Teacher Professional Development & Evaluation: An Overview of Research & Practices

August 14, 2012

2012 Adequacy

Purpose of Report

The purpose of this report is to summarize research on effective professional learning practices that contributes to student achievement gains and on teacher evaluation approaches that inform professional development (PD) programs.

The preponderance of evidence indicates that individually targeted, ongoing PD enhances teaching effectiveness associated with student achievement gains.

The effects of PD on student learning are maximized when based on individualized developmental plans, conjointly designed by an evaluator and the teacher, that specify details about content and skills to be learned though job-embedded modeling, practice teaching, classroom observation and feedback from principals, academic coaches, and effective peer teachers.

2

Goal of Professional Development

Professional learning is designed to specifically address individual weaknesses in content and skills in job-embedded PD.
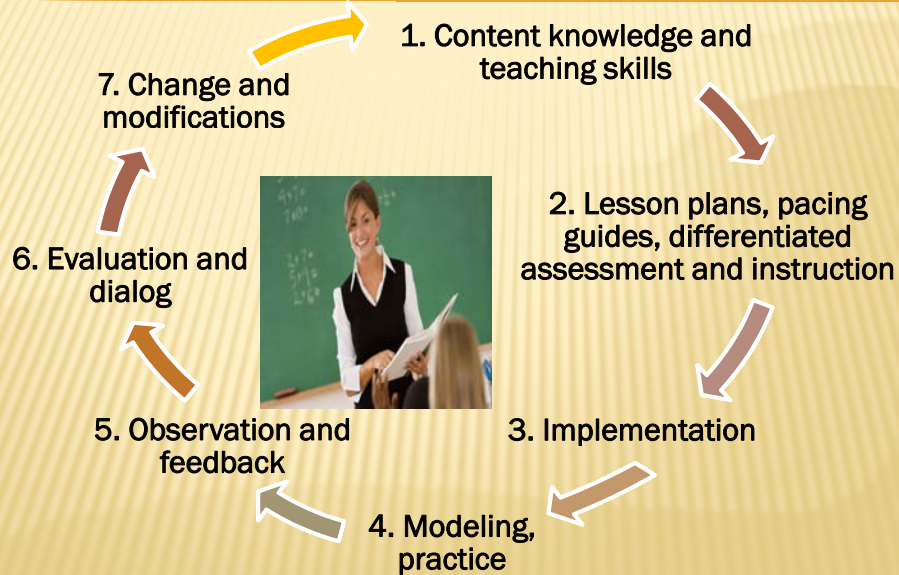
Effective Teaching Generates

Student Achievement

Context of Effective Teaching
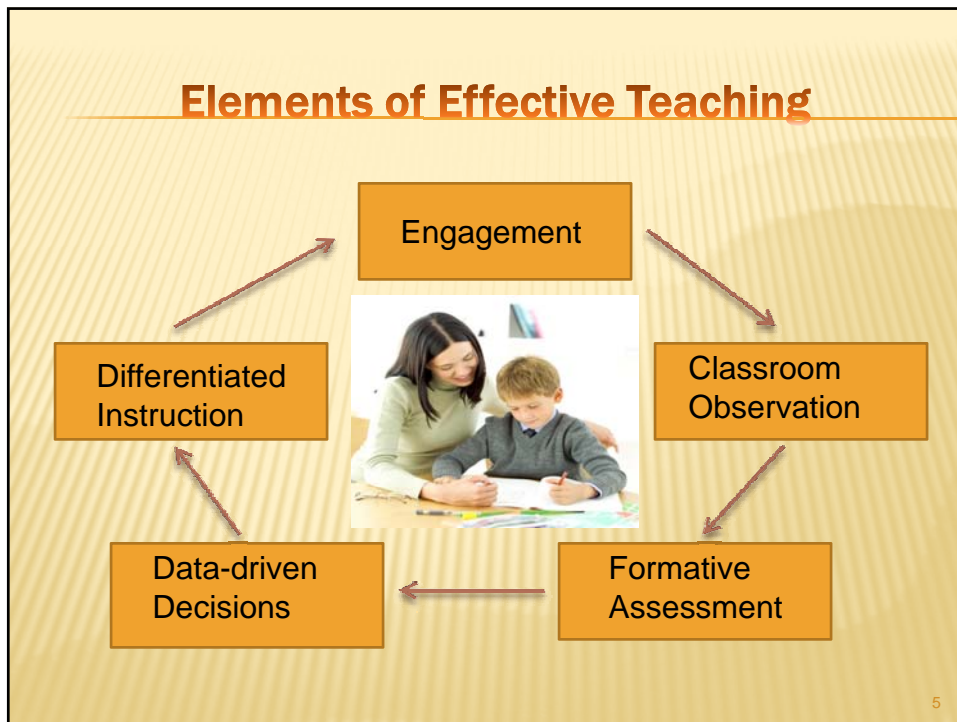
7. Change and modifications

1. Content knowledge and teaching skills

2. Lesson plans, pacing guides, differentiated assessment and instruction

6. Evaluation and dialog

3. Implementation

5. Observation and feedback

4. Modeling, practice

Elements of Effective Teaching

Engagement

Differentiated Instruction

Classroom Observation

Data-driven Decisions

Formative Assessment

5

Need for Professional Development

During most of the 20th century, the United States possessed peerless mathematical prowess, measured not only by the depth and number of mathematical specialists, but also by the scale and quality of its engineering, science, financial leadership, and even by the extent of math education in the broader population (U. S. Department of Education, 2008).

However, the average U.S. mathematics literacy score (487) on the Program for International Student Assessment (PISA) in 2009 was lower than the average scores of the 34 OECD (Organization for Economic Co-operation and Development) countries (U.S. Department of Education, 2011).

The NAEP math scores in 2011 show 40% of 4th graders in the country are proficient (grade-level) or above, whereas 35% of 8th graders are at that level. In 2011, 45% of the high school graduates who took the ACT met the Benchmark for college readiness in math; 25% met the Benchmark for all four exams (i. e., math, science, English, and reading).
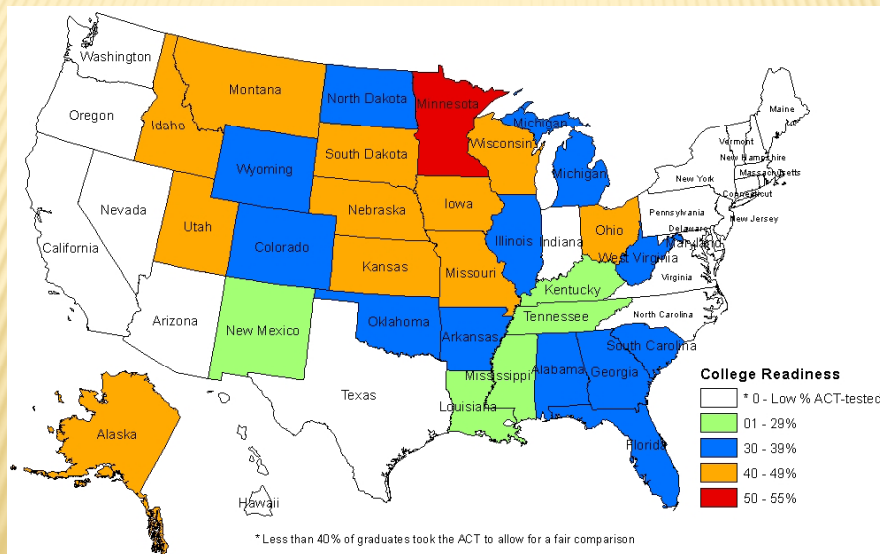
6

3

Table 1. Average Scores of 15-Year-Old Student on Mathematics 2009

OECD* Country and Average

Korea	546		New Zealand	519		Denmark	503
Finland	541		Belgium	515		Slovenia	501
Switzerland	534		Australia	514		Norway	498
Japan	529		Germany	513		France	497
Canada	527		Estonia	512		Slovak Republic	497
Netherlands	526		Iceland	507		OECD average	496
Austria	496		United Kingdom	492		Ireland	487
Poland	495		Hungary	490		Portugal	487
Sweden	494		Luxembourg	489		Spain	483
Czech Republic	489		United States	487		Italy	483
Greece	466		Turkey	445		Mexico	419
Israel	447		Chile	421			

Note: Green indicates scores significantly >U.S.; White = U.S.; Tan significantly <U.S.
Organization for Economic Co-operation and Development (OECD)
Program for International Student Assessment (PISA) scores vary 0 to 1000.

7

Percent of ACT-Tested High School Graduates Meeting Three or Four College Readiness Benchmarks by State, 2011

College Readiness
* 0 - Low % ACT-tested
01 - 29%
30 - 39%
40 - 49%
50 - 55%

*Less than 40% of graduates took the ACT to allow for a fair comparison

8

4

National Mathematics Advisory Panel Findings

➢ **Differences in teachers account for 12% to 14% of total variability in students' mathematics achievement gains during an elementary school year.**

➢ **When teachers are ranked according to their ability to produce student achievement gains, there is a 10 percentile point difference across the course of a school year between achievement gains of students of top-quartile teachers versus bottom-quartile teachers.**

➢ **Teachers must know the mathematical content they are responsible for teaching in detail, and its connections to other important elements of math, both prior to and beyond the level they are assigned to teach.**

➢**Teachers, especially below high school, do not know enough math to teach it. A college major, or even a concentration, in math is not required to teach math in virtually all states.**

9

Effective Professional Development

Effective instruction requires a comprehensive, in-depth knowledge of content taught, a thorough understanding of different learning styles, and an array of teaching skills to present complex ideas to a diverse group of learners.

In an effective professional learning system, leaders and teachers work with academic coaches, highly skilled peers, and outside experts when needed, to create a culture, structures (e.g., regular meetings), and educator dispositions that promote ongoing professional learning of individual teachers by targeting their specific needs.

Several narrative reviews and meta-analyses have identified key elements of effective professional learning (National Staff Development Council, Stanford University, 2010). They include the following:

10

Effective Professional Development

➢ **Effective PD is focused on specific content knowledge and effective teaching skills for students with diverse needs and abilities.**

➢ **PD is more effective when it is not approached in isolation (e.g., single-session workshops) but as a coherent part of the school's educational program.**

➢ **For substantial change to occur, curriculum, assessment, standards, teacher evaluations, and professional learning should be seamlessly linked.**

➢ **Effective PD is designed to engage teachers in active learning through modeling, practice teaching, observational feedback, and consultation with academic coaches, highly qualified peers, and outside experts.**

➢**Together, teachers and evaluators should develop a coherent, ongoing, sequential plan for acquiring content knowledge and teaching skills that is individually tailored for each teacher.**

11

Effective Professional Development

➢ **Content knowledge and teaching skills are more fully developed by making PD a priority, and dedicating a regularly scheduled time for reflection and discussion among teachers.**

➢ **Effective PD is based on systematic summative teacher evaluations that detail specific strengths and weaknesses, and a coherent developmental plan for remedying deficiencies.**

➢ **Research is clear that professional learning communities (PLCs) that actually provide learning experiences optimize the acquiring of content knowledge and instructional skills.**

➢ **Generally, existing evidence indicates that single-session, fragmented workshops have little, if any, positive impact on teaching or student achievement (Wei et al., 2010). Workshops often introduce innovations and needed expertise, while PLCs function to implement ideas and skills through modeling, practice teaching, and feedback.**

12

Effective Professional Development

➢ **PLCs provide a forum for teachers to discuss, reflect upon, and practice newly acquired ideas and skills in their specific setting, sharing individual experiences and understanding with one another. Teachers can learn to emulate successful colleagues.**

➢ **Collaboration among teachers with different content areas and skill levels, and divergent types of experiences, can capitalize on strengths of each other, compensate for individual weaknesses, increase shared knowledge and skill, and collectively enhance their school's overall professional capacity.**

➢ **Collaboration promotes mutual problem-solving, and the creation of a shared technical language and agreement on sound practices.**

➢ **Professional learning is greatly enhanced by effective teacher evaluation systems.**

13

Methods of Teacher Evaluation

By far the most common method of teacher evaluation involves observing a teacher's classroom instruction, typically structured by a protocol (or framework) that measures performance on an assessment metric, increasingly the use of scales such as Danielson's *unsatisfactory, basic, proficient*, and *distinguished*.

In addition to structured classroom observations, teacher performance ratings in most states, including Arkansas, are also based on "artifacts" or indicators (e.g., lesson plans, PD participation, samples of student work, formative and summative assessments, and prior teacher evaluations) of student and teacher performance. In Arkansas, artifacts are to be selected conjointly by the teacher and evaluator (Act 1209).

The ADE is in the process of training evaluators in incremental stages of learning, and one of the most noteworthy aspects of the new teacher evaluation system is its <u>direct</u> linkage to professional learning and development.

14

7

Arkansas Teacher Evaluation

In concert with the teacher, Arkansas evaluators are required to develop PD plans in which at least half of the 60 hours required by law are directly related to the teacher's area of teaching and identified teaching needs. For teachers in intensive support status, all PD hours, except those required by law, must be directly related to the individual teacher's needs (Act 1209).

Teachers who receive an unsatisfactory rating are placed in an intensive support status. Along with other teachers who are struggling with classroom instruction, these teachers have access to online video vignettes that model effective teaching in each domain.

A teacher who does not satisfactorily accomplish the goals and tasks of the intensive support status is recommended for dismissal or nonrenewal of the teacher's contract by the superintendent based on documentation of intensive support efforts (Act 1209).

15

Critique of Classroom Observation & Artifacts

In examining the literature, it appears that each approach to teacher evaluation offers advantages over other methods, as well as weaknesses that make it less desirable than alternative procedures.

For example, the traditional approach of principals and other evaluators observing instruction in the classroom has the advantage of direct observation of teaching and of student participation and interactions with one another and with the teacher.

Classroom observation evaluations can allow for unique, individualized approaches to teaching, and nuances in skills that go undetected in other methods of teacher evaluation, such as primary reliance on portfolios and formalized measures of student achievement gains.

Supplemented with artifacts (e.g., student work and assessments, PD activities, lesson plans and assignments, committee work, community service), the traditional classroom observation approach to evaluation offers direct indicators of a breadth of diverse teacher responsibilities.

16

Critique of Classroom Observation & Artifacts

Teacher evaluations based on classroom observations and artifacts rely on professional judgments, which are susceptible to subjective inferences and personal preferences. This subjectivity is regulated to an unknown extent by protocols and training, but professional judgments by nature are "black boxes" in terms of retracing the steps involved, the algorithm used to weight factors considered, and the validity of the ratings.

Although training is indispensable in trying to achieve uniformity in teacher evaluation ratings, rarely is inter-rater reliability examined to determine the extent of uniformity achieved among evaluators.

Failure to determine inter-rater reliability also plagues the second primary method of teacher evaluation, which entails extensive or exclusive use of portfolios. Failure to examine reliability of portfolio ratings is problematic because several diverse items are assigned differential rating weights.

17

Portfolio Teacher Evaluation

Contents of portfolios vary across districts nationally, but often include lesson plans, pacing guides, student work and assessments, PD activities, formative and summative evaluations, relevant parental comments, awards and recognitions, and community service. In states like Arkansas, portfolios are one of several artifacts that can be presented in the evaluation process.

Portfolios cover a broad range of teacher responsibilities and allow the teacher to present diverse indicators of performance in addition to classroom instruction, including student achievement measures.

The portfolio process often requires teachers to reflect on why artifacts were included and how they relate to particular standards. They may contain exemplary work as well as evidence that the teacher is able to reflect on lessons, identify problems, and make modifications for future lessons.

18

Value-Added Modeling

Finally, the most quantitative approach to teacher evaluation is "Value-Added Modeling" (VAM), which focuses directly on growth in student achievement scores.
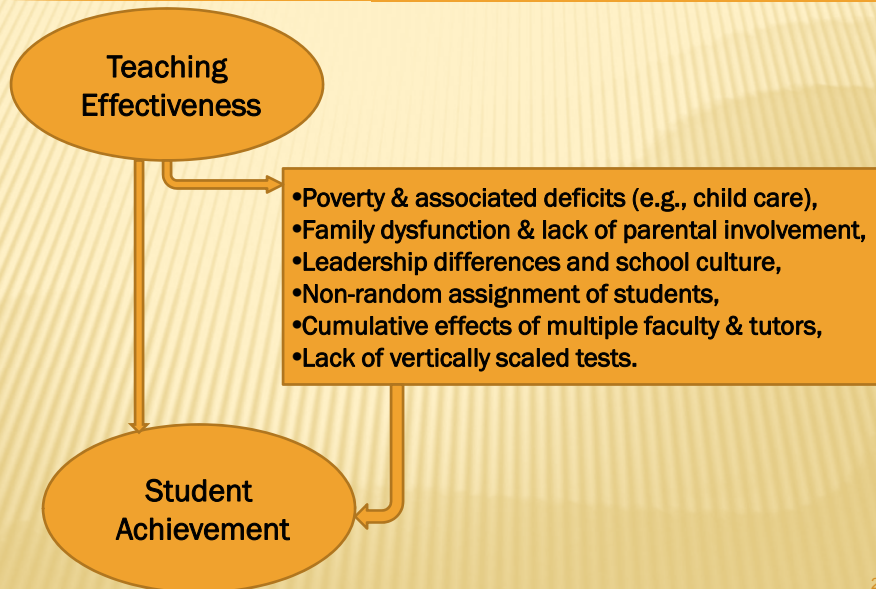
VAM entails elegant or sophisticated regression statistics that examine very complex interrelationships between interventions (e.g., PD), extraneous influences (e.g., student and community characteristics), and outcomes (e.g., student achievement gains).

Conceptually, VAM's promise of quantifying the "added value" that teachers produce in terms of student learning offers intuitive appeal because it is more objective than professional judgments of classroom teaching or portfolios.

Methodologically, VAM is attractive because it appears to offer a way to disentangle the "effects" of teaching from those of uncontrolled (or extraneous) factors, such as poverty and associated lack of resources, family influences, cumulative effects of multiple teachers shown on the next slide.

19

Model of Student Achievement

Teaching Effectiveness

•Poverty & associated deficits (e.g., child care),
•Family dysfunction & lack of parental involvement,
•Leadership differences and school culture,
•Non-random assignment of students,
•Cumulative effects of multiple faculty & tutors,
•Lack of vertically scaled tests.

Student Achievement
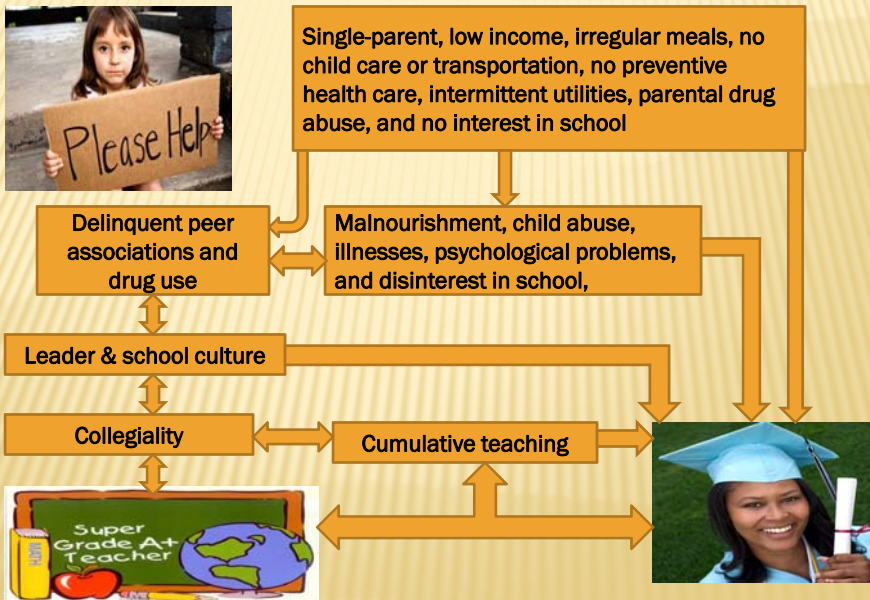
20

10

Weaknesses of Value-Added Modeling

However, despite its conceptual and methodological appeal, the use of VAM to estimate teacher effectiveness for high stakes purposes poses daunting and seemingly insurmountable challenges stemming from many factors, including but not limited to:

1) unavailable or inadequate measures of key influences; 2) use of tests of grade-level performance instead of vertically scaled tests measuring learning gains, 3) difficulties of disentangling the contributions to learning of multiple teachers, 4) effects of school culture, curriculum, resources, and leadership, 5) confounding effects of student characteristics (e.g., ability, motivation), and 6) complications of community characteristics (e.g., poverty, crime).

The complexities and interactive effects of these factors have not been fully analyzed and are not well understood. Too many of the potent influences are not represented in analyses owing to a lack of measures (e.g., cumulative effects of multiple teachers) or inadequate measures (e.g., poverty).

21

Hypothetical Model of Interactions

Single-parent, low income, irregular meals, no child care or transportation, no preventive health care, intermittent utilities, parental drug abuse, and no interest in school

Delinquent peer associations and drug use

Malnourishment, child abuse, illnesses, psychological problems, and disinterest in school,

Leader & school culture

Collegiality

Cumulative teaching

22

Examination of Teacher Evaluation Methods

Based on an extensive review of VAM research, scholars convened by the Economic Policy Institute (Baker et al., 2010, p. 9) concluded: *"Although value-added methods can support stronger inferences about the influences of schools and programs on student growth than less sophisticated approaches, the research reports cited….have consistently cautioned that the contributions of VAM are not sufficient to support high-stakes inferences about individual teachers.*

While VAM is too imprecise for use in high-stakes personnel decisions, these highly systematic statistical procedures do provide objective estimates of the relative effects of interventions and extraneous influences on outcomes such as student achievement gains.

Presently, it seems judicious to conclude that each of the teacher evaluation methods discussed, while suffering important limitations, offers useful information for evaluation, professional development, and strategies for improving student achievement.

23

Arkansas Teacher Evaluation

Therefore, the wisest approach to evaluating teachers may be to blend different approaches, as done in Arkansas, to get a more robust and multi-faceted understanding of a teachers' influence on student learning and achievement.

Some of the salient advantages of the Arkansas Teacher Evaluation and Support System include the use of a well-established framework for teaching (The Danielson Group, 2011), with empirically-grounded rubrics and performance ratings.

This system links teacher evaluation to PD, intensive support, and employment decisions. Also, recommended artifacts, half of which must be external measures, cover a wide range of teacher responsibilities in addition to classroom instruction, such as PD participation and collaborative research. This diversity of performance data comports well with the criticism directed at using "high stakes" testing as the sole measure of teacher effectiveness.

24

12

Arkansas Teacher Evaluation

ADE is devoting considerable resources and time to training evaluators, involving multiple sessions to ensure a thorough understanding and consistency among evaluators.

A final noteworthy aspect of Arkansas Teacher Excellence and Support System is the use of online video vignettes to model effective teaching in each domain, which is especially useful for teachers who are struggling with instruction.

25

For further information contact:

Dr. Brent Benda
Bureau of Legislative Research
bendab@blr.arkansas.gov
(501) 537-9146

Bureau of Legislative Research
Policy Analysis & Research Section

ARKANSAS

Appendix A

Table 1. Percent of Professional Development Provided by Source

Statistic	AETN	ADE	Co-op	District	District/ AETN	Contract	Other
Statistics							
Mean	2.98	7.96	28.88	45.69	0.61	4.41	2.48
Median	1.00	5.00	25.00	50.00	0	0	0
Std. Deviation	4.34	8.37	21.85	25.85	2.70	9.50	9.19
Minimum	0	0	0	0	0	0	0
Maximum	30.00	57.00	94.00	100.00	100.00	75.00	100.00
Percentiles							
20 Percentile	0	0	10.00	20.00	0	0	0
40 Percentile	1.00	5.00	20.00	41.80	0	0	0
60 Percentile	2.00	10.00	31.00	53.80	0	0	0
80 Percentile	5.00	15.00	45.00	70.00	0	10.00	0
Number of Districts							
Missing Data	43	1	0	1	70	63	94

Note: Missing data are number of districts, not percentages as reported in the rest of the table.

27

Appendix B

Table 2. Number of Formative Evaluations in a School Year

Teacher's Experience	< 2 Years of Experience	3 or > Years of Experience
Mean	3.75	3.21
Median	2.00	1.00
Standard Deviation	6.52	6.48
Minimum	0	0
Maximum	36	36
20th Percentile	1.00	1.00
40th Percentile	2.00	1.00
60th Percentile	2.00	2.00
80th Percentile	4.00	4.00

28

14

Appendix C

Table 3. Type of Summative Evaluation used in Arkansas

Type of Evaluation	Number of Schools	Percent
Open Format	2	2.7
Checklist	17	23.0
Protocol with open comments	6	8.1
Protocol with rating scale	40	54.1
Other	9	12.1

Appendix D

Table 4. Number of Teacher Separations from School Due to Performance in the Past 3 Years

Number of Teachers	Number of Schools	Percentage
0	51	68.9
1	13	17.8
2	6	8.1
3	3	4.1
4	1	1.4

Note: Number of certified teachers separated from the district for performance issues in the past 3 years.

National Teacher Evaluation Systems

http://www.arkleg.state.ar.us/education/K12/TeacherIssueDocuments/
12-097_Public_NationalTeacherEvaluationSystem_EdComm_BB_08-
14-2012.pdf