

# Better Policy Through Research: Pursuing High-Impact Research in State Education Agencies

Carrie Conaway  
Massachusetts Department of Elementary  
and Secondary Education

November 2015

© 2015 Edvance Research, Inc.

In spring 2014, my agency, the Massachusetts Department of Elementary and Secondary Education, was at a turning point.<sup>1</sup> We were nearing the end of our first year rolling out our new **educator evaluation framework** for teachers and principals, with about half our districts having gone through an initial phase and the other half set to begin. While we had general statewide agreement that the old evaluation system didn't do a good enough job differentiating strong, capable, and weak performance in teachers and principals—or identifying where educators could improve—the first year of implementation had been challenging. Districts struggled with training staff on the new model's components, finding time for evaluators to conduct and document classroom observations, and identifying appropriate student growth measures. Anecdotal concerns surfaced that some teachers were more likely to get a positive evaluation simply due to the nature of the grade levels, subjects, or students they taught rather than their actual performance, thus creating a perception of unfairness. We planned to use a May 2014 statewide meeting on educator evaluation to acknowledge the new evaluation system's challenges and districts' good-faith efforts to implement the new system, but also to rebut misperceptions where we had evidence that they were inaccurate.

Fortunately, we had invested in research and evaluation projects to inform our implementation of the evaluation initiative: that meant we had solid evidence to share while we addressed districts' challenges head on. We had examined the statewide distribution of performance ratings and compared them to student growth measures on our state assessment to understand how well the educator practice ratings aligned with the measures of educator impact on student growth. We had contracted for a large-scale program evaluation of the new system, including a representative sample survey of teachers and principals, as well as case studies and focus groups in a set of representative districts carrying out the new framework. With the survey conducted just a few weeks prior to the May meeting, we pushed for preliminary survey data so we could share new information on how educators perceived this major state initiative in time for the statewide event.

As a result, my agency's commissioner was able to share real-time data with the field in his **keynote address**. From our internal analyses, he could show that the ratings distribution was bell curve-shaped, with the majority of educators rated “proficient” on the summative performance rating, a measure of educator practice. This sharply contrasted with other states, where most educators rated in the highest performance level; our data demonstrated that evaluators were taking seriously the goal of better differentiating teaching performance. From the alignment analysis, he could show that higher professional practice ratings were associated with higher impact on student learning, supporting the rating system's validity. The external evaluation, meanwhile, found that most educators felt they had received sufficient training

on the system as well as timely and helpful feedback from their evaluators. Importantly, it also showed that while many educators had concerns that the *system* overall was unfair, nearly 90 percent of those evaluated felt that *their own* evaluation had been fair. The commissioner used these research findings to acknowledge the work our districts had done to implement the system with fidelity, to show that the new evaluation system had a statewide impact, and to begin to alleviate educators' anxiety about the system. Our research investment in this key state initiative positioned us to start to shift the state dialogue about the new evaluation system and to better support our statewide implementation.

State education leaders face many choices about whether and how to invest in research and how to structure these inquiries to best support the state's policy strategy. This essay will focus on where to begin: why research is valuable, what types of research activities state agencies might focus on, how to ensure the research is high impact, and how much it might cost.

### WHY RESEARCH IS VALUABLE

Our story is just one example of how research has informed our state's implementation of educator evaluation. We also relied heavily on research to develop our initiative in the first place. We combined the best evidence from existing research literature with stakeholder feedback to establish the program's broad framework. Educators receive a professional practice rating of "exemplary," "proficient," "needs improvement," or "unsatisfactory," undergirded by a five-step cycle of self-reflection and goal setting. They also receive a separate rating that gauges their impact on student learning, which involves multiple measures of student achievement and growth over at least two years of data. Research also informed many decisions on the detailed regulatory requirements. The inclusion of staff and student feedback as a required data source for the professional practice rating was influenced by the Measures of Effective Teaching study, which showed these data provide insight on an important dimension of educators' practice that other data sources fail to fully capture.<sup>2</sup> And our decision not to require a set percentage of an educator's rating to come from their impact on student learning was based in part on researchers' cautions that these estimates are often imprecise at the teacher level.<sup>3</sup> Findings from our commissioned program evaluation have shaped our technical assistance to districts: the state's 2014–15 school year priorities on district-determined measures of student growth, ratings calibration across educators, and human resources practices all stemmed from **evaluation findings** indicating that these were areas where educators felt they needed more support.<sup>4</sup> We have also deployed our internal analytical staff to monitor implementation, identify problem areas early, and redirect resources and assistance as needed.

But our agency’s research investment goes well beyond educator evaluation. Other recent projects have examined implementation of our 2010 curriculum frameworks in English language arts and mathematics, our school and district turnaround programming and supports, placement patterns of low-income students into special-education programs, the impact of the state’s college merit scholarship program, and the impact of career and technical education on students’ academic outcomes. Some of our research has been done internally, some by university researchers, and some by professional research and evaluation organizations. In all cases, findings are shared with the field and incorporated into the lifeblood of the agency’s decision-making processes.

Why have we spent so much time and effort on research? Because it makes our work better. It demonstrates to the field our commitment to gather and weigh their input and perspective as we implement major initiatives—and it does so in a systematic and democratic way so we do not hear only from those with the influence and power to communicate directly with state leaders. As our opening example demonstrates, research allows us to reframe public discourse about our initiatives and understand where the field faces challenges and needs more support. It helps us stay abreast of implementation challenges as they arise so we can adjust course as needed. It informs our strategic planning, both early on to help identify priorities and later to monitor rollout and outcomes. And it allows us to prioritize our financial and staff resources, determining what technical assistance is most valuable to the field and what resources can shift over time to programs with the greatest evidence of impact. Through all these mechanisms, research helps us continuously improve our program implementation to maximize results for our state’s educators and students.

## WHAT RESEARCH CAN DO

Research encompasses a broad range of activities and inquiries that state education agencies can pursue.<sup>5</sup> But all of them help support continuous improvement of policy implementation using various lenses, methodological approaches, and levels of sophistication.

### Among the options are:

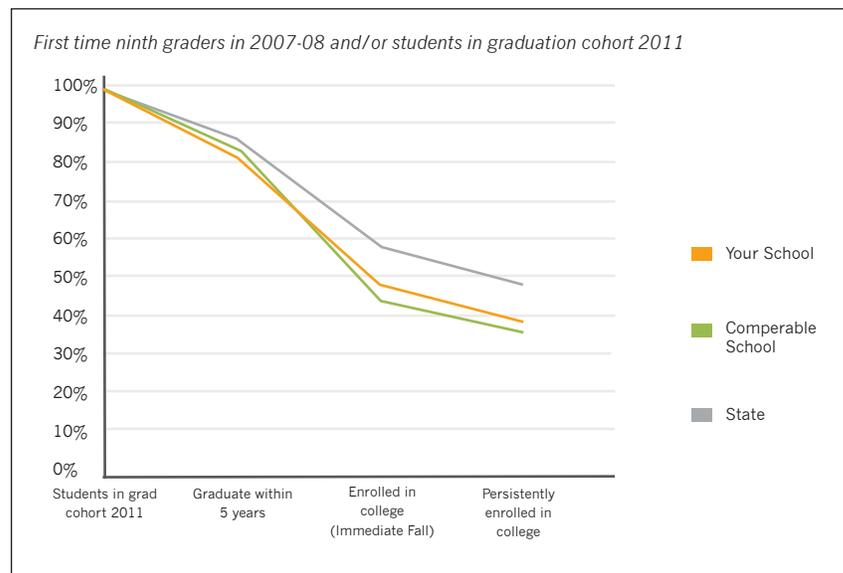
- **Literature reviews.** Particularly valuable in the early phases of policy development or at key decision points in a program’s implementation strategy, literature reviews let us examine what is already known about a particular policy approach’s likely impact. This can help narrow policy options to those most likely to succeed and can flag potential blind spots or opportunities to expand or refine the program. The federal government offers several useful resources: the **What Works Clearinghouse** reviews recent research on effective educational practices, the **Regional Education Laboratory** program has a reference desk service to provide literature

reviews to education stakeholders, including state education agency staff, and the **Comprehensive Center** program provides technical assistance to states focused both regionally and on key education policy topics.

- **Descriptive analyses.** Simple descriptive analyses—percentages, averages, medians, and the like—are the bread and butter of many policy analyses, with good reason. They are straightforward to calculate and widely understood by stakeholders. For years in Massachusetts, there was widespread knowledge of and concern over our public colleges' high remediation rates. But in 2008, we established for the first time that 37 percent of our state's public high school graduates who enrolled in our public colleges were taking at least one remedial course in their first semester.<sup>6</sup> This specific, concrete link between students moving from the public K–12 system to the public college system garnered far greater attention to and concern about our students' college and career readiness than the general statewide data previously available. It ultimately resulted in changes to the state's four-year college admissions requirements and a greater focus on the rigor of the high school curriculum. And it gave us a yardstick against which to measure our progress toward reducing remediation needs.
- **Comparative analyses.** Analyses that allow states, districts, and schools to compare themselves to one another and identify their strengths and areas for improvement are even more powerful. The Massachusetts research office produces a set of **District Analysis and Review Tools (DARTs)** that use enrollment and demographic data to identify the ten most similar districts for each district statewide, then display a range of simple charts and tables on student performance and other outcomes for these comparison districts. Identifying relevant comparison groups at the state level is particularly important, as it is hard for districts to know which others are most like themselves; the most similar district demographically might be on the other side of the state, far beyond the local district's radar. Analyzing at the state level rather than the district level allows local educators to focus on their most important work: spotting opportunities to strengthen their practice.
- **Longitudinal analyses.** The last decade's expansion of state longitudinal data systems has facilitated much greater access to data on trends over time. The DART tools generally display the five most recent years of data on most indicators, and many Massachusetts data series go back ten years or more. The simplest version of a longitudinal analysis compares across cohorts: for example, last year's 4th graders versus this year's. More powerful analyses link the same individual students over time and follow their educational trajectories. Massachusetts has used these linked data to

create college enrollment “waterfalls,” starting with the 9th grade cohort and measuring what percentage of them reach key college milestones: graduating from high school on time, enrolling in college, and persisting in college (see Figure 1.) These are published in the DART Detail: Success After High School tool.

**Figure 1. Student Progression from High School Through Second Year of Post-Secondary Education**



- **Policy modeling.** Research can also be beneficial in policy modeling: testing the potential impact of various policy options to help guide decisions. As Massachusetts developed our request for flexibility on certain provisions of the Elementary and Secondary Education Act, we modeled how many districts and schools would likely be affected by our proposals so we could reduce unintended consequences of our proposed accountability system changes. For example, we considered including student attendance in our accountability model, but once we tested it, we realized it made little difference in the determinations for most schools and decided to err on the side of simplicity.
- **Predictive modeling.** Predictive modeling uses prior data on a student (or school or district) to predict their future outcomes. Massachusetts uses this approach extensively in its **Early Warning Indicator System**, which predicts for each Massachusetts student in grades 1 through 12 his or her likelihood of missing key academic milestones: proficiency on grade 3 reading, proficiency on grade 6 English language arts and mathematics, successful completion of all grade 9 courses, and on-time high school graduation.

Each student is assigned a “low,” “moderate,” or “high” risk of missing the milestone based on patterns of similar students in prior years. We produce both individual student-level and aggregate reports of these data and make them available to the educators serving each student through our secure online data system. These data allow district staff to better prioritize their resources and focus on students at greatest risk. The reports are among the most popular in our district data reporting tool.

- **Qualitative data collection.** Quantitative analyses like those above are often fairly straightforward to run, particularly when they involve existing state data. But the nuance of the story—and the anecdotes that can prove so powerful when talking about policy to general audiences—typically comes from qualitative data collections such as interviews, focus groups, and observations. When done systematically, qualitative data collection can also shed light on implementation challenges. Our statewide educator evaluation implementation study included interviews and focus groups of educators in 12 case study districts. This qualitative data indicated that teachers of and staff in non-core academic subjects tended to look the least positively on the new evaluation system. When this finding became clear, our agency worked with four state associations representing these educator groups to develop additional resources to make sure the evaluation system was effectively applied to these roles.
- **Stakeholder perceptions.** Collecting systematic, representative data on stakeholders’ perceptions on policy implementation—whether qualitative or quantitative—can reap enormous benefits for the state. It is a tangible way the state can demonstrate that it is listening and actively responding to criticisms and concerns. And it wards against paying too much attention to squeaky wheels when their perceptions are at odds with the majority opinion. Massachusetts collects these data through satisfaction surveys for superintendents and principals, statewide surveys of stakeholder perceptions on key initiatives, commissioned research, and other sources.
- **Causal analysis.** We are often limited by available data to measuring only whether a particular policy is correlated with an outcome: for example, whether student achievement appears to increase along with implementation of a new policy. But some types of policies and programs lend themselves to what academic researchers call causal analyses: analytical approaches that can credibly claim that a particular intervention actually *caused* the observed outcome.

Of these, the simplest analytically—but often hardest to implement—is a randomized controlled trial, in which participants are randomly assigned to either receive an intervention or to serve as a control. Because assignment is

random, program participants and non-participants are on average statistically the same before the intervention begins—so therefore any difference between treatment and control groups after the intervention must be due to the treatment (intervention or program) itself. But random assignment is often difficult in real-world policy settings, so researchers have invented other analytical approaches that approximate random assignment. These approaches include comparing outcomes for students who fell just above and below a qualification for eligibility in a program (for instance, a minimum GPA or an income threshold), comparing the trajectory of student outcomes before and after a policy was implemented, comparing outcomes for “treated” students versus other students with similar measurable characteristics, and so forth.

When executed well, all these approaches allow us to be more confident that any difference we see between the treatment and control or comparison group is due to the intervention itself, not other factors that might have changed simultaneously. This can make for convincing evidence to share with legislators and funders. But these studies can also be complex and difficult to explain and can sometimes limit the generalizability of the findings. They are particularly worthwhile when a policy area is highly politicized and a definitive study is needed to address rebuttals, or when a program is new so it is easier to set up a random-assignment mechanism to study its impact.

In Massachusetts, probably the most influential research of this kind has been the work of the School Effectiveness and Inequality Initiative based at the Massachusetts Institute for Technology, which has conducted a series of studies examining the impact of the state’s charter schools on student outcomes by using the school lottery process as a type of random assignment.<sup>7</sup> This work has convincingly demonstrated that students enrolled in charter schools in our state’s urban areas are achieving remarkable performance gains relative to similar students who applied to charter schools but did not gain admittance—and that those gains are greatest among the most disadvantaged students.

- **Cost analysis.** It is valuable but often quite challenging to determine the actual cost of running a program, particularly when much of the cost comes from reallocating existing staff to the new work. Research can shed light on this through resource cost analyses that tie individual salaries and other expenditures back to the programs they support. My agency has used resource cost analysis to examine the cost of implementing Expanded Learning Time, a state grant program that provides \$1,300 per pupil to about 20 schools per year to implement a longer school day. We learned that the per-pupil program cost varied from approximately \$1,500 to \$4,300 depending on how districts chose to pay for the additional time required for staff and external partners to cover the longer day.<sup>8</sup> We

have used this insight to develop guidance for new schools entering the program. In cases where it is possible to measure a program's impact, we may be able to go a step further and analyze the program cost relative to its impact. Such cost-benefit analyses are infrequently done but are increasingly important as state and district leaders look to maximize their K–12 education investments.<sup>9</sup> Several research and technical assistance organizations are leading the way in this area by developing new methodologies and tools for districts and states to analyze how effectively they are using their resources.

### CONDUCTING HIGH-IMPACT RESEARCH

With myriad options for analyses, it is crucial that states make smart choices about the research they choose to support. The most important consideration is selecting the right program for study. Principally, the program should be visible and sit squarely on the agency's strategic agenda so that learning more about its implementation and impact will help drive agency priorities forward. Further, strong candidates for research are policies or programs that are malleable (potentially adjustable on the basis of findings) and durable (likely to be sustained for at least a few years), because this creates conditions where change is possible. If more information about a program's implementation or impact will not affect decisions about its course—for example, if it has strong support from funders or stakeholders in its current form—or if the program is likely to be short-lived, then further study is likely not worthwhile. That said, if the state could conduct a causal analysis, that methodological rigor might counterbalance the voices supporting the program. But this is as much a political consideration as a technical one.

Almost as important as selecting the right program for study is designing the right set of products from the research. Too often when state education agencies commission research, they require researchers to produce a lengthy document describing the program in great detail and including every imaginable analysis of its impact. This research often comes at the end of an implementation timeline—well past the time when the program is malleable or when its durability can be secured. Massachusetts made this mistake with its five-year evaluation of Expanded Learning Time. We hired an external evaluation firm Abt Associates, to conduct a study of both program implementation and impact.<sup>10</sup> The firm interviewed stakeholders, conducted student and staff surveys, and analyzed state data to measure program outcomes, taking a comprehensive look at the program statewide each year. But we requested just a single deliverable: an annual report covering all findings. Each year the researchers had to wait until well into the fall for our state assessment data from the prior year to become available so they could measure student outcomes from the program. They then needed time

to analyze and write up their findings. As a result, we typically didn't have a report on the preceding school year until around February of the following year—far too late to influence program design decisions. Further, the 200-page document put off all but the most dedicated program staff from reading it, further diminishing its usefulness. And though the researchers spent considerable time thinking about how to fairly and accurately measure the extent and variability of program implementation, those ideas rarely crossed over to inform our agency's program monitoring work, let alone into tools districts or schools could use to assess their own implementation. Abt Associates did just what we asked them to do. Unfortunately, the report we explicitly asked for wound up not being useful.

Having learned our lesson the hard way, we now require shorter, differentiated products to more quickly get research findings into the agency's discourse. We rarely commission the classic end-of-year tome unless the legislature or a grant funder requires it. We only "do the tome" when a program is ending or substantially changing course and we want definitive documentation of the program's history to date. We now ask the researchers we hire to produce smaller, quick turnaround reports immediately after pieces of data collection are completed so that we can learn from their work in near real time. These might be internal memos, briefings for program leadership, or short summaries of findings intended for district audiences. Even when we are simply sharing data with academic researchers, rather than explicitly commission their work (see box), we require them to produce **short, field-oriented summaries** of their research papers that we then share with agency staff, superintendents, and principals.

## Working With Academic Researchers

University-based researchers are a great resource for conducting high-quality research. Academics interested in education policy want their research to be used and are often eager to contribute to state agency work. They can identify clever ways to answer research questions with existing data or new questions you had not thought to ask. They often can develop a more rigorous research design than your own staff. They can be particularly valuable in politically charged waters, for their work will often be viewed as more independent and even-handed. And if graduate students are involved in a project, investing time in them can pay off in a stream of work for years to come as they become familiar with the state's data and policy context.

Working with academics, however, also comes with some challenges. The academy values a different skill set than the policy sector. Academics are expected to narrow their focus to a very small question and expend as much time and resources as needed to answer it; policymakers must view each question in a broader context and likely face trade-offs in how much time and effort to expend across different program areas in a dynamic, relatively fast-moving environment. Similarly, complaints about academic writing are legendary; sometimes it seems the academy's *lingua franca* is the Greek equation. But most challenging is the subset of researchers who view their interaction with a state agency as an opportunistic request for data with no real spirit of collaboration or partnership to answer questions of mutual interest.

State agencies seeking to work with university-based researchers would be wise to clearly set expectations: What is the expected timeline for the work? Does the state want briefings on preliminary results along the way or an opportunity for agency staff to work alongside the university-based researchers? Should the researchers plan to present their findings at state board meetings or other public forums? An up-front investment in clarifying roles and responsibilities goes a long way toward establishing productive, enduring partnerships.

We also carefully examine variation in program implementation to gain insight into the conditions under which programs work best.<sup>11</sup> To this end, we generally require that researchers produce school- and district-level results, not just statewide aggregations, with the specific schools and districts named whenever confidentiality is not a concern. We also break down results for key subgroups, such as student demographic groups or types of educators. This gives us much better insight into the context of the findings and informs our technical assistance plans.

Finally, we increasingly look for opportunities to turn researchers' data collection methods into concrete tools for districts. For example, the research team analyzing our educator evaluation implementation came up with a methodology for assessing whether educators' evaluations met various criteria of our statewide evaluation system; for example, establishing goals, providing effective feedback, and so forth. We are turning this into a toolkit for districts to do their own self-assessments, more efficiently leveraging the dollars invested in that research work. Similarly, we have shared with districts the **questionnaires** we used for our statewide analyses so they can gauge their own educators' perceptions of the evaluation system.

## RESOURCES NEEDED

To successfully build a central role for research in policy making, sufficient agency resources must be dedicated to the work. A key resource of any state education agency is its longitudinal data system. This can be used both by internal staff and external researchers to answer key policy questions. For internal staff, access to these data and training on how to properly analyze them may be needed to make the most of the data set. (See the "Getting Started" section regarding building an internal research team.) Since external researchers need timely access to these data to conduct most studies, setting up systems to facilitate appropriate researcher access is an important investment in promoting research. The Federal Education Rights and Privacy Act allows state agencies to provide personally identifiable student information to researchers under certain conditions. State agencies should develop memoranda of understanding with researchers that detail those conditions and any additional requirements they may want. For example, states might want to require researchers to provide them with an advance copy of findings before public release or require them to use industry-standard encryption to secure data files.

Another resource, of course, is staff time. Internal staff will need to carve out time away from other priorities to make time for conducting actionable policy analysis. Making sufficient space for this work in your analysts' portfolios is probably the single most important thing a state leader can do to increase

the use of data and research in policy making. This might be made possible by automating or simplifying some required data collection or reporting activities or by growing the number of staff with analysis skills by retraining existing staff or hiring new staff. Your analysts will also need time to assist external researchers to ensure that these external parties are using state data accurately. A simple first step is to create a researcher's guide to your state data, to save staff from repeatedly answering the same basic questions. This guide should include an overview of what data are available to researchers and under what conditions, links to data codebooks with details of how variables are coded and formatted, and any key business rules researchers may need to know to use the data correctly.

It is helpful to have a designated research and evaluation coordinator or at least to have a staff person with those duties as part of a broader work portfolio. This person can help define and execute the agency's research agenda and serve as a liaison for researchers to help them use state data effectively. But, equally importantly, they can also work with senior leadership and program staff to ensure that the agency gets the most possible out of the research work, whether conducted internally or externally. Because our research coordinator in Massachusetts has a background in both research and program implementation, she can speak credibly to both sides and serves as a valued connector and translator between the two. When we hire external researchers, she ensures our procurement documents include the right information for bidders to be able to design appropriate research projects, getting us more for our money. She also manages most of our evaluation projects to ensure that vendors produce high-quality deliverables and answer our program staff's research questions.

A common rule of thumb in hiring an external research firm or academic researcher is for an evaluation to cost about 5 to 10 percent of the program budget (lower budgets may be feasible depending on circumstances.) In most cases, the data collection strategy is the key factor driving costs, with three main considerations:

- **The nature of the research questions.** Research questions that can be answered with quantitative analysis of existing data are least expensive; questions that require extensive interviewing, focus groups or other qualitative data collections cost substantially more. Questions that require classroom observations to answer are often the most expensive, since researchers typically need to visit a large number of classrooms to ensure a representative snapshot. That said, classroom observations can also offer the most valuable data since they give evidence of whether your policy or program is affecting classroom instruction—the ultimate aim of most education interventions.

- **The structure of the program.** Programs carried out similarly across all program sites are least expensive to study, since fewer sites need to be contacted to get a representative look. If program sites have flexibility about what or how to implement, the sample selection will need to be more complex and, therefore, more costly.
- **The desired representativeness of the sample.** If all you want is a rough statewide picture on a particular research question, researchers may only need to collect data from a few districts or schools. But if for political or practical reasons you want to ensure that all districts can participate in the study, or you want the researchers to stratify their sample to ensure that different types of districts are included (e.g., rural vs. suburban vs. urban, high- and low-performing, regions within the state), costs may increase.

Another cost factor besides the data collection strategy is embodied in the classic trade-off: “Cheap, quick, and good: pick two.” You can always make a project cheaper if you are willing to sacrifice quality or speed. But if you need an answer quickly, or if the results need to be unassailable, costs will likely be higher.

To give some specific cost examples, in Massachusetts we dedicated about 5 percent of the state’s Race to the Top grant to external program evaluation and hired two additional analysts to support our reporting, analysis, and evaluation work. This gave us sufficient funds to do in-depth implementation analyses of our major Race to the Top initiatives, as well as run specialized studies on other select initiatives. Project budgets ranged from \$40,000 for a one-time, small-scale analysis of how districts were using results from a statewide educator survey to a multi-year \$625,000 study on educator evaluation, which included substantial qualitative components and a statewide representative sample survey of principals and teachers. A from-scratch, moderate-length statewide survey of all superintendents and principals we developed cost about \$40,000 including survey development, deployment, and results reporting for respondents overall and broken down by several district types. When we re-administered a similar survey the next year and requested similar reports, the cost was about half that. We have generally found that a case study done well costs \$20,000 to \$30,000 per site to look at implementation and outcomes from a single program. Those costs can quickly add up if you want case studies representing the range of districts or schools in the program, in terms of either district characteristics (e.g., urban/rural, size, etc.) or quality of implementation.

Apart from cost, another key resource that supports research in state agencies is agency leadership. Leadership's role is to consistently signal that research is high-priority work that you want staff to undertake, and that you expect staff to use evaluation results and data analyses as part of your continuous improvement strategy. With competing demands for staff time, research can easily be relegated to the back burner and ultimately never completed. And all too frequently, after the research is completed, the results are left to gather dust. Leadership can drive a culture of interest in both the research findings and in using findings to drive agency and system improvement. Clear direction from the top is critical to ensure a state agency research program's success.

### GETTING STARTED

If you are ready to start boosting your state's research capacity, below are steps to get started.

#### 1. Build the Team

You probably already have staff analysts who prepare your state's federal reporting and other special data analyses. These staff may be helpful in conducting research, but do not assume the required skills are identical. Ideally, your research staff should be more than just number crunchers (see box for sample job description and Figure 2 for an organizational chart). They should substantively understand both research methodology and education policy, as well as data analysis. That said, staff do not need doctoral degrees; often master's students are well qualified for state analyst roles. Look for candidates who have taken beyond the minimum-required graduate coursework in statistics or research methodology and who have worked previously as a research assistant or analyst (not just class research projects.) Demonstrated skill in writing about technical topics for a general audience is key; you don't want your staff to produce work that you yourself (let alone stakeholders) can't understand. In Massachusetts we require a writing sample that demonstrates analyst candidates' skill in writing about data. Second-round interviewees do an exercise in which they analyze state data and produce a short report similar to what they might do on the job. The exercise requires that they make judgments about which data to highlight and clearly explain the results.

## Sample Policy Analyst Job Description

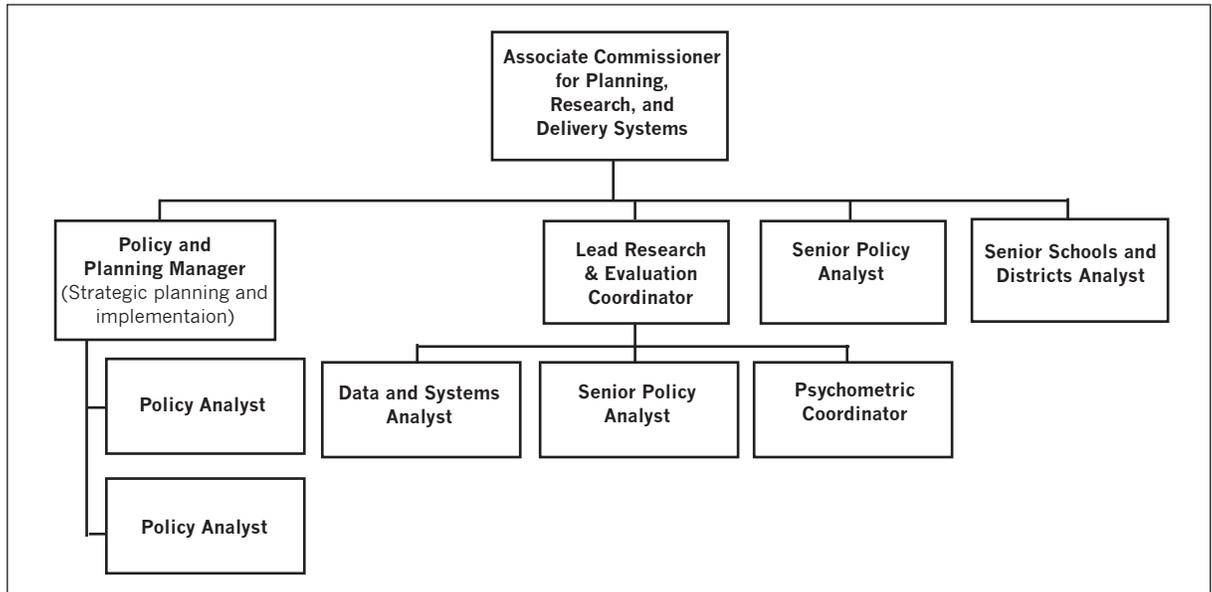
### Job duties:

- Analyze qualitative and quantitative data on the implementation and impact of the state's education initiatives
- Develop comparative, longitudinal analyses of state, district, and school performance
- Conduct analyses to model the potential impact of proposed policy changes
- Prepare reviews of recent research literature on key topics
- Work with program staff to identify program goals, implementation benchmarks, and outcome measures to help evaluate their programs
- Write memoranda, presentations, briefing materials, and other reports summarizing findings for program managers and senior executives

### Qualifications:

- Strong quantitative analysis skills, preferably including prior work experience in statistical analysis and reporting
- Successful completion of substantial coursework in statistics or econometrics, with coursework in multivariate statistical models and/or research methodology (preferred)
- Ability to use information gathered from research and key stakeholders to identify effective policy levers and promote improved outcomes
- Strong ability to write effectively for general audiences
- Strong project management skills
- Excellent interpersonal skills and collaborative and consultative approach

Figure 2. Massachusetts Department of Elementary and Secondary Education, Office of Planning and Research



You may opt to invest in improving existing staff’s research skills in addition to (or in lieu of) hiring outside. Most analysts who completed graduate work more than five years ago will not have had much exposure to today’s more sophisticated analytical techniques. My agency bridged this gap by hiring an advanced graduate student from a local university to teach seminars for the agency’s analysts on those techniques. Our goal was for our analysts to at least understand when such techniques could be appropriately used, even if the analysts were not employing the techniques themselves. Other options for strengthening your existing staff’s research skills are having them participate in national programs such as Harvard University’s Strategic Data Project, sending them to research conferences by organizations such as the Association for Education Finance and Policy or the American Education Research Association, or providing them with training offered by local universities or national centers like the Interuniversity Consortium for Political and Social Research. Staff might also develop formal partnerships with any external researchers the agency is working with to build their skills as they are embedded in their job. All these opportunities will expose them to and connect them with the broader education policy research world.

## 2. Build the Agenda

If you are just beginning your research work, start small. Pick just one program or policy to study, ideally one that meets the criteria described above: visible, malleable, and durable. Your analyst staff may help you further narrow options by determining which programs have existing data

or program information (therefore making them easier to study). Consider launching a pilot project; this allows you to embed an evaluation design from the program's inception, ensuring a knowledge base is built on the program implementation and impact throughout the pilot period. As your agency knowledge and capacity grow to manage the work, more projects become feasible. Ultimately, you should build an agency **research agenda** that tightly aligns with your agency's strategic plan, with each goal in your plan carrying a set of attached research questions and planned deliverables to answer them. This will help strategically guide your agency to improve program implementation and heighten impact.

### 3. Organize the Data

Your state longitudinal data system is a great research resource, but it may take effort to organize it in a way that allows researchers to easily use it for analysis. The first step is to document the available data and business rules, as noted earlier, to save time answering basic questions. Creating special research files that integrate data across sources and removing some identifying student data (such as name and date of birth) may be helpful. That way, as researchers request data or as internal research requests arise, your staff will have everything at hand rather than having to make special files each time. Finally, you should establish a standard data-use agreement or memorandum of understanding that all external researchers must sign to gain access to your agency data.

### 4. Mobilize Resources

Since your internal staff will be major contributors to your agency's research, they should play a key role from the get-go in producing analyses to support your agency's research agenda. That said, some research is best done by external partners, whether local universities, for-hire research firms, or the federal Regional Education Laboratory program (which offers research and data analysis to state education agencies). Nate Schwartz's essay in this volume provides more details on how to make the "build or buy" choice and how to manage work most effectively.

Consider applying for grant funding to support your research. The Institute for Education Sciences, part of the U.S. Department of Education, recently created several grant programs that can help launch or sustain collaborations between state education agencies and research organizations. In fiscal year 2015, available grants to support emerging partnerships topped out at \$400,000 over two years; grants to support major research and evaluation projects hit \$5 million over five years. Other local and national funders are increasingly interested in supporting research partnerships as well.

## 5. Formalize Feedback Loops

All this research effort is for naught if it does not improve your agency's work. In Massachusetts, we use the **U.S. Education Delivery Institute** strategic planning and implementation systems, which provide us with routines and structures for sharing research with agency staff and leadership. Even without this degree of formality, simple steps such as asking analysts or external researchers to write short summaries of their findings and provide timely briefings to leadership and program staff can go a long way to infusing research into the agency's work. The key is consistently making clear to staff that you expect to have data and information available as you make decisions and that you expect them to use those same data to improve their own work. Having program staff help build the research agenda and questions for their work establishes continuous improvement as an agency goal and creates a way for it to actually happen.

## CONCLUSION

State education agencies have much to gain from investing in their research capacity. Developing a research agenda that dovetails with the agency's strategic direction and making thoughtful choices about how to carry out that research agenda gives agencies a better chance of having the right information at hand when they need to shift course or make major decisions. State agencies can also more credibly demonstrate to the field that their input makes a difference in how the agency operates. These factors make research critical to advancing productivity in K–12 education as decision makers become better equipped to make informed decisions and to shift resources to maximize impact. Ultimately, research supports a cycle of continuous improvement that yields better programs and smarter investments.

## ENDNOTES

1. I would like to thank Heather Peske of the Massachusetts Department of Elementary and Secondary Education and the CRPE staff, along with two external reviewers, for helpful comments on previous drafts of this essay.
2. See Thomas Kane and Douglas Staiger, *Gathering Feedback for Teaching: Combining High-Quality Observations With Student Surveys and Achievement Gains* (Seattle, WA: The Bill & Melinda Gates Foundation, 2012).
3. See Douglas Harris, *Value-Added Measures: What Every Educator Needs to Know* (Cambridge, MA: Harvard Education Press, 2011).
4. SRI International, *Research Brief: Early Implementation of the Massachusetts Educator Evaluation Framework* (Menlo Park, CA: SRI International, 2014).
5. Many in this line of work distinguish between research and program evaluation. As with any academic distinction, perspectives abound as to whether a difference between the two exists and, if so, where precisely the line between them is. In the end, the techniques used to conduct studies—the ones I list in this section—are largely the same, irrespective of whether the study is called research or an evaluation. My view is that, from the perspective of a state education agency leader, the distinction is not important.
6. Massachusetts Board of Higher Education and Massachusetts Department of Elementary and Secondary Education, *Massachusetts School to College Report: High School Class of 2005* (Boston, MA: 2008).
7. See Atila Abdulkadiroglu et al., “Accountability and Flexibility in Public Schools: Evidence from Boston’s Charters and Pilots,” *Quarterly Journal of Economics* 126, no. 2 (2011): 699-748; see also Joshua D. Angrist et al., *Stand and Deliver: Effects of Boston’s Charter High Schools on College Preparation, Entry, and Choice* (Cambridge, MA: National Bureau of Economic Research, 2013).
8. Tammy Kolbe and Fran O’Reilly, *More Time in School: An Analysis of the Costs Associated with Schools Implementing the Massachusetts Expanded Learning Time Initiative. Prepared on behalf of the Massachusetts Department of Elementary and Secondary Education.* (Malden, MA: Massachusetts Department of Elementary and Secondary Education, 2012).
9. See Matthew M. Chingos and Grover J. Whitehurst, *Class Size: What Research Says and What it Means for State Policy* (Washington, DC: Brookings Institution, 2011); see also Douglas Harris, “Toward Policy-Relevant Benchmarks for Interpreting Effect Sizes: Combining Effects With Costs,” *Educational Evaluation and Policy Analysis* 31, no. 1 (2009): 3-29.
10. See Amy Checkoway et al., *Evaluation of the Massachusetts Expanded Learning Time (ELT) Initiative Year Five Final Report: 2010–2011* (Cambridge, MA: Abt Associates, 2012).
11. See Anthony S. Bryk et al., *Learning to Improve: How America’s Schools Can Get Better at Getting Better* (Cambridge, MA: Harvard Education Publishing, 2015).