

# Research on Teacher Evaluation:

## A Review of Statute, Regulation and Litigation in the Region

Helen M. Hazi

West Virginia University

*Limited research has been done to examine teacher evaluation in rural schools. This article presents an analysis of legislation and regulation of teacher evaluation in selected rural states, highlights their commonalities and differences, reports their litigation, and speculates on potential problems that can result in rural schools. It ends with recommendations for states to consider now that the Every Student Succeeds Act (formerly No Child Left Behind) has passed, and states have the option to reconsider their teacher evaluation plans.*

**Keywords:** teacher education, rural, law

### Introduction

Teacher evaluation has become a dominant reform strategy to address teacher quality in the states since 2009. Teacher evaluation as reform can be traced back to *A Nation at Risk* (The National Commission on Excellence in Education, 1983), was echoed by *A Blueprint for Reform* (U.S. Department of Education, 2010), was financially encouraged by *Race to the Top* (U.S. Department of Education, 2012), and was kept alive by flexibility waivers from the requirements of *No Child Left Behind* (U.S. Department of Education, 2002). Now that the *Every Student Succeeds Act (ESSA)* replaces *No Child Left Behind*, states can make their own decisions about the use of student test scores, because of ESSA's silence on teacher evaluation (Klein, 2015).

While little is reported about the implementation of new teacher evaluation systems in rural schools (e.g., Eady & Zepeda, 2007; Egley & Jones, 2004; Lasswell, Pace & Reed, 2008), this research is offered to stimulate future reports. It presents findings of an analysis of the legislation and regulation of teacher evaluation in the central states of Arkansas, Illinois, Missouri, Oklahoma and Texas. It includes highlights of their commonalities and differences, a summary of their litigation, and a discussion of a few problems that can come with these new evaluation systems. The article ends with recommendations for states to consider at this important crossroad that ESSA has provided.

### Methods

States were selected based on projected attendance at the annual fall conference of the National Rural Education Association (personal communication, Sandra Watkins, President, May 12, 2015). The website of each state's department of education was searched for information relevant to teacher evaluation from statute, regulation, handbooks, and news releases. Tables were constructed based on a similar regional analysis (Hazi, 2015a) and factors relevant to teacher evaluation.

While there can be variation among the states (and within states among different types of teachers), this analysis reports the highlights of their commonalities and differences. Since this is a review of website documents, this research is limited by the information that states choose to make public.

### Results

Three of the 5 states in this region have collective bargaining as seen in Table 1. Most (4 of 5) states allow local control of teacher evaluation as characterized by a low rating of 1 or 2 for level of state control. Arkansas, Illinois, Missouri and Oklahoma allow local control but specify the minimum contents of district policy, approve or monitor local plans, and/or obtain results. In states such as Texas with a high rating of 4, state control is more extensive and can include: approving it or an alternative, specifying details (e.g., walkthroughs, pre-observation conference, evaluation follows teachers to another district), monitoring what is done, setting guidelines for improvement plans, and annually evaluating implementation to make changes (Hazi & Arredondo Rucinski, 2009).

Table 1

*Factors Related to Change in Teacher Evaluation in the Central States*

<u>State</u>	<u>Collective Bargaining</u>	<u>Level of State Control</u>	<u>Early R2T Adopter</u>	<u>Student gains Criteria</u>	<u>Large scale instrument of 80s</u>	<u>Tenure Status Change</u>
Arkansas	No	1				2011
Illinois	Yes	2				2011, 13
Missouri	Yes	1				2011, 13
Oklahoma	Yes	1				2011,13
Texas	No	4		X	X	

While none of the states in the region got a head start on evaluation as an early adopter of Race to the Top funds, Texas did change its evaluation statute to incorporate student achievement as a criterion for teacher evaluation and did implement a statewide evaluation system in the 1980s.

Four states made changes to limit tenure. A review of the Education Commission of the States (ECS) database on tenure revealed that probationary status ranges from 3 years (e.g., AR) to 5 years (e.g., OK).

When the teacher evaluation law was enacted new teachers were required to have proficient or excellent ratings to receive tenure, while reduction in force became based on performance ratings (except AR that requires objective criteria) (ECS, 2015).

Table 2 shows that the frequency of teacher evaluation varies in the region. Beginning teacher evaluation is done on an annual basis, but for those veteran teachers with satisfactory scores, evaluation can be done every 2, 3 or 5 years with provisions such as annual plans with goals.

Table 2

*Features of Teacher Evaluation in the Central States*

<u>State</u>	<u>Frequency</u>	<u>% Student Test Scores</u>	<u>Instrument</u>	<u>Evaluation Influence</u>	<u>SLO/SGO</u>	<u>Training</u>
Arkansas	Annual Or 3 yrs*	2 consecutive yrs of SOAR Growth in future	TESS	Danielson	Goals in PGP	Teachscape Focus Proficiency Test ADE Website IDEAS modules
Illinois	Once/2yrs tenured Annual untenured	At least 30% Growth or 50% if unable to agree % specified for each of 4 ratings with % increasing **	Plan aligned to state standards or State Model Attendance Subject competency Lesson plan submitted	Danielson	2-4 SLOs aligned to school plan	State/local Prequalification process & passed test Content specified Retraining every 5 yrs

Missouri	Ongoing	Weighting locally determined but based on Evidence used not just test scores	Teacher Growth Guide State developed with 9 Standards & each with Quality Indicators	Multiple+	2 of 3 indicators focus on student learning	MOST ETS video library
Oklahoma	Annual	VAM calculated with no stakes until 2015-16 when 35%	OK Teacher & Leader Effectiveness Evaluation System (TLE)	Checklist with criteria for Approved instruments++	Postponed for 2014-15	Videoconferences, webinars, varied instrument training
Texas	Annual or Once every 5 yrs if proficient	20% Student Growth in 2014-15. May change in 2016-17 with VAM in tested subjects/SLOs/portfolios/districts pre-post tests	State developed T-TESS or locally created+++	NIET & Danielson	Self-Assessment (10%) includes goals, SLOs	NIET trainer of trainer sessions SLO online training

\*360 video technology may be used for observation

\*\* IL: Adjusts for special education, ELL, low SES

+RMC Corporation justified state built system with writings and studies, including Danielson, Marzano, Hattie and Lemov.

++Approved instruments to include: Danielson, Marzano, TAP, Tulsa TLE.

+++Instructional coaches and dept chairs can be trained & certified to assist in conducting evaluations.

Four states use student growth in calculating teacher scores, although some (AR, OK, TX) had “no stakes” in 2014-15. The growth percentiles range from 20-30% with some states anticipating increases. Only Missouri allows its districts to determine whether to use test scores and, if so, their weighting. In student growth calculations only Illinois adjusts for the variables of special education, English language learner status and low SES.

The second column also shows that all states in the region give the most weight to observation. Most established a state instrument but allow a locally chosen one (3<sup>rd</sup> column). Charlotte Danielson’s Framework seems to have influenced the region the most (4<sup>th</sup> column). Missouri developed its own instrument with a research-based rationale.

Most states in the region use the Student Learning Objective (SLO). SLOs are being viewed as a way to measure student growth in an array of subjects and grades, especially the non-tested, without the cost of new statewide assessments (EducationCounsel, 2013). Used in at least 30 states, teachers develop measurable SLOs individually, in teams, or school-wide, based on data and/or standards

and usually approved by a principal (Lacireno-Paquet, Morgan & Mello, 2014).

All states offer some form of training in the instrument or in writing student learning objectives. Training is available through websites, trainer of trainers, webinars, online modules and in some cases, video conferences. Arkansas uses Danielson’s Teachscape, while Missouri uses the ETS Classroom Video Library. Arkansas and Illinois require a test, while Illinois requires retraining every 5 years.

In 2012 teachers of non-tested subjects filed the first suit against their teacher evaluation system in Florida. Since then teachers in 7 other states and the District of Columbia have filed suits. Most are at the complaint stage working their way through the system. Complaints are typically 14<sup>th</sup> Amendment challenges to evaluation provisions that are vague, and to actions that are arbitrary and capricious. Teachers claim their scores give them ratings that deny them bonuses, damage their reputation, or put them in line for dismissal. Most complaints focus on the value-added (or student growth) score that tends to make their overall rating ineffective (Hazi, 2015b).

In this region in *Houston Federation of Teachers v. Houston Independent School District*

(2014) seven teachers who were dismissed filed a complaint in U.S. District Court of the Southern District. Houston officials implemented value-added assessments in 2007 before the Race to the Top initiative. Teachers claimed the tests, accounting for 50% of their score at the time, did not assess their curriculum, that the formulas were incomprehensible, there was a vague definition of student growth, that they were not informed on how to improve, and that their deficiencies were manufactured to match their test scores (Amrein-Beardsley, 2014). As the case unfolds perhaps the lessons of Texas can inform others in the region.

As in Texas, teachers in other states have challenged the use of the value-added measure of student growth because of its “complex mathematical formulas that can supposedly factor out all of the other influences and emerge with a valid assessment of how effective a particular teacher has been” (Strauss, 2014, np). However, a judge could conclude as it did in Florida that “This case, however, is not about the fairness of the evaluation system. The standard of review is not whether the evaluation policies are good or bad, wise or unwise; but whether the evaluation policies are rational within the meaning of the law” (Sawchuk, 2014, np).

### Discussion

Limited information exists about how schools are implementing teacher evaluation, but for the episodic incidents of: resignation (Burris, 2015), burning of evaluations (e.g., Bush, 2015), the strike (e.g., Babwin, 2012), the court case (e.g., Sawchuk, 2014) and suicide (e.g., Lovett, 2010). While scholars have begun to critique teacher evaluation in panels (e.g., at the 2014 and 2015 American Educational Research Association), in articles (e.g. Berliner, 2014), in blogs (e.g., Amrein-Beardsley, 2015) and in *The New York Times*, the consequences for rural schools and their teachers are still speculative.

Are there incidents with teachers like these in rural schools? If there are no incidents, does their absence indicate that rural schools are “getting evaluation right” (Darling-Hammond, 2013)? Teachers are tending to sue, when student test scores are weighted at 50% (Hazi, 2015b). Some states in this analysis delayed the high stakes (e.g., OK), while others (e.g., IL) weighted student test scores less than 50%. Since weighting at 50% is an important factor, then there may be fewer problems and litigation when the weighting is less. Use of additional measures such as student or parent surveys, or student learning objectives, could also temper the effect of student test scores. In addition, when districts involve stakeholders in selection, design and

revision, their criticism is minimized and buy-in is increased (SREB, 2015).

In rural schools, the principal carries the burden of implementation. Principals eager to make evaluation work, may become office-bound, and spend more time after hours with increased paper work, instead of in classrooms and giving feedback to teachers. Some may become lenient with their ratings to maintain their relationships, to “keep peace” in their small, family-like school, to avoid a “gotcha” evaluation, or to emphasize professional growth instead of the summative rating (Derrington, 2014).

Superintendents in rural schools may not be able to provide the support that principals need. To commit to this time-intensive endeavor, principals may need additional training, time-saving technology, and personnel such as retired principals and directors to conduct some of the evaluations (Derrington, 2014). If teachers are used to assist principals in this endeavor, superintendents should be cautious not to remove their best teachers from the classrooms for long blocks of time, and endanger student test scores.

Principals, as well as teachers, tend to trust observer ratings instead of student test scores to evaluate teachers (Goldring, et al., 2015). If this becomes the trend, then principals can come to rely on rubrics that they believe clearly define teaching and help them better understand instruction (Derrington, 2014), only to later learn that the rubric feedback may neither improve instruction nor increase student achievement.

The real business of evaluation is teacher improvement with professional development about how effective teaching looks in different subjects, for different ability levels of students, and in different grade levels. In this research some states (e.g., AR and MO) provide a video library to help this process. However, in addition to the comprehensive on-line video library, states are providing enhanced rubrics, e-learning courses, and materials that support teacher discussion (SREB, 2015). Do states provide generic training, or do schools customize and embed it locally? These types of discussions about instruction are best handled locally by teachers with principals and in small groups, rather than by remote.

Even if state departments provide local discretion on the observation instrument, rural schools with limited funds may most likely choose the state’s model, rather than to select its own and purchase a preferred instrument and its validity, electronic platforms to collect data, and formulas to calculate growth for assessments beyond the state tests. Rural schools may also select the state model, since it will usually come with its own evaluation support to include training, website and manuals.

Rural schools should not have to choose between diesel fuel and evaluation support services or professional development (e.g., Eady & Zepeda, 2007; Harmon, et al., 2007).

### Conclusions

States have been rushing to implement statute and regulation for waivers from the requirements of *No Child Left Behind*. At this crossroad states must decide whether to proceed as planned, or to reconsider their options and minimize damages. In this rush states may have overlooked that there is no evidence that teacher evaluation improves instruction or increases student achievement. States should:

- Establish a *state task force* to collect strengths and challenges to evaluation. As critics have pointed out, many are anticipated. A state would be wise to begin problem solving to make these evaluations more manageable for teachers and administrators.

- Keep *student growth gains weighted under 50%*. The current criticism and litigation of teacher evaluation focus on the use of value-added measures and their unintended consequences. Value-added measures are appealing, but unproven in high stakes personnel decision-making, since they are imprecise, unstable and can misidentify effective and ineffective teachers (AERA Council, 2015; Darling-Hammond, 2015).

- Declare a *moratorium on the use of test scores* to evaluate teachers and students, as they did in New York. A value-added score will be reported for teachers but not used during its four year moratorium (Taylor, 2015).

- When dismissal, establish an *appeals process* that offers 3<sup>rd</sup> party evaluators, reviews documentation, and assures due process. While districts may not be able to afford a 3<sup>rd</sup> party in the evaluation itself, they can ensure that the process is being followed and that no effective teachers are dismissed.

- If implementing *performance pay*, then promote school-wide pay. Performance pay has been unsuccessful in the past (Goldstein, 2014). Critics believe the high stakes climate will encourage competition rather than collaboration. School-wide merit may foster competition among schools rather than between teachers.

Since teachers tend to understand and trust principal observations (Goldring et al., 2015), then

- Teachers should attend the same *training* with principals. Professional development should include information about teaching and learning and its improvement, not just the instrument (SREB, 2015).

- Allow *local discretion* for the instrument. The thinking about teaching and learning that comes from discussions about the items of an instrument is important because there is no one best instrument that reflects a consensus about effective teaching (Kennedy, 2010), and that instrument reflects what its developers value about teaching (Sergiovanni, 1984).

- If districts want to involve others in evaluation, then use instructional coaches (free of evaluation responsibilities) to *help teachers improve* (e.g., Eady & Zepeda, 2007), especially in the content area. Not much attention has been given to teacher improvement in state policy. Instead states have focused on how to get better data. Now is the time to begin to focus on improvement.

### References

- Amrein-Beardsley, A. (2014, May 4). Breaking news: Houston teachers suing over their district's EVAAS use. Retrieved from Vamboozled blog at <http://vamboozled.com/category/lawsuit/>
- Amrein-Beardsley, A. (2015, April 10). Value-added in the courthouse. Retrieved from <http://vamboozled.com/value-added-in-the-courthouse/>
- Babwin, D. (2012, September 10). Chicago teachers strike in bitter contract dispute. Retrieved from <http://www.boston.com/news/education/2012/09/10/chicago-teachers-strike-bitter-contract-dispute/sfiCONiSqSFZ854cz2UjL/story-1.html>
- Berliner, D. (2014). Exogenous variables and value-added assessments: A fatal flaw. *Teachers College Record* 116 (1). Accessed: 8/22/2015 from <http://www.tcrecord.org> ID Number: 17293.
- Burris, C. (2015, June 4). Principal: I'm retiring because common core puts test scores before children. Retrieved from <http://hechingerreport.org/principal-im-retiring-because-common-core-puts-test-scores-before-children/>
- Bush, M. (2015, May 21). APS teachers burn their evaluations in protest. *Albuquerque Journal* Retrieved from <http://www.abqjournal.com/587949/news/aps-teachers-burn-their-evals.html>
- Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York: Teachers College Press.

- Darling-Hammond, L. (2015). Can value-added add value to teacher evaluation? *Educational Researcher* 44 (2), 132-137.
- Derrington, M.L. (2014). Teacher evaluation initial policy implementation: Superintendent and principal perceptions. *Planning and Changing* 45(1/2), 120-137.
- Eady, C., & Zepeda, S. (2007). Evaluation, supervision and staff development under mandated reform; The perceptions and practices of rural middle school principals. *The Rural Educator*, 28 (2), 1-7.
- Education Commission of the States (ECS). (2015). Teacher tenure/continuing contract data base. Retrieved June 1, 2015 from [http://www.ecs.org/html/educationIssues/teachingQuality/teacherdb\\_intro.asp](http://www.ecs.org/html/educationIssues/teachingQuality/teacherdb_intro.asp)
- EducationCounsel. (2013, January 11). An emerging state framework for designing SLO processes for use in educator evaluation. Retrieved from <http://www.educationcounsel.com/articles-archive>
- Egley, R., & Jones, B. (2004). Rural elementary administrators' views of high-stakes testing. *The Rural Educator* 26(1), 30-39.
- Goldring, E., Grissom, J., Rubin, M., Neumerski, C., Cannata, M., Drake, T., & Schuermann, P. (2015). Make room value-added: Principals' human capital decisions and the emergence of teacher observation data. *Educational Researcher*, 44 (2), 96-104.
- Goldstein, D. (2014). *Teacher wars: A history of America's most embattled profession*. New York: Knopf Doubleday Publishing Group.
- Harmon, H., Gordanier, J., Henry, L., & George, A. (2007). Changing teaching practices in rural schools. *The Rural Educator*, 28 (2), 8-12.
- Hazi, H.M. (2015a). Teacher evaluation in statute, regulation and litigation: A view of the Mid-Atlantic region with a focus on Pennsylvania. An article submitted to *Pennsylvania Educational Leadership*.
- Hazi, H.M. (2015b). Teacher evaluation in the states: An update. A paper to be presented at the annual meeting of the Education Law Association, November 6, Cleveland, OH.
- Hazi, H.M., & Arredondo Rucinski, D. (2009). Teacher evaluation as a policy target for improved student learning: A fifty state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives*, 17(5). Retrieved 3/09/09 from: <http://epaa.asu.edu/epaa/v17n5/>
- Houston Federation of Teachers et al. v. Houston Independent School District. (2014). Complaint filed in the U.S. District Court for the Southern District of Texas, Houston Division.
- Kennedy, M. (2010). Approaches to annual performance assessment (pp. 225-249). In M. Kennedy (Ed.) *Teacher assessment and the quest for teacher quality: A handbook*. San Francisco: Jossey-Bass.
- Lacireno-Paquet, N., Morgan, C., & Mello, D. (2014, March). How states use student learning objectives in teacher evaluation systems: A review of state websites. Retrieved from [http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL\\_2014013.pdf](http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/REL_2014013.pdf)
- Lasswell, T., Pace, N., & Reed, G. (2008). Weighing in: Rural Iowa principals' perceptions of state-mandated teaching evaluation. *The Rural Educator*, 29 (3), 40-44.
- Lovett, I. (2010, November 9). Teacher's death exposes tensions in Los Angeles. *The New York Times*. Retrieved from <http://www.nytimes.com/2010/11/10/education/10teacher.html?emc=tnt&tntemail1=y>
- The National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Author.
- Sawchuck, S. (2014, May 7). Florida judge: Teacher evaluation system unfair, but legal. *Education Week*. Retrieved from [http://blogs.edweek.org/edweek/teacherbeat/2014/05/florida\\_judge\\_teacher-evaluati.html](http://blogs.edweek.org/edweek/teacherbeat/2014/05/florida_judge_teacher-evaluati.html)
- Sentell, W. (2013, September 4). Teacher ratings: Only 4 percent 'ineffective.' The Advocate. Retrieved from <http://theadvocate.com/home/6956165-125/32-percent-of-teachers-highly>
- Sergiovanni, T. (1984). Expanding conceptions of inquiry and practice in supervision and evaluation. *Educational Evaluation and Policy Analysis*, 6 (4), 355-365.
- Southern Regional Education Board (SREB). (2015, February). Toward trustworthy and transformative classroom observations: Progress, challenges and lessons in SREB states. Retrieved from <http://SREB.org>.
- Strauss, V. (2014, April 30). Houston teachers sue over controversial teacher evaluation method. *The Washington Post*. Retrieved from <http://www.washingtonpost.com/blogs/answer-sheet/wp/2014/04/30/houston-teachers-sue-over-controversial-teacher-evaluation-method/>
- United States Department of Education. (2002, January). No Child Left Behind Act of 2001. *Public Law 107-110, 107th Congress*. Retrieved from

<http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf>  
United States Department of Education. (2010). *A blueprint for reform: The reauthorization of the elementary and secondary education act*. Washington, DC: Author.

United States Department of Education. (2012, February). ESEA flexibility review guidance. Retrieved June 30, 2015 from <https://www.ed.gov/sites/default/files/review-guidance.doc>

**About the author:**

Helen M. Hazi has been a teacher, a Supervisor of Curriculum and Instruction, an expert witness and a professor who prepares school administrators. She writes about teacher evaluation and about school reform in West Virginia, a rural state. If you have teacher evaluation problems, contact her at [Helen.Hazi@mail.wvu.edu](mailto:Helen.Hazi@mail.wvu.edu).