

POLICY

SNAPSHOT

BY LAUREN MATLACH

MARCH 2015

Evaluating Evaluation Systems: Policy Levers and Strategies for Studying Implementation of Educator Evaluation

In recent years, states have invested significantly in the development and implementation of educator evaluation systems. Studies of these new systems are critical to understanding whether the systems are producing accurate and understandable results that can be used to make valid and reliable inferences about professional development needs and to inform human capital decisions. Studies can provide feedback on implementation, support continuous improvement, and increase understanding of evaluation systems' impact on teaching and learning. However, despite the importance of studying educator evaluation systems, prioritizing and funding studies can be challenging. Successful studies require expertise, time, and a shared commitment to conduct research from the state education agency down to the educators participating in the study.

In this Policy Snapshot, we focus on how states might prioritize and fund studies of educator evaluation systems. We highlight the importance of studying educator evaluation implementation, share strategies for prioritizing and funding studies, and share examples of successful funding strategies from various states and districts.*

Why Study Educator Evaluation Systems?

Effective educator evaluation systems provide actionable data that inform educators on how to improve teacher practice. Educator evaluation results also are used to inform high-stakes decisions related to hiring, tenure, and compensation. Therefore, it is important to ensure that measures are valid and produce consistent and useful results. Studies of evaluation systems can inform system improvement as well as build system credibility and buy-in from educators.

Depending on study design, data and analyses from studies of educator evaluation systems can provide the state with multiple pieces of information, including the following:

- Documentation of the evaluation system design and resources
- Evidence of how the system works in practice (e.g., how districts collect, store, access, and use evaluation data to make decisions)
- Strengths and weaknesses of the design, supporting resources, and implementation strategy

* Although we share examples in this brief, we do not endorse any of the states, districts, or studies mentioned.

- Challenges to anticipate during full implementation
- Stakeholder perceptions
- Quality control checks on the fidelity of implementation and the validity and reliability of results across districts and schools
- Evidence of ways in which educator practice has changed over time
- Areas of educator practice in need of professional development
- Impact on student growth
- Information on return on investment

Strategies for Prioritizing and Funding Studies of Educator Evaluation Systems

State education agencies (SEAs) and policymakers have used different policy levers and strategies to prioritize and fund educator evaluation systems. This section highlights different strategies that SEAs and policymakers have used and provides important considerations associated with each approach.

- 1. Create legislative or regulatory mandates to study educator evaluation systems.** Some state regulations require validation or early implementation studies of educator evaluation systems. For example, Public Act No. 12-116 in Connecticut required the Neag School of Education at The University of Connecticut to conduct a study of the state's teacher evaluation and support pilot program. Regulations described the purposes of the study and prescribed a deadline by which researchers needed to report findings. This approach's strength was that it codified the need for a study. However, requiring a study with legislation does not necessarily address the issue of funding. For example, the study in Connecticut, which was conducted with focus groups, interviews, and surveys in 14 districts, cost the state \$245,000 (Lambeck, 2014).

Sample Regulatory Language

The following is a sample of regulatory language from Public Act No. 12-116 of Connecticut:

Sec. 53. (a) The Neag School of Education at The University of Connecticut shall study the implementation of the teacher evaluation and support pilot program described in section 52 of this act. Such study shall (1) analyze and evaluate the implementation of the teacher evaluation and support program adopted pursuant to subsection (b) of section 10-151b of the general statutes, as amended by this act, for each local or regional board of education participating in the teacher evaluation and support pilot program, (2) compare such teacher evaluation and support program adopted by each local or regional board of education pursuant to subsection (b) of section 10-151b of the general statutes, as amended by this act, to the teacher evaluation and support program guidelines adopted by the State Board of Education pursuant to subsection (c) of said section 10-151b, and (3) compare and evaluate the use of student performance data on the state-wide mastery examination, pursuant to section 10-14n of the general statutes, and the use of student performance data on progress monitoring tests approved by the State Board of Education as an indicator of and method for student academic growth and development.

(b) Upon completion of such study, but not later than January 1, 2014, the Neag School of Education at The University of Connecticut shall (1) submit to the State Board of Education such study and any recommendation concerning validation of the teacher evaluation and support program guidelines adopted by the State Board of Education pursuant to subsection (c) of section 10-151b of the general statutes, as amended by this act, and (2) submit such study to the joint standing committee of the General Assembly having cognizance of matters relating to education, in accordance with the provisions of section 11-4a of the general statutes.

Source: S.B. 458, p. 119 (<http://www.cga.ct.gov/2012/act/pa/pdf/2012PA-00116-R00SB-00458-PA.pdf>)

Nevada

Designing and implementing a valid educator evaluation system was paramount to leaders in Nevada. In recent years, Nevada state leaders have advocated for resources to implement a validation study, required district participation in a validation study, appropriated funding, and received results from Year 1 of the study.

A call for rigorous study. In Nevada's *Summary of Anticipated Final Recommendations and Implementation Considerations*, written in November 2012, the Teachers and Leaders Council recommended that the State Board of Education adjust timelines for educator evaluation implementation to allow time for validation work. The summary noted that "system integrity is paramount to the success for this new accountability and support endeavor" (Fitzpatrick & Salazar, 2012, p. 4). Because evaluation results will eventually be tied to high-stakes human capital decisions, the council called for a rigorous validation study of the new frameworks' pilot implementations and for subsequent, timely studies that advise revisions to the model.

Legislation and appropriations. The Nevada legislature passed *Senate Bill No. 407* in 2013, which required the state to conduct a validation study of the statewide educator framework. All districts were required to participate in the validation study during the 2013–14 and 2014–15 school years, and researchers selected a representative sample from those collective data. According to the 2013 *appropriations report*, the legislature also approved a new budget account with \$320,000 of general funds for the evaluation throughout the 2013–15 biennium. Additional details can be found in the *state regulations*.

The study. During Year 1 of the study, researchers administered Web-based educator surveys in 86 schools, conducted focus groups in 14 schools, and conducted 40 principal interviews. Results indicated that educators expressed the need for more time and training on the evaluation system and did not believe they were ready to implement the system statewide (WestEd, 2014).

Sources: http://www.doe.nv.gov/Educator_Effectiveness/Educator_Develop_Support/NEPF/NEPF_Study_FAQ_s_-_August_2014/
http://www.doe.nv.gov/Educator_Effectiveness/Educator_Develop_Support/NEPF/NEPF_Val_Study_Year_1_Report___052214_CH__DRAFT_to_NDE/
https://ccea-nv.org/images/stories/pdfs/TLC_Recommendations_Summary_11_14_12_pdf.pdf

- 2. Conduct studies internally.** States and districts may leverage internal expertise to study implementation of educator evaluation reforms. Such an approach reduces reliance on outside resources and may increase the likelihood of the continual study of implementation. However, to conduct studies internally, states and districts need to allocate sufficient resources and personnel to the task. Staff members need to have sufficient experience in study design, data collection, data analysis, and writing for lay audiences. They also need sufficient time to conduct the study. If staff members are already overwhelmed with responsibilities, studying implementation may require shifting responsibilities or may not be feasible.

Tennessee

In August 2012, the Tennessee Department of Education (DOE) created an Office of Research and Policy as part of its efforts to ensure that research informs the agency's work. The office is charged with "generating a program of study to ensure that policy decisions within the department are grounded in rigorous internal data analysis and guided by the findings of independent research studies" (Tennessee DOE, n.d.). Each year, Office of Research and Policy staff members meet with each department of the Tennessee DOE to talk about their priorities for the upcoming year and any research questions they might have. Based on these meetings, the Office of Research and Policy staff develops a research agenda. When conducting research, the office leverages expertise of internal staff, external partners, and Strategic Data Project (SDP) partners.

Internal staff. The Office of Research and Policy consists of a director and four policy analysts. The office has released reports on implementation of teacher evaluation in Tennessee as well as a series of policy briefs and presentations, including the following (Tennessee DOE, 2015):

- Making Every Observation Meaningful: Addressing Lack of Variation in Teacher Evaluation
- Teacher Retention in Tennessee: Are We Keeping Our Best Teachers?
- Teacher Characteristics and Effectiveness

SDP partners. Tennessee DOE is an agency partner with SDP, which places fellows with partner agencies and offers ongoing support related to data use. Especially during the early stages of building the Office of Research and Policy, SDP fellows increased the analytic capacity of the office. For example, SDP fellows examined [statewide trends in the relationship between retention and effectiveness](#) using teacher evaluation data (Pratt & Booker, 2014).

External partners. Tennessee DOE partners with external agencies to conduct additional research. For example, the department worked with the Tennessee Consortium on Research, Evaluation, and Development to produce a [report of the findings from the Race to the Top survey](#), which included questions about educators' perceptions of the new teacher evaluation system. Other collaborations include an [evaluation of a partnership pilot project](#) conducted by John Papay, John Tyler, and Eric Taylor (Tennessee DOE, 2013).

Funding. Because many federal grants require research and reporting, the Tennessee DOE is able to fund the department using both federal and state funds.

Sources: <http://team-tn.org/innovation/research/>
<http://tn.gov/education/data/research.shtml>
<http://tnclassroomchronicles.org/survey-captures-teacher-perceptions-of-evaluation-system/>

Texas

The Austin Independent School District in Texas conducts internal studies of its programs. The district has a separate department, the Department of Research Evaluation (DRE), staffed with professional researchers and evaluators with expertise in measurement and research methods.

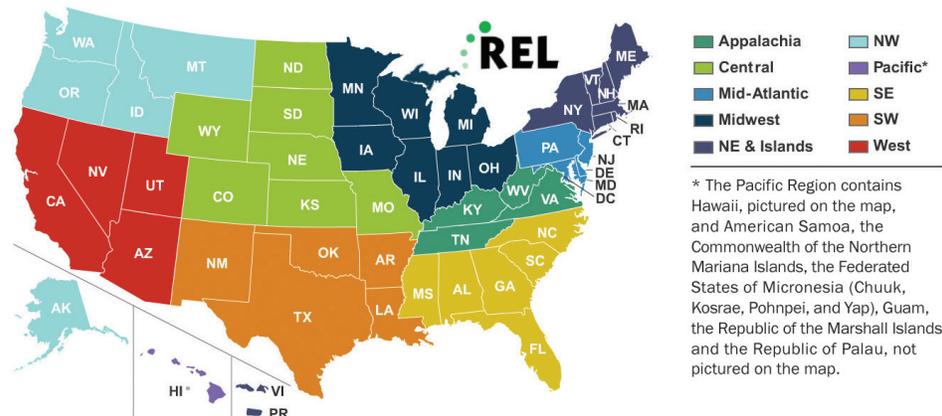
Evaluation plans. Each year, DRE staff members create evaluation plans that identify the scope of work for the coming year. Evaluation and program staff, the chief performance officer, and other executive-level district staff work to identify programs to be evaluated and services to be provided by DRE staff during the coming year. The evaluation plan includes program descriptions, the purpose of the evaluation, fiscal considerations, scope and methods, required reporting, and program support (ongoing support provided to program staff during the year).

Reports. To date, researchers have prepared numerous reports on the pilot teacher appraisal system as well as the district's performance pay program. These reports include briefs as well as longer reports.

Source: <http://www.austinisd.org/dre/about-us>

3. Leverage support from Regional Educational Laboratories (RELs). RELs have conducted studies of educator evaluation in multiple states. RELs are federally funded centers that work in partnership with state departments of education, school districts, and other stakeholders “to use data and research to improve academic outcomes for students” (IES, n.d.). Each REL has a set of research alliances that usually includes an educator effectiveness research alliance. Each alliance is comprised of practitioners, policymakers, researchers, and others who share an interest in the alliance’s policy area and work together to shape the REL’s research agenda. Each research alliance identifies its own topics and areas of interest and then creates its own method of prioritizing research questions to develop a scope of work, which in some cases has included evaluations of educator effectiveness initiatives. For example, REL West studied teacher evaluation in Arizona, REL Northeast and Islands studied educator evaluation implementation in New Hampshire, and REL Mid-Atlantic studied principal evaluation implementation in New Jersey and is assessing the internal consistency and validity of Pennsylvania’s principal evaluation rubric. Generally, states must be represented in a research alliance in order to have studies conducted in their states; however, each REL can reserve up to 15 percent of its funding to do work outside of the research alliances. States can find links to their RELs as well as contact information at <http://ies.ed.gov/ncee/edlabs/regions/>.

Regional Educational Laboratories



Source: <http://ies.ed.gov/ncee/edlabs/regions/>

Arizona

The Arizona Department of Education (ADE) partnered with Regional Educational Laboratory West (REL West) at WestEd to study pilot implementation of its teacher evaluation model and to examine the statistical properties of the model's component measures.

Perceptions of implementation. A recent study by Ruffini, Makkonen, Tejwani, and Diaz (2014) examined educator perceptions of pilot implementation of teacher evaluation systems. Arizona officials requested the study, and it was conducted by researchers from REL West in collaboration with the West Comprehensive Center. In five pilot districts, it tested the new state teacher evaluation model, and in five other districts, it received feedback on locally developed models. Data collection and analyses consisted of the following:

- West Comprehensive Center researchers conducted semistructured interviews and focus groups with a sampling of participating teachers and principals.
- ADE developed and administered an online end-of-year survey for participating teachers.
- REL West reviewers analyzed all data and wrote the final report. During the pilot, ADE's Research and Evaluation Division sought to maintain consistency in how the pilot districts calculated the student academic progress component of teachers' evaluations. The department trained staff in charge of student data management on data entry and monitored data management closely throughout the year. Feedback from the study helped ADE identify challenges and unintended consequences reported by teacher and principal participants, and based on the findings, ADE made multiple changes to its model and supports.

In response to study findings, ADE did the following:

- Trained principals in pilot districts in managing time and leading instructional conversations.
- Required evaluators to pass an online proficiency examination prior to conducting observations.
- Used a revised student learning objective process that made evaluations more standard across teachers.
- Planned to collect additional data on the use of student learning objectives.
- Planned to track certain findings of the study over time (e.g., monitoring the extent to which teachers assess their own practice over time).

Statistical properties. ADE also partnered with REL West to analyze the statistical properties of the component measures within its new state teacher evaluation model piloted in the five study districts. The study examined questions in three research areas: (1) statistical properties of the teacher observation instrument, (2) relationships between domain, item observation scores, and student academic progress, and (3) statistical relationships among the three components of the evaluation system (i.e., observed teaching performance, stakeholder survey input, and student academic progress). Based on report findings, Lazarev, Newman, and Sharp (2014) recommend that every teacher's evaluation "appropriately accounts for teaching environment, including grade level, classroom type, and student characteristics" (p. 17). The authors also recommend that ADE explore "more efficient ways to create summative scores for teachers by further examining the structure underlying the item scores or employing strategies to account for the statistical properties of the component (measure) and potential differences across groups of teachers" (Lazarev et al., 2014, p. 18).

Sources: http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_2015062.pdf
http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_2014050.pdf

- 4. Partner with local universities and programs.** Multiple states and districts, including New Jersey, Connecticut, and Prince George's County Public Schools (PGCPS) in Maryland, contracted with local universities to conduct studies of educator evaluation. University researchers are often based locally, which potentially reduces travel costs for site visits or meetings.

Partnership Spotlight

Prince George's County Public Schools and University of Maryland-College Park

A partnership between Prince George's County Public Schools (PGCPS) and the University of Maryland (UMD)-College Park has enabled multiple cohorts of graduate students to gain valuable in-the-field experience as well as class credit through a research apprenticeship while providing the district with high-quality evaluation studies at a reduced cost. Between 2008 and 2011, graduate researchers led by two UMD professors conducted two formative evaluations of Financial Incentive Rewards for Supervisors and Teachers (FIRST), a Teacher Incentive Fund program that included a new teacher evaluation system. The first study focused on the design and pilot-year implementation of FIRST, and the second focused on the implementation, payouts, and perceived effects of FIRST. Since 2011, UMD researchers have conducted an unrelated study on student-based budgeting as part of the partnership.

- 5. Apply for a grant from the Institute of Education Sciences.** Each year, the Institute of Education Sciences (IES) funds a limited number of studies through its three competitions. States seeking funding for a study can partner with a local university or research organization to apply for a grant to explore a particular research question. States interested in applying for IES grants first need to determine the IES research competition that best applies to the state and then identify any partners needed to conduct the study. The application process is intensive, so it is recommended that states partner with universities or organizations that have the experience and capacity to write proposals and have a history of winning IES grants. Application periods and deadlines vary by competition, so states should consult the IES website at <http://ies.ed.gov/funding/> for additional details.

Partnership Spotlight

District of Columbia Public Schools, the University of Virginia, and Stanford University

Partnership and application. District of Columbia Public Schools (DCPS), the University of Virginia (UVA), and Stanford University established a partnership to examine ways in which student achievement is affected by IMPACT, the educator evaluation system in DCPS. The partners submitted a proposal to the Researcher-Practitioner Partnerships in Education Research program, which funds partnerships of research institutions with SEAs or local education agencies. According to IES (2014), the partnerships should:

...identify an education issue with important implications for improving student achievement that is of high priority for the education agency, carry out initial data analyses regarding the education issue, and develop a plan for further research on the issue culminating in an application to one of the Institute's research grant programs. The ultimate goal of the partnerships is to conduct research that has direct implications for improving programs, processes, practices, or policies that will result in improved student outcomes.

Award. IES awarded the partnership \$398,332 to analyze how IMPACT has affected teacher applicant characteristics, teacher mobility, and teacher practice over time. In addition, researchers will examine the psychometric properties of evaluation system measures. Another goal of the study is to increase DCPS research staff's capacity. The partner organizations will meet every two weeks to discuss progress, and DCPS research staff will have the opportunity to receive mentoring from UVA faculty.

Source: http://ies.ed.gov/funding/ncer_rfas/partnerships.asp

- 6. Seek funding from foundations.** Multiple foundations have missions focused on improving outcomes for students, and many of them have supported efforts to improve educator evaluation systems and related outcomes measures, including the following:
- **The Wallace Foundation**—One of the foundation’s major initiatives is school leadership. To date, the foundation has funded multiple studies focused on the evaluation of school leaders.
 - **The Bill & Melinda Gates Foundation**—This foundation provided funding for the Measures of Effective Teaching (MET) project. Areas of the foundation’s education focus are teaching, learning, and innovation.
 - **The Joyce Foundation**—This foundation is dedicated to improving the quality of life in the Great Lakes region and nationwide. The Joyce Foundation funded the landmark study, *Rethinking Teacher Evaluation*, by the Consortium on Chicago School Research. The foundation accepts applications throughout the year.

SEAs also should consider local and regional foundations as potential sources of funding.

- 7. Consider setting or revising policies related to research data requests.** University researchers and graduate students seeking to study educator evaluation need access to data sources. Granting researchers access to data can be mutually beneficial: Researchers gain the access data needed for their study, and the SEA gains access to findings and can potentially further its own research agenda. Explicitly articulating the SEA’s criteria for evaluating requests—such as alignment to a specific SEA research agenda—can encourage researchers to pursue projects that reflect the SEA’s needs to seek collaboration.

Allowing researchers to place data requests can seem like a no-cost or low-cost approach to increasing research on educator evaluation. However, SEAs should consider what each request would require of its own staff resources. SEA staff members, data technicians, institutional review boards, and, in some cases, legal counsel, may need to review requests (National Forum on Education Statistics, 2012). In addition, staff time may be needed to share the data, provide oversight of the data requests, and train researchers to use data responsibly (National Forum on Education Statistics, 2012). Although it may be tempting to view this methodology as a no-cost approach to increase the amount of research conducted, states should consider the benefits and tradeoffs associated with allowing research data requests.

Key Resource

The National Forum on Education Statistics offers a [guide to supporting data access for researchers](#). This document provides an overview of potential benefits for data partnerships and identifies core practices for effectively managing data requests. The document identifies costs associated with providing data access to researchers and also highlights important considerations for state education agencies.

Source: <http://nces.ed.gov/pubs2012/2012809.pdf>

State Spotlight

Florida

The Florida Department of Education (DOE) has an [established process for handling research data requests](#). It revised its process in 2011 to ensure that all granted data requests were of potential benefit and use to the Florida DOE. When submitting research proposals, researchers must demonstrate how their study will support the Florida DOE’s [research agenda](#).

Source: <http://www.fldoehub.org/Research/Pages/default.aspx>

References

- Fitzpatrick, R., & Salazar, P. (2012). *Teacher and Leaders Council: Summary of anticipated final recommendations and implementation considerations*. Las Vegas, NV: Clark County Education Association. Retrieved from https://ccea-nv.org/images/stories/pdfs/TLC_Recommendations_Summary_11_14_12_pdf.pdf
- Institute of Education Sciences. (n.d.). *Regional Educational Laboratory Program* [website]. Retrieved from <http://ies.ed.gov/ncee/edlabs/regions/>
- Institute of Education Sciences. (2014). *Examining the Effects of IMPACT on Students Achievement: DCPS-UVA Research Partnership* [website]. Retrieved from <http://ies.ed.gov/funding/grantsearch/details.asp?ID=1486>
- Lambeck, L. C. (2014). *UConn: Teacher evaluation system off to a good start*. Retrieved from <http://www.ctpost.com/local/article/UConn-Teacher-evaluation-system-off-to-a-good-5109625.php#photo-5672694>
- Lazarev, V., Newman, D., & Sharp, A. (2014). *Properties of the multiple measures in Arizona's teacher evaluation model* (REL 2015-050). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Educator Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_2014050.pdf
- National Forum on Education Statistics. (2012). *Forum guide to supporting data access for researchers: A state education agency perspective* (NFES 2012-809). Washington, DC: U.S. Department of Education, National Center on Education Statistics.
- Nevada Legislative Council Bureau. (2013). *Education*. Retrieved from http://leg.state.nv.us/Division/fiscal/Appropriation%20Reports/2013AppropriationsReport/6_Education.pdf
- Pratt, T., & Booker, L. (2014). *Teacher retention in Tennessee: Are we keeping our best teachers?* Nashville, TN: Tennessee Department of Education Office of Research and Policy. Retrieved from http://www.tn.gov/education/data/doc/retention_report.pdf
- Ruffini, S. J., Makkonen, R., Tejwani, J., & Diaz, M. (2014). *Principal and teacher perceptions of implementation of multiple-measure teacher evaluation systems in Arizona* (REL 2015-062). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory West. Retrieved from http://ies.ed.gov/ncee/edlabs/regions/west/pdf/REL_2015062.pdf
- S.B. 407, 2013 Leg., 77th Sess. (Nv. 2013). Retrieved from <https://www.leg.state.nv.us/Session/77th2013/Bills/SB/SB407.pdf>
- S.B. 458, 2012 Gen. Assem. (Ct. 2012). Retrieved from <http://www.cga.ct.gov/2012/act/pa/pdf/2012PA-00116-ROOSB-00458-PA.pdf>
- Tennessee Department of Education. (n.d.). *Research* [website]. Retrieved from <http://team-tn.org/innovation/research/>
- Tennessee Department of Education. (2013). *Classroom Chronicles: Survey captures teacher perceptions of evaluation system* [website]. Retrieved from <http://tnclassroomchronicles.org/survey-captures-teacher-perceptions-of-evaluation-system/>
- Tennessee Department of Education. (2015). *Research and policy briefs* [website]. Retrieved from <http://tn.gov/education/data/research.shtml>
- WestEd. (2014). *Nevada Educator Performance Framework (NEPF) validation study: Final report*. San Francisco, CA: Author. Retrieved from http://www.doe.nv.gov/Educator_Effectiveness/Educator_Develop_Support/NEPF/NEPF_Val_Study_Year_1_Report__052214_CH_DRAFT_to_NDE/



I WANT TO KNOW MORE!

For more information or resources on approaches to studying educator evaluation, please contact gtlcenter@air.org.

Lauren Matlach is a consultant providing technical assistance for the Center on Great Teachers and Leaders.