordinarily careful to protect the use of these data in accordance with the FERPA regulations to safeguard students' privacy and their own ability to conduct research as needed to improve their programs.

We would also like to recognize the fact that commissioning research on priority policy areas—particularly when those topics represent cornerstones of an individual's policy agenda—takes courage and is not without risk for state superintendents and commissioners. There is always the very real possibility that the research will yield an answer opposite the one that was anticipated, whether because the policy was actually not effective or because its costs were too high given the benefits. Conversely, some programs have strong political backing and are unlikely to be eliminated even in the face of

strong evidence of their ineffectiveness; research on these programs has only risks, not benefits, from the state agency perspective. Even in states such as ours that are committed to understanding what works and why, sometimes there are questions to which, for good reason, we do not want to know the answer.

State Research Needs: Beyond SLDS

Even if we could address all the challenges of aligning research agendas, internal capacity, and external partnerships to meet our research needs, and even if we had no missing data, privacy, or political considerations to worry about, we would still have difficulty enacting our agendas with SLDSs alone. The data in SLDSs come nowhere close to answering all of a state agency's research questions. We also want to know how well our programs are being implemented in the field, what our stakeholders think of them, and whether their benefit outweighs their cost. These types of questions require access to data not typically contained in SLDS systems, and often data that are difficult or time-consuming to collect. Some detailed examples of research issues important to states that are not addressed well with SLDS data include the following.

Technical Calculations

States are frequently called upon to identify the lowest performing districts or schools in the state, or to develop an index or ranking system for a set of indicators, or to identify which districts are most similar to a particular focal district. For instance, state applications for waivers from NCLB involve complex calculations to create the indices used for assigning schools and districts into various accountability designations. These types of analyses often use longitudinal student data, but the research questions involved are fundamentally methodological rather than longitudinal. Considerable evidence shows that these systems stretch state data capacity to the breaking point, yet researchers rarely offer their expertise in these areas (Grew & Sheldrake, 2013). We need assistance from the research community to ensure that we are developing fair systems of comparison and that we avoid generating unintended consequences.

Accountability Decisions

High-stakes policy decisions are rarely made solely on the basis of SLDS data. For instance, as our states determine which of our lowest performing schools have made enough progress to be removed from the underperforming label, we consider not only quantitative, longitudinal data about the performance of students in the school over time and relative to the state but also qualitative information gathered from sources such as district accountability reviews and school site visits to measure whether the district and schools have the systems in place to continue to support improvement. These data are collected systematically but not coded into the SLDS.

Real-Time Program Evaluation

Program managers are eager to improve their offerings and thus seek real-time feedback on program performance. They often use logic models to define expected intermediate outcomes from their programs—for instance, changes in teacher practice that might be expected to eventually yield improvements in student outcomes, or results from district-based benchmark or formative assessments that are not relayed to state data systems. Although these approaches lend themselves to research, it is research of a different nature than the type typically conducted with state administrative data.

Fidelity and Variability of Implementation

States often want to capture not just the average impact of a program or policy but also the range of outcomes achieved across program sites or subgroups and the reasons why outcomes varied. What factors drove variation in the quality of implementation? What characteristics describe the sites that achieved the best outcomes, and how much do they differ from those that fared the worst? This helps states improve their programming and deliver better results over time. Yet many quantitative studies aim only at measuring the average impact, not the range, and the qualitative questions around variation in implementation take the difficult and costly work of observations, interviews, and focus groups to answer, not SLDS data.

Stakeholder Perceptions

A key element in any policy decision is the likely concerns of stakeholders and whether and how those concerns could be mitigated. Massachusetts, for instance, has created a variety of research-based feedback loops to gather input from the field for policymaking, including customer satisfaction surveys of superintendents and principals, a biannual statewide teacher survey, data collection on implementation of major initiatives, and commissioned program evaluations. Although this type of information can sometimes be gathered quantitatively and longitudinally, it is not a part of SLDSs, and often the most valuable information is qualitative rather than quantitative.

Resource Allocation

Some of the most important decisions state agencies make are how to allocate their scarce resources. For example, should a state invest in providing direct services to its lowest performing districts, or should it instead provide those districts with grants? Should it expand after-school and out-of-school programming for the students with the greatest needs, or should it expand the school day for all students? Answering these questions requires understanding not only the potential impact of the program or policy options but also their associated costs and the counterfactual of what would happen in the absence of a change. But few studies take on these challenging issues. Indeed, of the 782 studies that met What Works Clearinghouse evidence standards with or without reservations as of July 2014, none of them included “cost” as a keyword. Only 31 studies out of the over 10,000 in the database used that keyword at all, and most were deemed ineligible for review (IES, 2014).

Summary

In all of these areas, we can and do engage with researchers to help inform our work. But much of the research community continues to focus on what it can do with the already collected administrative data available in the SLDS, rather than also helping us to collect new sources of data to shed light where the SLDS cannot. When we commission our own research, of course, we can insist that these broader data sources be developed and incorporated. But when we

partner with researchers and must negotiate a shared research agenda that all agree to, this type of work is too often an afterthought.

Conclusion

Our three states have sought to increase research capacity and use our longitudinal data more efficiently and strategically for policymaking. We have developed research agendas tied to our agencies’ priorities. We have strengthened our agencies’ ability to handle research requests and to do research internally. And we have built strong partnerships with external organizations to help us fill in the gaps between our capacity and our needs. We all value working with researchers and recognize that we cannot do our work well without them.

We look forward to exploring new models for research partnerships, enhanced by the ever-growing data sets contained within SLDSs but grounded in the needs of state agencies. Successful models would adapt research agendas to state strategic priorities, find new ways to synthesize information across individual studies to create actionable recommendations, draw upon new sources of flexible funding to allow for both long-term, causal studies, and diagnostic analysis, and consistently integrate cost–benefit analyses and questions of resource allocation into their core findings.

Finding ways to meet these needs is a challenge on par with the work we have already undertaken to build SLDSs and make them more usable for research. But it is this hard work of building research skills within our organizations and nurturing mutually beneficial partnerships with the research community that will ultimately help us reach our goal: to increase the use of evidence—from SLDSs or elsewhere—in the policymaking process.

Authors’ Note

The authors are listed in alphabetical order. All contributed equally to this work. All opinions are our own and do not necessarily reflect the positions of the Massachusetts Department of Elementary and Secondary Education, the Michigan Department of Education, or the Tennessee Department of Education.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Abdulkadiroglu, A., Angrist, J., Dynarski, S., Kane, T., & Pathak, P. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *Quarterly Journal of Economics*, *126*, 699–748.
- Center for Local, State, and Urban Policy, Gerald R. Ford School of Public Policy, University of Michigan. (2010). *Mandating merit: Assessing the implementation of the Michigan merit curriculum*. Retrieved from <http://closup.umich.edu/files/pr-13-michigan-merit-curriculum.pdf>
- Cohodes, S., & Goodman, J. (2013). *Merit aid, college quality, and college completion: Massachusetts' Adams scholarship as an in-kind subsidy* (Working Paper Series RWP13-005). Cambridge, MA: Harvard Kennedy School.
- Conaway, C. L. (2013). The problem with briefs, in brief. *Education Finance and Policy*, *8*, 287–299.
- Data Quality Campaign. (2014). *State analysis by essential element: Element 5*. Retrieved from <http://www.dataqualitycampaign.org/your-states-progress/10-essential-elements?element=five>
- Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. (1994). Economic deprivation and early childhood development. *Child Development*, *65*, 296–318.
- Grew, J. R., & Sheldrake, W. J. (2013). *Examination of Indiana's A to F school accountability model* (Report to the Legislature). Retrieved from http://www.in.gov/legislative/pdf/Accountability_Model_Exam.pdf
- Institute of Education Sciences. (2011). *Grants for statewide, longitudinal data systems*. Retrieved from http://ies.ed.gov/funding/pdf/2012_84372.pdf
- Institute of Education Sciences. (2014). *What Works Clearinghouse*. Retrieved from <http://ies.ed.gov/ncee/www/>
- Institute of Education Sciences and National Science Foundation. (2013). *Common guidelines for education research and development*. Retrieved from <http://ies.ed.gov/pdf/CommonGuidelines.pdf>
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high quality observations with student surveys and achievement gains*. Seattle, WA: Bill & Melinda Gates Foundation.
- Massachusetts Department of Elementary and Secondary Education. (2014). *ESE Research Agenda for 2012–14*. Retrieved from <http://www.doe.mass.edu/research/agenda.html>
- Massell, D., Goertz, M. E., & Barnes, C. A. (2012). State education agencies' acquisition and use of research knowledge for school improvement. *Peabody Journal of Education*, *87*, 609–626.
- Office of Research and Policy. (2013). *The impact of the 2012 TNCore math training on teaching practices and effectiveness*. Tennessee Department of Education. Retrieved from http://www.tn.gov/education/data/doc/impact_of_TNCore_Training.pdf
- Office of Research and Policy. (2014a). *Advanced placement strategy: A framework for identifying school-level barriers to AP success*. Tennessee Department of Education. Retrieved from http://www.tn.gov/education/data/doc/AP_report.pdf
- Office of Research and Policy. (2014b). *Teacher retention in Tennessee: Are we keeping our best teachers?* Tennessee Department of Education. Retrieved from http://www.tn.gov/education/data/doc/retention_report.pdf
- Papay, J. P., Murnane, R. J., & Willett, J. B. (2010). The consequences of high school exit examinations for low-performing urban students: Evidence from Massachusetts. *Educational Evaluation and Policy Analysis*, *32*, 5–23.
- Roderick, M., Easton, J., & Sebring, P. (2009). *The consortium on Chicago school research: A new model for the role of research in supporting urban school reform*. Retrieved from <https://ccsr.uchicago.edu/publications/ccsr-new-model-role-research-supporting-urban-school-reform>
- Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review*, *102*, 3628–3651.

Authors

CARRIE CONAWAY is the associate commissioner for planning, research, and delivery systems for the Massachusetts Department of Elementary and Secondary Education. She oversees the agency's work in research, program evaluation, strategic planning and implementation, district accountability, data collection, and data reporting. She holds a bachelor's degree in sociology from Oberlin College; a master's degree in policy analysis and labor policy from the Humphrey Institute of Public Affairs, University of Minnesota; and a master's degree in sociology and social policy from Harvard University.

VENESSA KEESLER is a deputy superintendent of accountability services at the Michigan Department of Education (MDE). The accountability services division includes the offices that oversee K–12 assessment, accountability, data systems, and strategic research, as well as teacher licensure and certification and educator preparation institution accountability. She previously

served as the director of the Office of Evaluation, Strategic Research, and Accountability at the MDE and as a classroom teacher in Massachusetts, teaching seventh-grade social studies. She holds a bachelor's degree in sociology from Harvard University and a PhD in measurement and quantitative methods from Michigan State University.

NATHANIEL SCHWARTZ is the director of research and policy at the Tennessee Department of Education.

His primary research interests include research-practice partnerships, large-scale instructional intervention, and teacher evaluation. He received his PhD and MPP from the University of Michigan and his BA from Harvard University.

Manuscript received September 30, 2013

Revision received August 5, 2014

Accepted January 23, 2015