

Do We Still Need Teacher Evaluation Systems?

06.2017 • Angela Minnici



How Did We Get Here?



50 Years in the Making

- Equality of educational opportunity
- Accountability
- Advances in statistical methods
- Teacher professional growth

President Obama and Teacher Evaluation



An Argument for Teacher Evaluation



Twin Goals



Consistent Findings, Persistent Gaps

- Studies consistently show that poor and minority students (and often low-performing students) are more likely to be taught by less-qualified or effective teachers (Demonte & Hanna, 2014).
- Despite ongoing efforts to address such education inequities, even very recent studies, along with analyses of state equity plans, suggest that equity gaps persist (Glazerman & Max, 2011).

Teacher Evaluation Ratings in Louisiana, 2012–13

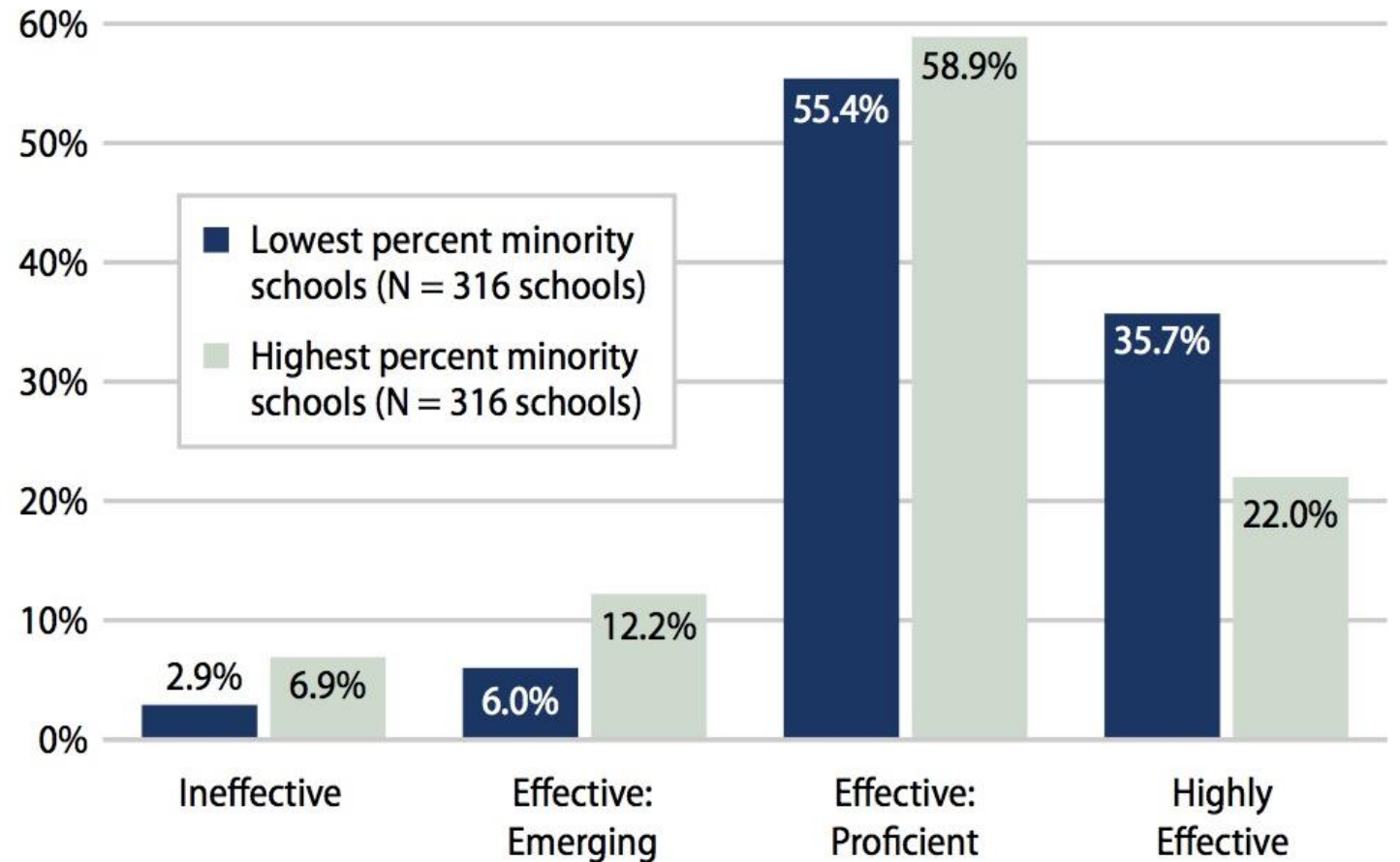
Top Versus Bottom Quartile Schools by Minority Enrollment

Source: DeMonte & Hanna, 2014, p. 4

FIGURE 2

Teacher evaluation ratings in Louisiana, 2012-13

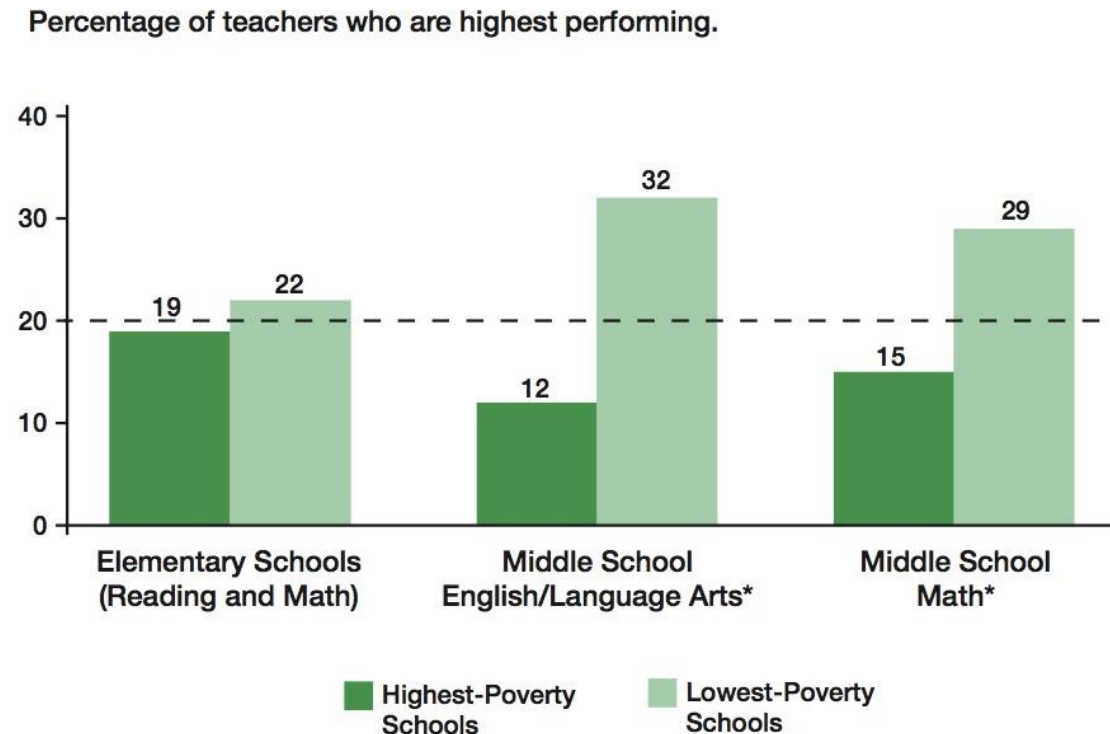
Top versus bottom quartile schools by minority enrollment



Prevalence of Highest Performing Teachers in the Highest and Lowest Poverty Schools (10 Districts)

Source: Glazerman & Max, 2011, as shown in Institute of Education Sciences, 2014, p. 6

Figure 2. Prevalence of Highest-Performing Teachers in the Highest- and Lowest-Poverty Schools



Teaching is Rocket Science

“Teaching is perhaps the most complex, most challenging, and most demanding, subtle, nuanced, and frightening activity that our species has ever invented. The only time a physician could possibly encounter a situation of comparable complexity would be in the emergency room of a hospital during or after a natural disaster.”

– Lee Shulman

Assessment of Teaching is Key



The National Conversation on Performance Management

Other Industries Are Ahead of Education



A Familiar Conversation

- Fifty-eight percent of executives responding to a public survey conducted by Deloitte believed that their current performance management approach drove neither employee engagement nor high performance.
- “We wondered if we could somehow shift our investment of time from talking to ourselves about ratings to talking to our people about their performance and careers — from a focus on the past to a focus on the future.”

(Buckingham & Goodall, 2015)

What Can We Learn from Business?

- Strengths-based approaches drive higher performance.
- Time with leaders matters for fueling performance.
- Performance management must move beyond simply ensuring compliance with mandates.
- Building strong teams is key to high performance.

(Minnici & Fipaza, 2015)

Some Companies Are Taking on Long-Standing Beliefs & Practices

The Case of Netflix

- “The best managers figure out how to get great outcomes by setting the appropriate context, rather than by trying to control their people.”
- “We develop people by giving them the opportunity to develop themselves, by surrounding them with stunning colleagues and giving them big challenges to work on.”
- “Career planning is not for us. Formalized development is rarely effective and we don’t try to do it (e.g., mentor assignments, career paths).”

(McCord, 2014)

Implementing Teacher Evaluation

Lessons Learned

Purpose of Evaluation

Districts have invested significant time and money in new teacher evaluation and support systems.

- In a healthy teacher evaluation system:
 - Teachers consistently receive and respond to high-quality, actionable instructional feedback.
 - Teachers, evaluators, and leaders focus on professional growth for individual teachers and across schools and districts.
 - The system includes supports such as coaching and professional learning that are targeted to individual and system needs.
 - All measures are appropriate and aligned with expectations for student learning.
 - Teachers trust the system and their evaluators.

Challenges With Evaluation

A healthy system cannot be fully realized when there is

- Inaccurate or unreliable measures of instructional practice
- Teacher distrust in the system
- Lack of adequate stakeholder engagement
- Premature use of data for high-stakes consequences
- Low quality or lack of professional growth opportunities
- Inadequate capacity to implement
- Poorly designed systems overall

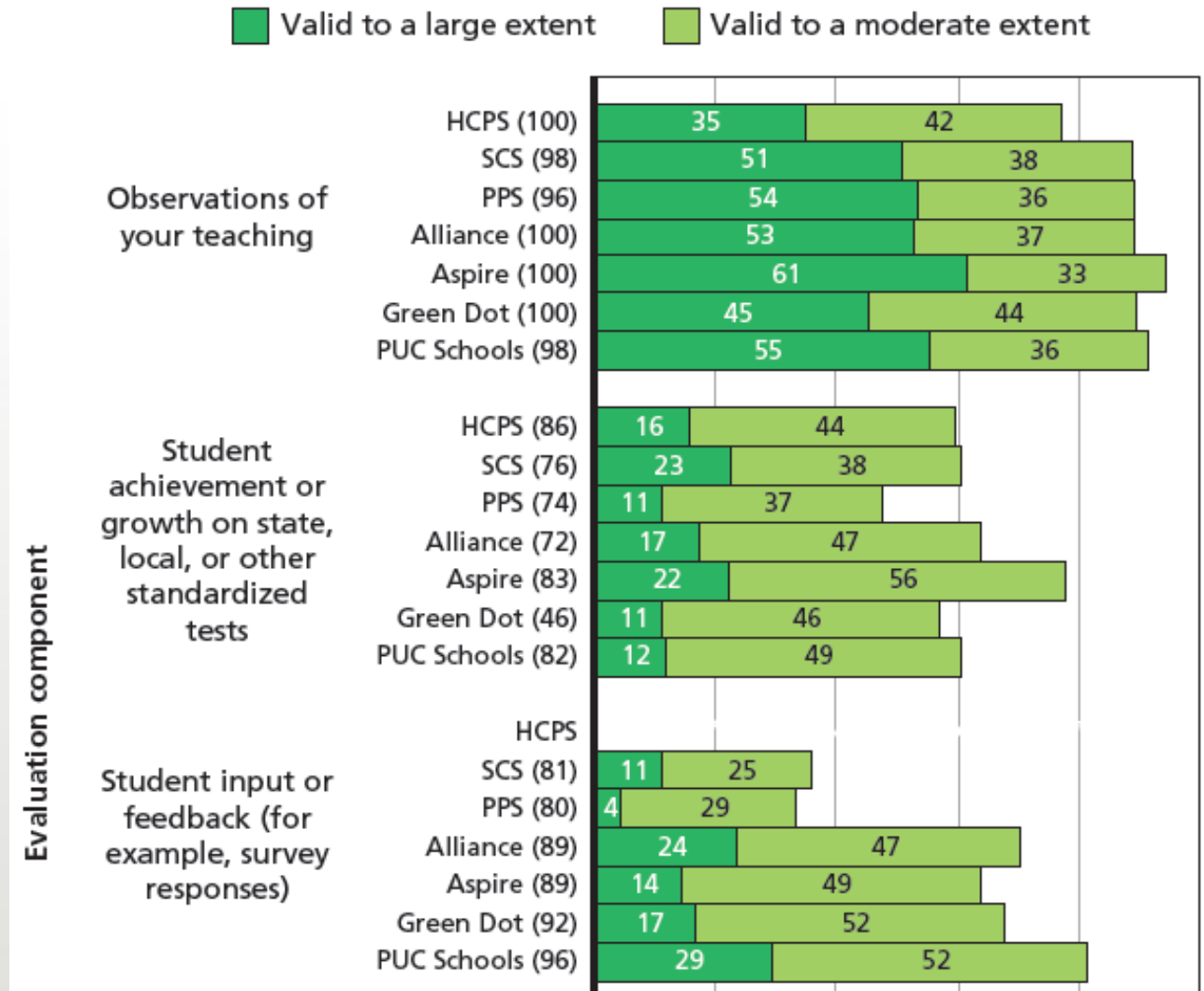
Challenges in Developing Evaluation Systems that Support Improvement: Three Big Ideas

- Quality of information/measures is low
- No opportunities to use data or no supports for doing so
- No trust or engagement in the system

Teachers' Views on Validity of Evaluation Components

Source: Stecher et al., 2016

Percentage of Teachers Reporting That Evaluation Components Are Valid to a Large or Moderate Extent, School Year 2013–2014



What We Know: Quality of Information

It still feels just like a snapshot to me. . . . I can put on a show for one day; something could go wrong for 20 minutes. I just don't feel like it gives a full representation even if I have five or six evaluations. To me, that's still a snapshot.

My only real complaint is that it's designed for you to see everything in one class period, especially in science. A lot of people do a dog and pony show, but I'm not going to do that and pretend that I do this every day. The way it's set up, if you have no evidence, it's a zero. I don't think that's a great way of setting it up.

(Stecher et al., 2016)

Research Gives Some Cause for Concern

Teacher observation systems may not be differentiating among teachers or providing them with actionable feedback.

- A study of the first year of implementation of evaluation measures in eight districts found that “nearly all teachers had classroom observation overall scores in the top two performance levels, limiting the degree of differentiation between lower- and higher-performing teachers” (Wayne et al., 2016, p. 23).

Research Gives Some Cause for Concern

(cont. from prior slide)

- Teacher observations can be a significant source of bias, penalizing teachers of lower-performing students (Steinberg & Garrett, 2016).
- In Chicago (Jiang & Spote, 2016), correlations between evaluation ratings and several student and teacher characteristics revealed that:
 - ▶ Thirty percent of the lowest-scoring teachers from classroom observations are found in the highest-poverty schools, while only 9 percent are in schools with the lowest poverty.
 - ▶ Minority teachers, male teachers, and less-experienced teachers also had lower value-added and observation scores.

Research Gives Some Cause for Concern

(cont. from prior slide)

- AIR identified instructional shifts required by the Common Core State Standards; eight common to both English language arts (ELA) and mathematics; three ELA-specific and four mathematics-specific.
- Forty-five teacher evaluation rubrics were analyzed for alignment with the instructional shifts (including rubrics from charter schools, states, districts, and nonprofit organizations).

(Welch et al., 2016)

Research Gives Some Cause for Concern

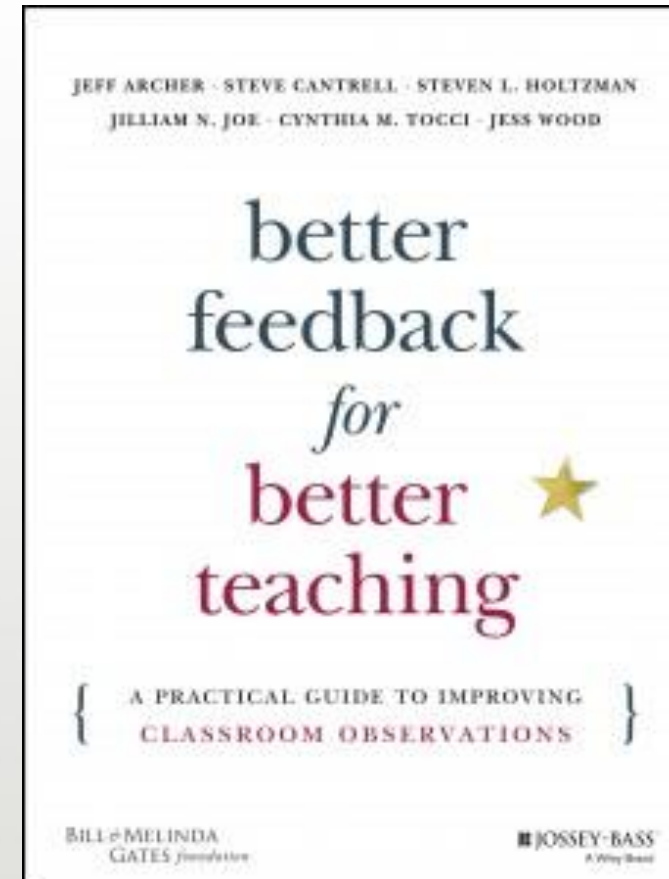
(cont. from prior slide)

- Scoring the alignment 0 (no alignment) to 1 (full alignment), only two rubrics scored above 0.5.
- In general, the report found a low alignment between evaluation rubrics and Common Core instructional shifts, meaning that teachers are not receiving feedback through their evaluations on these important instructional shifts.

(Welch et al., 2016)

Research Gives Some Cause for Optimism

Better Feedback for Better Teaching (Archer, Cantrell, Holtzman, Joe, Tocci, & Wood, 2016) is based on the knowledge of key partners in the Measures of Effective Teaching (MET) project, which carried out one of the largest studies ever conducted on classroom observations and of states, districts, and charter management organizations implementing observations.



(cont. on following slide)

Research Gives Some Cause for Optimism

(cont. from prior slide)

Best Foot Forward Study on Video Observations (2013):

- Treatment teachers collected more than twice the required number of video lessons; very few control teachers collected any videos.
- Treatment teachers were more critical of their own practice and regularly noticed new trends in their practice or in student behaviors.
- Treatment teachers perceived evaluators as more supportive and their observations as more fair.
- Administrators had more time to observe, spent less time on paperwork, and could be more flexible when they fulfilled observation duties.

Research Gives Some Cause for Optimism

(cont. from prior slide)

- Teacher choice in video did not impact ability to identify struggling teachers and teacher-chosen videos were generally representative of teacher quality.
- Teachers and administrators became more supportive of video through project.

(Kane, Gehlbach, Greenberg, Quinn, & Thal, n.d.)

What We Know: Teacher Trust and Engagement

Forty-seven percent of surveyed teachers reported that the evaluation process took more effort than the results were worth.

(Jack & Stratos, 2015)

What We Know: Teacher Trust and Engagement

(cont. from prior slide)

Several teachers from one district expressed fear that they could lose their jobs or fail to receive a pay increase as a result of one poor evaluation, despite the fact that neither of these consequences would have been imposed on the basis of a single evaluation.

– Intensive Partnership for Effective Teaching

Lessons Learned on Communication and Engagement

Best Foot Forward Study on Video Observations (2013):

- Develop a solid communication plan and revisit it often.
- Identify and use “communication ambassadors” (e.g., teachers who are designated to help inform staff, pass along feedback, and answer questions about the evaluation system).
- Consider engaging stakeholders to pilot and test new systems, to support the decision-making process, and to advise and provide input and feedback.

(Kane, Gehlbach, Greenberg, Quinn, & Thal, n.d.)

What We Know: Use of Data and Supports

- Sixty-six percent of teachers reported that the evaluation process helped them improve their practice.
- Sixty-five percent reported that the process helped them improve as professionals.

(Stecher et al., 2016)

What We Know: Use of Data and Supports

(cont. from prior slide)

Challenges:

- Knowing how to use the data to inform professional growth and improve practice.
- Easier to rate a teacher as effective than to give a lower rating when the principal does not know how to help the teacher improve.
- Difficulty of customizing professional development (PD) to the needs of individual teachers.
- Record-keeping for individualized PD.

(Center on Great Teachers and Leaders, 2016)

Collaborative for the Continuous Improvement of Educator Effectiveness (Collaborative)

Arizona



Colorado



Delaware



Louisiana



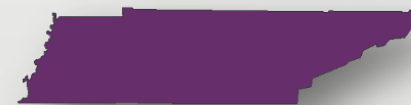
Massachusetts



Rhode Island



Tennessee



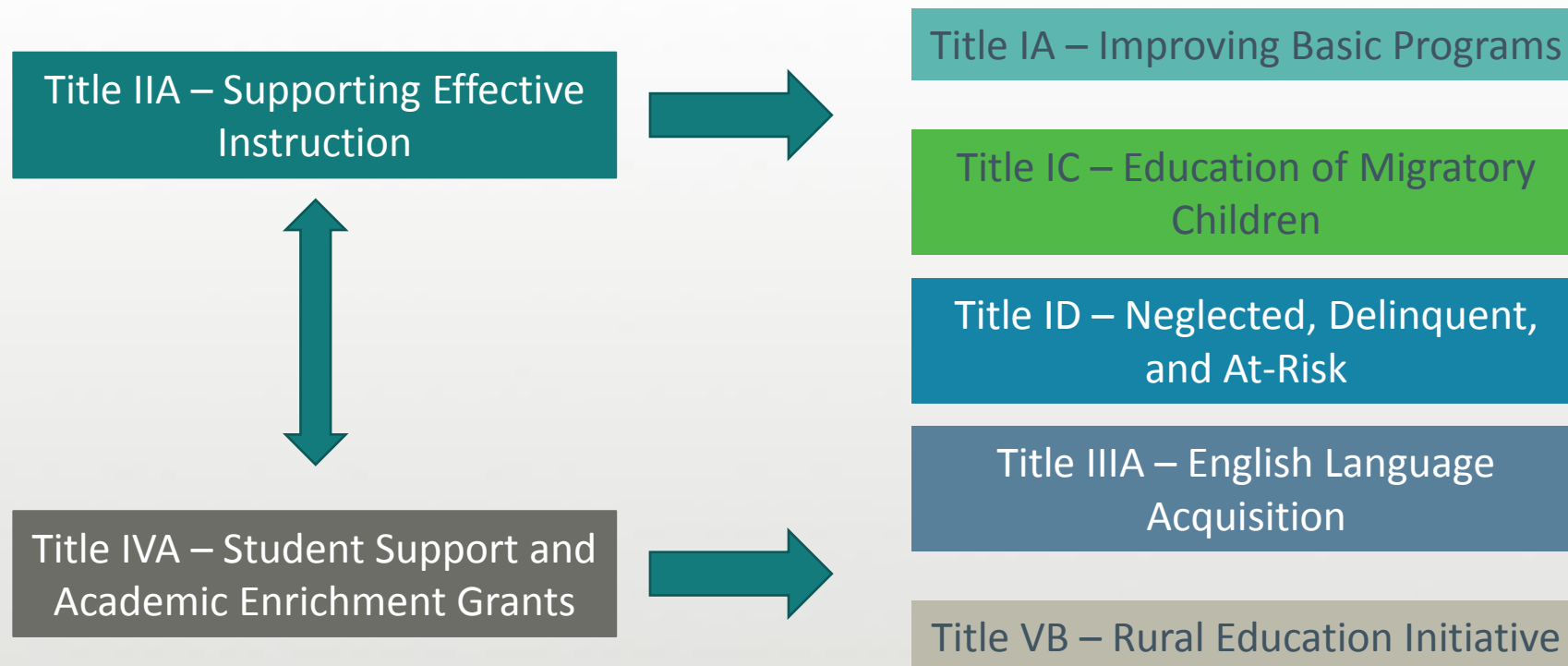
ESSA, Talent Management and Teacher Evaluation



ESSA Provides Opportunities to Support Your Talent Management Goals

- Funding is flexible
- “Braiding” funds for greater impact
- Title 2 Part A: 21 separate categories of allowed uses
- Fewer statutory requirements
- Less federal regulation
- Equity plans are part of ESSA plans
- Title IV can support equity goals

Fund Transferability



Time to Rebrand, Revise and Renew



References

- Archer, J., Cantrell, S., Holtzman, S., Joe, J., Tocci, C., & Wood, J. (2016). *Better feedback for better teaching: A practical guide to improving classroom observations*. San Francisco, CA: Jossey-Bass.
- Buckingham, M., & Goodall, A. (2015, April). Reinventing performance management. *Harvard Business Review*. Retrieved from <https://hbr.org/2015/04/reinventing-performance-management>
- Center on Great Teachers and Leaders. (2016). [Interviews with states on teacher evaluation systems]. Unpublished survey data.
- Demonte, J., & Hanna, R. (2014, April 11). *Looking at the best teachers and who they teach: Poor students and students of color are less likely to get highly effective teaching*. Center for American Progress. Retrieved from <https://cdn.americanprogress.org/wp-content/uploads/2014/04/TeacherDistributionBrief1.pdf>
- Glazerman, S., & Max, J. (2011). *Do low-income students have equal access to the highest-performing teachers?* NCEE Evaluation Brief. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
- Jack, J. & Stratos, K. (2015). *Teacher perceptions of educator effectiveness systems: Results from the second annual teacher evaluation survey*. Philadelphia, PA: School District of Philadelphia. Retrieved from: http://www.academia.edu/10682947/Teacher_Perceptions_of_Pennsylvanias_Educator_Effectiveness_System_Results_from_the_2nd_Annual_Teacher_Evaluation_Survey.

References

- Jiang, J., & Sporte, S. (2016). *Teacher evaluation in Chicago: Differences in observation and value-added scores by teacher, student, and school characteristics*. Chicago, IL: University of Chicago Consortium on School Research. Retrieved from <https://consortium.uchicago.edu/publications/teacher-evaluation-chicago-differences-observation-and-value-added-scores-teacher>
- Kane, T. J., Gehlbach, H., Greenberg, M., Quinn, D., & Thal, D. (n.d.). *The Best Foot Forward Project: Substituting teacher-collected video for in-person classroom observations — First year implementation report*. Cambridge, MA: Center for Education Policy Research, Harvard University. Retrieved at https://cepr.harvard.edu/files/cepr/files/l4a_best_foot_forward_research_brief1.pdf
- McCord, P. (2014). How Netflix reinvented HR. *Harvard Business Review*. Retrieved from <https://hbr.org/2014/01/how-netflix-reinvented-hr>
- Minnici, A. & Fipaza, J. (2015). *From the boardroom to the classroom: Can corporate America inform teacher evaluation?* (Blog post). Retrieved from <https://www.air.org/resource/boardroom-classroom-can-corporate-america-inform-teacher-evaluation>
- Sporte, S. E., Stevens, W. D., Healey, K., Jiang, J., & Hart, H. (2013, September). *Teacher evaluation in practice: Implementing Chicago's REACH students*. Chicago, IL: University of Chicago Consortium on Chicago School Research. Retrieved from https://consortium.uchicago.edu/sites/default/files/publications/REACH%20Report_0.pdf

References

- Stecher, B., Garet, M. S., Hamilton, L. S., Steiner, E. D., Robyn, A., Poirier, J., ... De Los Reyes, I. B. (2016). *The intensive partnerships for effective teaching*. Santa Monica, CA: The Rand Corporation. Retrieved from https://www.rand.org/content/dam/rand/pubs/research_reports/RR1200/RR1295/RAND_RR1295.pdf.
- Steinberg, M., & Garrett, R. (2016, June). Classroom composition and measured teacher performance: What do teacher observation scores really measure? *Educational Evaluation and Policy Analysis*, 38(2), 293–317. Retrieved from <http://journals.sagepub.com/doi/pdf/10.3102/0162373715616249>
- Walsh, K., Joseph, N., Lakis, K., & Lubell, S. (2017). *Running in place: How new teacher evaluations fail to live up to promises*. Washington, DC: National Council on Teacher Quality. Retrieved from http://www.nctq.org/dmsView/Final_Evaluation_Paper
- Wayne, A. J., Garet, M. S., Brown, S., Rickles, J., Song, M., & Manzeske, D. (2016, November). *Early implementation findings from a study of teacher and principal performance measurement and feedback: Year 1 report* (NCEE 2017–4004). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://ies.ed.gov/ncee/pubs/20174004/pdf/20174004.pdf>
- Welch, M. J., Davis, S., Isaia, R. G., Johnston, W. T., Stein, L. B., Jenuwine, H., & Macdonald, K. (2016, October). *Aligning evaluation: How much do teacher evaluation rubrics emphasize Common Core instruction?* Washington, DC: American Institutes for Research. Retrieved from <http://www.air.org/sites/default/files/downloads/report/Teacher-Evaluation-Common-Core-Alignment-October-2016.pdf>

Angela Minnici Senior Director of State Strategy WestEd

aminnic@wested.org

202-471-2439