Title: Summary of research on criteria used for teacher evaluations

Date: September 2014

Question: >> Could you provide research on the criteria used for teacher evaluations, including reviews of state practices?

Response:

This memo includes research reports and articles on the criteria used in teacher evaluations, as well as on state practices in teacher evaluation.

➢ Citations include a link to a free online version.
➢ Citations are accompanied by an abstract, excerpt, or summary written by the author or publisher of the document.

We have not done an evaluation of the rigor of these resources, but provide them for your information only.

References

Bill and Melinda Gates Foundation. (2011). *Measures of Effective Teaching (MET) Project.* Bellevue, WA: Author. "Research reports" page retrieved on September 2, 2014, from http://www.metproject.org/reports.php

Summary: Two-thirds of American teachers feel that traditional evaluations don't accurately capture the full picture of what they do in the classroom. They want information that they can trust from measures that are fair and reliable. The Measures of Effective Teaching (MET) Project was designed to find out how evaluation methods could best be used to tell teachers more about the skills that make them most effective and to help districts identify and develop great teaching. The project brought together 3,000 teacher volunteers in six different school districts with dozens of education experts and researchers. The result was a ground-breaking three-year study whose culminating findings were released in 2013. Based on this work, the MET project and its partners have provided practical guidance and tools to support quality evaluation and feedback. [*REL West note:* See also http://www.metproject.org/more/components.php for the classroom observation tools being studied by the MET project.]

Brandt, C., Thomas, J., & Burke, M. (2008). *State policies on teacher evaluation practices in the Midwest Region* (REL Technical Brief, REL 2008–No. 004). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved on September 2, 2014, from http://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/techbrief/tr_00408.pdf

Summary: This technical brief describes state-level policies and procedural requirements for guiding teacher evaluation practices at the district level in the seven states served by the Midwest Regional Educational Laboratory: Illinois, Indiana, Iowa, Michigan, Minnesota, Ohio, and Wisconsin. Building on the Issues & Answers report *Examining district guidance to schools on teacher evaluation policies in the Midwest Region* (Brandt et al., 2007), this technical brief reveals how teacher evaluation practices are addressed by state policies and other state-level initiatives that include teacher evaluation features. During June–August 2007 researchers completed a systematic search for publicly available policy data and conducted short interviews with key department of education staff in each of the seven Midwest Region states based on 13 key characteristics of teacher evaluation procedures used in previous large-scale investigations. The information is intended to provide policymakers and practitioners with a view of teacher evaluation policy at the state level in the Midwest Region.

Chaplin, D., Gill, B., Thompkins, A., & Miller, H. (2014). Professional practice, student surveys, and value-added: Multiple measures of teacher effectiveness in the Pittsburgh Public Schools (REL 2014–024). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. Retrieved on September 2, 2014, from http://www.mathematica-mpr.com/~/media/publications/pdfs/education/rel_profpractice.pdf

Summary: Responding to federal and state prompting, school districts across the country are implementing new teacher evaluation systems that aim to increase the rigor of evaluation ratings, better differentiate effective teaching, and support personnel and staff development initiatives that promote teacher effectiveness and ultimately improve student achievement. States and districts are implementing richer measures of professional practice alongside "value-added" measures of student achievement growth and in some cases are incorporating additional measures, such as student surveys. Pittsburgh is a leader in the nationwide movement to evaluate, enhance, and reward effective teaching. The analyses presented in this report were conducted to assist Pittsburgh Public Schools in refining its multiple measures of teacher effectiveness, to create a rich, valid, and comprehensive combined measure. The analyses help to assess how well Pittsburgh's measures differentiate among teachers and to establish how strongly they are correlated with each other. In addition, because each of Pittsburgh's three primary measures of teacher effectiveness is based on an approach that is being used or considered elsewhere, the findings have important implications for districts and states across the country that are developing and refining their measures of teacher effectiveness. The Pittsburgh Public Schools teacher evaluation system includes three types of multicomponent measures. The first measure—the Research-based Inclusive System of Evaluation (RISE), based on Charlotte Danielson's widely used Framework for Teaching—is an observation-based professional practice measure that relies on principals' assessments. The second measure is based on a student survey called the 7Cs, which incorporates students' perceptions of teachers' practices and was developed by Ronald Ferguson of Harvard University as part of the Tripod Project and administered by Cambridge Education. The third measure is a value-added measure that uses changes in student test scores to estimate each teacher's contribution to student achievement over up to three years of teaching. This study used

2011/12 data to describe how the ratings on the three measures are distributed across teachers and how the ratings are correlated.

Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002). *Teacher evaluation: A review of the literature.* Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA. Retrieved on September 2, 2014, from http://files.eric.ed.gov/fulltext/ED464915.pdf

Abstract: The teacher evaluation literature was reviewed to discover insights to further the understanding of teacher evaluation. Information was gathered from a review of the empirical research and theoretical literature related to teacher evaluation through more than 30 ERIC searches. A matrix was developed to summarize the findings from each research study and theoretical article or book. Findings were coded by topic, and the topics were grouped by general theme. Findings were categorized into five criteria that can be used to determine the effectiveness of a teacher evaluation system. These criteria are: (1) purposes for evaluation match the methods or procedures; (2) district commitment is evident in sound policies and practices and appropriate allocation of time and resources; (3) teacher evaluation is tightly connected to district priorities and school functions; (4) educational leaders play a strong, positive role in evaluation; and (5) the evaluation environment is supportive of ongoing professional learning. These criteria served as a foundation for effective teacher evaluation beginning in the 1980s and can serve as a framework for guiding research and educator efforts to strengthen teacher evaluation. An appendix compares research-based criteria with expert guidelines.

Daley, D., & Kim, L. (2010). *A teacher evaluation system that works* (Working Paper). Santa Monica, CA: National Institute for Excellence in Teaching. Retrieved on September 2, 2014, from http://www.tapsystem.org/publications/wp_eval.pdf

Abstract: Status quo approaches to teacher evaluation have recently come under increasing criticism. They typically assign most teachers the highest available score, provide minimal feedback for improvement, and have little connection with student achievement growth and the quality of instruction that leads to higher student growth. A more comprehensive approach has been demonstrated for ten years by TAP™: The System for Teacher and Student Advancement. This system includes both classroom observations and student achievement growth measures, provides feedback to teachers for improvement, is aligned to professional development and mentoring support, and provides metrics for performance-based compensation. This paper describes the TAP system, and examines data from a large sample of teachers to assess the distribution of TAP evaluations and their alignment to student achievement growth. We find that TAP evaluations provide differentiated feedback, that classroom observational scores are positively and significantly correlated with student achievement growth, that TAP teachers increase in observed skill levels over time, and that TAP schools show differential retention of effective teachers based on these evaluation scores.

Doherty, K. M., & Jacobs, S. (2013). *State of the states 2013. Connect the dots: Using evaluations of teacher effectiveness to inform policy and practice.* Washington, DC: National Council on Teacher Quality. Retrieved on September 2, 2014, from http://www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report

Excerpt: The National Council on Teacher Quality (NCTQ) has long advocated that any meaningful understanding of "effective" teaching must be rooted in results for kids. Whatever else they accomplish in the classroom, effective teachers improve student achievement. It seems like

commonsense. Yet, until recently, it has been an exceptional way of thinking about teacher quality, totally out of step with teacher policy across the states. As part of the annual State Teacher Policy Yearbook, NCTQ has systematically collected and analyzed state policies on teacher preparation, training, retention, compensation and other personnel policies. In this paper we provide:

1. A detailed and up-to-date lay of the land on teacher evaluation policies across the 50 states and the District of Columbia Public Schools;
2. An in-depth look at policy in states promising ambitious teacher evaluation systems (states requiring student growth and achievement to be a significant or the most significant factor in teacher ratings), including states' efforts to "connect the dots" and use teacher evaluation results in meaningful ways to inform policy and practice;
3. A compilation of some of the important lessons learned, pitfalls and successes states have experienced on the road to improving teacher evaluation systems.

Dounay Zinth, J. (2010). *Teacher evaluation: New approaches for a new decade*. Denver, CO: Education Commission of the States. Retrieved on September 2, 2014, from http://www.ecs.org/clearinghouse/86/21/8621.pdf

Excerpt: Whether spurred by the hope of Race to the Top funds or the research that students in ineffective teachers' classrooms do not make the academic gains of children taught by more effective educators, a number of states have enacted legislation in the last year to amend teacher evaluation policies. The summaries that follow provide highlights of these new provisions.

Doyle, D., & Han, J. G. (2012). *Measuring teacher effectiveness: A look "under the hood" of teacher evaluation in 10 sites*. New York: 50CAN; New Haven, CT: ConnCAN; and Chapel Hill, NC: Public Impact. Retrieved on September 2, 2014, from http://50can.org/sites/50can.org/files/Measuring%20Teacher%20Effectiveness.pdf

Excerpt: This report highlights 10 of the most advanced and talked-about teacher evaluation systems nationally: Delaware; Rhode Island; Tennessee; Hillsborough County, FL; Houston, TX; New Haven, CT; Pittsburgh, PA; Washington, DC (referred to throughout just as Washington); Achievement First (a charter management organization, or CMO); and the Relay Graduate School of Education in New York City. Together, these systems serve more than 1.6 million students each year. These are not the only systems taking on this work, but they are trailblazers. None of the sites we feature would say they have fully "figured out" teacher evaluation. Nor do we hold them up as examples of perfection. Yet they have all worked long and hard to carefully and thoughtfully tackle the most difficult challenges related to developing and implementing a high-quality teacher evaluation system. As more states, districts, school systems, and schools of education strive to develop and implement teacher evaluation systems, they will wrestle with these same challenges. This report is designed to help share information and lessons learned about this difficult, yet essential, work. Our report therefore strives to "get under the hood" of the teacher evaluation systems at these 10 sites. We have collected information about these emerging systems in a single place, using a consistent format that allows users to look across sites and see where organizations are proceeding in similar ways, and where approaches diverge. Although we recognize that there are many uses for data related to teacher effectiveness, and that many sites are grappling with how best to use this information to improve teacher quality, this report focuses specifically on evaluation practices.

Harris, D. N. (2012). How do value-added indicators compare to other measures of teacher effectiveness? What we know series: Value-added methods and applications (Knowledge Brief 5). Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved on September 3, 2014, from http://www.carnegieknowledgenetwork.org/wp-content/uploads/2012/10/CKN_2012-10_Harris.pdf

Abstract: In the recent drive to revamp teacher evaluation and accountability, measures of a teacher's value added have played the starring role. But the star of the show is not always the best actor, nor can the star succeed without a strong supporting cast. In assessing teacher performance, observations of classroom practice, portfolios of teachers' work, student learning objectives, and surveys of students are all possible additions to the mix. All these measures vary in what aspect of teacher performance they measure. While teaching is broadly intended to help students live fulfilling lives, educators must be more specific about the elements of performance that contribute to that goal—differentiating contributions to academic skills, for instance, from those that develop social skills. Once educators have established what aspect of teaching they intend to capture, the measures differ in how valid and reliable they are in capturing that aspect. Although there are big holes in what they know about how evaluation measures stack up on these two criteria, they can draw some important conclusions from the evidence collected so far. In this brief, the author will show how existing research can help district and state leaders who are thinking about using multiple measures of teacher performance to guide them in hiring, development, and retention.

Hinchey, P. H. (2010). *Getting teacher assessment right: What policymakers can learn from research*. Boulder, CO: National Education Policy Center, School of Education, University of Colorado. Retrieved on September 3, 2014, from http://nepc.colorado.edu/publication/getting-teacher-assessment-right

Abstract: It is well established that teacher quality makes a difference in student learning. Since the implementation of No Child Left Behind in 2002, staffing every classroom with a high-quality teacher has been an official national priority. That goal entails an implicit requirement to assess teacher and teaching quality more rigorously than has been the case in the past. Despite decades of research on how best to assess teacher performance, however, no consensus has evolved on any single assessment strategy or collection of strategies—indicating that the problem of designing adequate and appropriate assessment is inherently complex and controversial. Such complexity has not, however, prevented the Obama administration from encouraging policymakers to define "good" teachers as those who produce gains in student achievement, measured by gains in standardized test scores. Notwithstanding the federal enthusiasm for test scores, many researchers have warned against using a single measurement of any kind as the primary basis for such important personnel decisions as teacher retention, dismissal or pay. While there are important questions about what achievement scores can—and cannot—indicate about individual teachers, there is no question that placing excessive emphasis on test scores alone can have unintended and undesirable consequences that undermine the goal of developing an excellent teaching force. Given the experience to date with an overwhelming focus on student achievement scores as a basis for high-stakes decisions, policymakers would do well to pause and carefully examine the issues that make teacher assessment so complex before implementing an assessment plan. To facilitate such examination, this brief reviews credible research exploring: the feasibility of combining formative assessment (a basis for professional growth) and summative assessment (a basis for high-stakes decisions like dismissal); the various tools that might be used to gather evidence of teacher effectiveness; and the various stakeholders who might play a role in a teacher assessment system. It also offers a brief overview of successful exemplars. Based on the research reviewed, it is recommended that policymakers employ an assessment system that targets both continual improvement of the teaching staff and timely dismissal of teachers who cannot or will not improve.

Steps toward that goal include that policymakers: (1) Be clear about the purposes of any assessment before selecting strategies. Where formative and summative assessment are to be combined, plan to address the challenges of dual-purpose systems; (2) Involve all key stakeholders in system design; (3) Rather than employing a single assessment tool, gather evidence from multiple sources. Combine strategies so that the weakness of any single tool is offset by the strengths of another; (4) Be sure that the criteria for assessing performance, artifacts or other factors are credible and are well understood by teachers and assessors; (5) Provide high-quality, ongoing training for assessors and routinely calibrate their efforts to ensure consistent application of criteria; (6) Look to high-quality research on existing tools and programs to inform the design of assessment systems; and (7) Commit sufficient resources to produce high-quality, productive assessment. Brief Summaries of Teacher Evaluation Methods is appended.

Jerald, C. D., & Van Hook, K. (2011). *More than measurement: The TAP System's lessons learned for designing better teacher evaluation systems*. Santa Monica, CA: National Institute for Excellence in Teaching. Retrieved on September 2, 2014, from http://tapsystem.niet.org/publications/eval_lessons.pdf

Abstract: As the longest-standing and most successful effort to radically revamp teacher evaluation using multiple measures including student achievement gains, TAP: The System for Teacher and Student Advancement provides a powerful set of lessons learned. In this paper, researcher and writer Craig Jerald worked with Kristan Van Hook of NIET and TAP practitioners to distill the key elements of TAP's successful teacher evaluation into ten essential recommendations useful to states and districts in redesigning their own systems.

Kane, T. J., Kerr, K. A., & Pianta, R. C. (Eds.). (2014). *Teacher evaluation systems: New guidance from the Measures of Effective Teaching Project*. San Francisco, CA: Jossey-Bass. Retrieved on September 2, 2014, from http://www.metproject.org/downloads/Designing%20Teacher%20Evaluation%20Systems_freePDF.pdf

Book description: A book on 15 original studies by MET project partners using the MET project's extensive data set to provide insights on new teacher evaluation methods and on the current state of teaching practice in U.S. classrooms. Published by Jossey-Bass, the book balances detailed quantitative analysis with discussion of implications for policy and practice.

Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains* (Research paper, MET Project). Bill & Melinda Gates Foundation. Retrieved on September 3, 2014, from http://www.metproject.org/downloads/MET_Gathering_Feedback_Practioner_Brief.pdf

Abstract: There is a growing consensus that teacher evaluation in the United States is fundamentally broken. Few would argue that a system that tells 98 percent of teachers they are "satisfactory" benefits anyone—including teachers. The nation's collective failure to invest in high-quality professional feedback to teachers is inconsistent with decades of research reporting large disparities in student learning gains in different teachers' classrooms (even within the same schools). Many states and school districts are looking to reinvent the way they do teacher evaluation and feedback, and they want better tools. With the help of nearly 3,000 teacher-volunteers, the Measures of Effective Teaching (MET) project is evaluating alternative ways to provide valid and reliable feedback to teachers for professional development and improvement. In this report, the authors focus on the value of classroom observations. They test five different approaches to classroom observations. Each observation instrument is designed to do two things: (1) focus an observer's attention on specific aspects of teaching practice and (2) establish common

evidentiary standards for each level of practice. In this report, the authors investigate the properties of the following five instruments: (1) Framework for Teaching (or FFT, developed by Charlotte Danielson of the Danielson Group); (2) Classroom Assessment Scoring System (or CLASS, developed by Robert Pianta, Karen La Paro, and Bridget Hamre at the University of Virginia); (3) Protocol for Language Arts Teaching Observations (or PLATO, developed by Pam Grossman at Stanford University); (4) Mathematical Quality of Instruction (or MQI, developed by Heather Hill of Harvard University); and (5) UTeach Teacher Observation Protocol (or UTOP, developed by Michael Marder and Candace Walkington at the University of Texas-Austin).

Milanowski, A. T. (2011).*Validity research on teacher evaluation systems based on the Framework for Teaching*. Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans, LA. Retrieved on September 2, 2014, from http://eric.ed.gov/?id=ED520519

Abstract: After decades of disinterest, evaluation of the performance of elementary and secondary teachers in the United States has become an important educational policy issue. As U.S. states and districts have tried to upgrade their evaluation processes, one of the models that has been increasingly used is the Framework for Teaching. This paper summarizes validity evidence pertaining to several different implementations of the Framework. It is based primarily on reviewing the published and unpublished studies that have looked at the relationship between teacher evaluation ratings made using systems based on the Framework and value-added measures of teacher effectiveness. The research results summarized suggest that teacher ratings made using evaluation systems based on or related to the Framework have a correlation with value-added estimates of teacher effectiveness, in the 0.2 to 0.3 range. Measurement error in both the value-added estimates and evaluation ratings likely bias these correlations downward. Though fragmentary, the evidence reviewed suggests that Framework-based evaluation ratings can show substantial inter-rater agreement and, if multiple observers and multiple occasions of observation are used, the ratings can be reliable. However, a score based on a single observation by one observer is likely to have low reliability. The paper also discusses ways to extend and improve validity research on teacher evaluation systems. "Inter-rater Agreement Details for Cincinnati" is appended to this document.

National Council on Teacher Quality. (2013). *State policy on evaluation of effectiveness*. Washington, DC: Author. Retrieved on September 2, 2014, from http://www.nctq.org/statePolicy/policyIssueFindings.do?policyIssueId=6&masterGoalId=11&stateId=

Excerpt: Nineteen states require student growth to be the preponderant criterion. Seven states require student surveys. Forty-three states require more than two rating categories. [*REL West note:* This is a state-by-state searchable database.]

National Council on Teacher Quality. (2013). *State of the states: Trends and early lessons on teacher evaluation and effectiveness policies*. Washington, DC: Author. Retrieved on September 2, 2014, from http://www.nctq.org/p/publications/docs/nctq_stateOfTheStates.pdf

Excerpt: In this report, NCTQ provides: (1) a detailed picture of the teacher evaluation policy landscape across the states; (2) an in-depth analysis of states with some of the most ambitious teacher effectiveness policies; and (3) a set of early observations on the development and implementation of performance-based teacher evaluations.

Sartain, L., Stoelinga, S. R., & Brown, E. R. (2011). Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation. Chicago, IL: University of Chicago, Consortium on Chicago Schools Research. Retrieved on September 2, 2014, from http://www.joycefdn.org/assets/1/7/Teacher-Eval-Report-FINAL1.pdf

Excerpt: This report summarizes findings from a two-year study of Chicago's Excellence in Teaching Pilot, which was designed to drive instructional improvement by providing teachers with evidence-based feedback on their strengths and weaknesses. The pilot consisted of training and support for principals and teachers, principal observations of teaching practice conducted twice a year using the Charlotte Danielson Framework for Teaching, and conferences between the principal and the teacher to discuss evaluation results and teaching practice. Although the findings from this report focus on a specific pilot in a specific city, they have broad implications for districts and states nationwide that are working to design and develop evaluation systems that rely on classroom observations to differentiate among teachers and drive instructional improvement. Overall, we found that the Excellence in Teaching Pilot was an improvement on the old evaluation system and worked as it was designed and intended, introducing an evidence-based observation approach to evaluating teachers and creating a shared definition of effective teaching. At the same time, the new system faced a number of challenges, including weak instructional coaching skills and lack of buy-in among some principals.

Shakman, K., Riordan, J., Sánchez, M.T., DeMeo Cook, K., Fournier, R., & Brett, J. (2012). *An examination of performance-based teacher evaluation systems in five states* (Issues & Answers Report, REL 2012–No. 129). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved on September 2, 2014, from http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/rel_2012129.pdf

Abstract: This study reports on performance-based teacher evaluation systems in five states that have implemented such systems. It investigates two primary research questions: (1) What are the key characteristics of state-level performance-based teacher evaluation systems in the study states? and (2) How do state teacher evaluation measures, the teaching standards the evaluations are designed to measure, and rating categories differ across states that have implemented statewide systems? To answer these questions, the study team reviewed state education agency websites and publicly available documents for all 50 states to identify states whose performance-based teacher evaluation systems met the following criteria: (1) Was required for practicing general educators; (2) Was operational statewide as of the 2010/11 school year; (3) Included multiple rating categories; and (4) Used multiple measures of teacher effectiveness, such as observations, self-assessments, and professional growth plans. Five states (Delaware, Georgia, North Carolina, Tennessee, and Texas) met these criteria. Key study findings include the following: (1) Of the five states that met the criteria, three have new systems (1–3 years old), and two have systems that are more than 10 years old; (2) One state (Georgia) requires full annual evaluations for all teachers. In the other states, evaluations are annual for teachers whom the state defines as novice and less frequent or less comprehensive for more experienced teachers; (3) All five states include self-assessments and observations of classroom teaching as part of teacher assessment. States differ in who conducts the observations, how often evaluations are conducted, and what scoring parameters are used; (4) In each of the five states, teacher evaluation rubrics and scoring forms reflect most or all of the 10 teaching standards set forth by the Interstate Teacher Assessment and Support Consortium (InTASC). These standards relate to teachers' understanding of how students learn, content knowledge, instructional practice, and professional responsibilities. All 10 standards are

reflected in the teacher evaluation rubrics in North Carolina and Texas, 9 are reflected in Georgia, and 8 are reflected in Delaware and Tennessee. One InTASC standard—specifying that teachers demonstrate an understanding of how students learn—is absent in two states' evaluation rubrics (Georgia and Tennessee); and (5) States differ in the number of rating categories used and how they compute scores and determine passing scores.

Southern Regional Education Board (SREB). (2013). *Toward better teaching: A view of evaluation policies, practices and lessons in SREB states*. Atlanta, GA: Author. http://publications.sreb.org/2013/13E08_Teacher_Eff.pdf

Excerpt: This report provides a timely update on the policies SREB states have adopted in recent years to change how teachers are evaluated, with the ultimate goal of helping all teachers become more effective. Though the policies differ from state to state, they have great potential to provide teachers, school leaders, state policymakers and the public with much more rigorous and multifaceted information about teachers' performance and impact than has ever been available before … Part 1 of this report describes recent teacher evaluation policies undertaken by SREB states, either through legislative action or as part of the Race to the Top competition or the NCLB waiver process. The goal is to highlight and reflect on some of the important distinctions among the various states' policies. Part 2 focuses on implementation of these policy reforms and reflects on lessons learned to date.

White, T. (2014). *Evaluating teachers more strategically: Using performance results to streamline evaluation systems*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved on September 2, 2014, from http://www.carnegiefoundation.org/sites/default/files/CF_evaluating_teachers_strategically.pdf

Excerpt: This brief explores differentiation strategies in nine districts, two charter management organizations, and three states, Tennessee, Delaware, and Ohio. Interviews with system leaders and analyses of teacher evaluation policies reveal that these systems now vary the format or frequency of formal evaluation cycles, the format or frequency of classroom observations, or the type of observer conducting classroom observations, based on what is known about teachers' needs, strengths, and goals. Many of these school systems have embraced differentiation strategies as a way to conserve teacher evaluation resources or to deploy existing resources more efficiently. In some of the systems, however, differentiation strategies have required increased resources, as system leaders have introduced more frequent classroom observation or trained additional observers (e.g., peers, coaches). But even in those instances, officials claim differentiation strategies have helped make evaluation systems more attentive to teachers' individual needs for supervision and support. And that, they say, is likely to lead to more effective teacher evaluation systems.

Methods

Keywords and Search Strings Used in the Search

"Teacher evaluation" AND ("criteria" OR "state practices") AND "research"

Search of Databases

EBSCO Host, Google, and Google Scholar

Additional Organizations/Websites Searched

Education Commission of the States; U.S. Department of Education, Institute of Education Sciences

Criteria for Inclusion

In general, when REL West staff review resources, we consider—among other things—four factors:

- **Date of the Publication:** The most current information is included, except in the case of nationally known seminal resources.

- **Source and Funder of the Report/Study/Brief/Article:** Priority is given to IES, nationally funded, and certain other vetted sources known for strict attention to research protocols.

- **Methodology:** Sources include randomized controlled trial studies, surveys, self-assessments, literature reviews, and policy briefs. Priority for inclusion generally is given to randomized controlled trial study findings, but the reader should note at least the following factors when basing decisions on these resources: numbers of participants (Just a few? Thousands?); selection (Did the participants volunteer for the study or were they chosen?); representation (Were findings generalized from a homogeneous or a diverse pool of participants? Was the study sample representative of the population as a whole?).

- **Existing Knowledge Base:** Although we strive to include vetted resources, there are times when the research base is limited or nonexistent. In these cases, we have included the best resources we could find, which may include newspaper articles, interviews with content specialists, organization websites, and other sources.