



Albuquerque Teachers Federation

Teacher Evaluation, Compensation and Student Outcomes

How can we link these components into a fair, reliable system that will change public education for the better?

Measuring What We Value

We should measure what we value, not the other way around. All too often in education, we value what is easy to measure, but overlook elements that are essential for effective practice. If we value outcomes for our students, such as the ability to understand essential concepts, work in groups, think critically, and solve problems, we must measure those abilities.

If we value those attributes in student learning, then we are obligated to value the same qualities in teaching. A teacher evaluation system that includes an over-reliance on standardized measures undervalues—to the point of ignoring—the exact outcomes we need for our students and must value in our teachers.

Teachers and their unions are often characterized as being against reform. We are *not* against reform; we *are* against slogan reform. Slogans like “prioritizing student academic gains” and “recruiting, recognizing, and retaining ‘rock star’ teachers” translate into shallow reforms that result in an emphasis on superficial teaching and learning. More than just being a distraction, slogan reforms do harm to students.

We believe that changes in the teacher evaluation system can be a powerful tool for positive change. High-quality, thoughtful evaluation systems can change public education for the better. Misguided evaluation approaches, like those primarily based on standardized test scores, will not. The effect on schools—and most importantly, on the students we serve—will be harmful.

If New Mexico succumbs, as many states have done, to quick-and-easy slogan reforms like the proposed evaluation system based at least 50% on standardized tests, the worst outcomes of this ill-conceived and unproven reform, will happen:

- Increasing reliance on inadequate measures of student learning.
- Creating incentives that focus on the wrong factors.
- Entrenching an educational system that runs contrary to a focus on deep learning, which is the intent of the Common Core Standards that New Mexico has adopted.
- Maintaining a system that creates judgments based on scarce evidence and which never really gives teachers meaningful feedback.
- Basing “effective” teaching on simply teaching to a bad test.

We know our evaluation system needs reform. Our pass/fail evaluation system provides little or no feedback that a teacher can use to improve; nor does it correctly identify those who are struggling or ineffective. However, if the current administration succeeds in its attempts to revamp our evaluation system without the experience and wisdom of the practitioners, teaching will not improve and student learning will not advance. Our union has a better way.

We know that evaluation serves multiple purposes, such as:

- Improving the overall quality of the teacher workforce by identifying and building upon individual and collective teacher strengths.
- Improving instruction and other teacher practices to improve student learning.
- Identifying exemplary teachers.
- Identifying ineffective teachers and developing a system of support to remediate their skills.
- Ensuring fair and valid employment decisions.
- Confirming what teachers are doing well.

Under “No Child Left Behind” (NCLB), successful educators are those who produce test score gains—not learning gains. We must put an end to this false definition of success for both teachers and students, and stop the widening gulf between what we say we value and what we measure. In addition, there is no evidence that teachers are motivated to improve if they are evaluated or compensated using test scores. There is ample evidence that the over-reliance on standardized tests has many well-documented consequences, such as:

- Excessive test preparation.
- A narrow curriculum.
- An arbitrary and inordinate focus on students who are on the cusp of proficiency as measured by standardized tests—a focus that undercuts the right of all students to a well-rounded, content-rich curriculum that encourages them to develop critical thinking skills.

Creating a system that includes appropriate measures of student learning takes real thought, as well as time to develop them and to learn how to implement them appropriately. When slogan reform takes the place of that time and effort, the impact on learning is disastrous. When contrived and unproven reforms like “merit pay” are added to a flawed testing and evaluation system, the impact is worse.

Evaluation of the complex task of teaching requires well-reasoned, multiple measures. We know what those measures should be. Now is the time for us to stand together to educate the rest of those whose knowledge comes from slogan reform. Our request—our conviction—is *With Us, Not To Us*. We stand ready to be partners in *real* reform.

Background & Context

“Race to the Top” has replaced the NCLB emphasis on “highly qualified” teachers with “highly effective” teachers. Some have assumed that teacher effectiveness can be measured by linking student test scores to teachers’ evaluations. The problem is that it can’t be.

The accountability tests that currently have such a profound impact on our nation’s schools are instructionally insensitive. Student standardized test scores correlate well with things like a student’s socioeconomic status and whether or not their parents read to them as young children, but not well with any measure of teacher quality, unless “teacher quality” is defined circularly, measuring it only with standardized test scores.

Experts confirm that while there are good reasons for concern about the current system of teacher evaluation, there are also serious reasons to be concerned about claims that measuring teachers’ effectiveness largely by student test scores will lead to improved student achievement.

With Us Not To Us

Teachers and their union leaders know the current evaluation system is ineffective and needs improvement. Good teaching is much more than a student test score, yet teacher evaluation cannot ignore the importance of student learning. As teachers, we embrace our professional accountability. We are ready for our evaluation system to take into account the outcomes of our work, as long as:

- Our collective expertise is included in meaningful and ongoing ways.
- The student outcomes actually indicate what students know and are able to do.
- The evaluation system is fair and the multiple measures of teachers’ work are valid and reliable.
- The measures of student learning are fair and instructionally-sensitive.

As unionized professionals, of everything within our control, nothing matters more than the knowledge, skills and effectiveness of our members.

Teacher Evaluation—the ATF Way

What Should an Evaluation System Look Like?

ATF believes an evaluation system should:

- Be based on a set of standards of practice that takes into account the complexities of teaching.
- Be both formative and summative.
- Include evidence of teaching and student learning from multiple sources.
- Be subject to continuous updating of instruments and processes as research

- on practice leading to valued student outcomes becomes available.
- Focus on providing continuous professional development and growth for teachers by addressing the skills, knowledge and needs of teachers depending on where they are on a career continuum (e.g., novice, mid-career, veteran).
- Promote teacher leadership.

What we have now is an inadequate teacher evaluation system with “drive-by” evaluations where teachers are observed in an often meaningless and cursory way by principals and where virtually all teachers meet competencies.

What does the current governor want for New Mexico’s future? She proposes a fundamentally-flawed system, where teachers are arbitrarily ranked largely based on results from students’ multiple-choice standardized test scores that have been subjected to opaque statistical procedures referred to as value-added methods (VAM).

What our union proposes is to truly improve teacher evaluation by using multiple measures that inform teaching and learning. In order to understand the absolute necessity of using multiple measures to evaluate teaching, you need to understand what value-added methods are and what they are not.

Value-Added Methods (VAM)

In the current national craze, identifying “highly effective” teachers using a “performance-based” system relies heavily on standardized test scores. Most educators and some policymakers know to avoid an over-reliance on a single measure, especially one that inadequately measures what students know and are able to do, and even less effectively identifies the impact of the teachers. As a result, many have turned to VAM.

The basic concept in VAM is that teachers should add “value.” In their simplest form, VAM compare data at 2 points in time rather than at a single point in time—in theory, this measures growth. In practice, in order to measure growth and determine the impact of a teacher, mathematical formulas are applied to each student’s anticipated trajectory in an effort to adjust for out of school factors. There is no one VAM—each district and state has created their own formula. There is broad agreement among statisticians, psychometricians, and economists that even the most sophisticated value-added measures are not reliable or valid to use for teacher evaluation.

A Move to Multiple Measures

Because VAM are not yet sufficiently developed to measure the impact of a teacher, and because we have yet to develop excellent performance-based assessments that would truly measure student learning, many districts/states are working to create evaluation systems with multiple measures. Does this sound easy? Well, it isn’t; nor should it be. Teaching is a complex endeavor and measuring whether a teacher is

effective with a classroom of 28 or so individual students in any given year is a complex task. That’s why it requires multiple measures—that is, assessing student learning and a teacher’s ability using more than one tool. What are those measures? Here’s what we think.

Measure 1: Classroom Observations with Substance – 40%

More People

Principals are essential, yet not sufficient, in assessing teacher competence. It is time for New Mexico to diversify our system by requiring that observations, an essential part of a teacher evaluation system, include instructionally-focused accountability. This can be achieved by ensuring that teachers are observed in classroom instruction several times a year by multiple trained observers, including principals and master teachers. Evaluators would be trained and certified. A system would be put in place to monitor the reliability and consistency of evaluations.

Certified Evaluators:

- Must be able to accurately assess teacher practice.
- Must be able to communicate effectively.
- Could help teachers establish their PDPs.

Better Observation Tools

In the APS School Improvement Grant (SIG) pilot, an observation tool has been developed that will provide a more comprehensive and useable picture of teachers’ ability. We have developed a rubric by combining Charlotte Danielson’s 4 *Framework for Teaching* domains with the 9 NM Teacher Competencies—a set of clear expectations for what effective teaching looks like along a continuum of growth.

For example: In Danielson’s Framework Domain 1(Planning and Preparation), Element 1 is “Knowledge of Content”. This corresponds directly with NM Competency 1 (*The teacher accurately demonstrates knowledge of the content area and approved curriculum.*), indicators C and D.

More Differentiation

When a teacher is evaluated, he or she will be given a performance rating based on the indicators in each domain using rubrics for several dimensions of instructional quality.

The four domains are:

- Designing and planning instruction
- Classroom learning environment
- Instruction
- Teaching responsibilities

Domain standards are scored during classroom observations, except the “teaching responsibilities” domain, which is assessed at the end of the year.

Each domain has several elements (we currently call those indicators) and for each element under the domain a teacher is rated based on a 4-category system:

- Unsatisfactory
- Basic
- Proficient
- Distinguished

Frequency

Teachers should have a summative evaluation every three years, with formative evaluations taking place yearly. The principal would observe teachers during their summative years, while only the evaluation team would observe during the formative years. By establishing this system, there will be the time necessary to make the evaluation process meaningful and teachers will be able to address and develop their teaching practices in areas in which they may be weak, preferably improving their teaching skills before the summative evaluation.

Any time a teacher is unsatisfactory in any given element or indicator, he or she would be given time and support to improve. A teacher who is unsatisfactory in a domain would be placed in a Peer Assistance and Review (PAR) program. Once in PAR, the teacher is given intensive support, and is either able to improve or is dismissed from the profession.

Measure 2: Student Outcomes – 30%

The increased awareness of how much teaching matters, along with widespread concerns about the severe limitations of using standardized testing as a way to measure teaching and learning, has led to new approaches that connect student growth to teacher performance. One proven strategy that's gaining attention is to include Student Learning Objectives (SLOs) in teacher evaluation systems. A number of school districts and states are adopting or considering SLOs for both tested and non-tested subjects.

The process of creating SLOs supports teachers to think critically about their teaching objectives and how they can be reliably measured to indicate student growth. SLOs are designed to both strengthen teaching and improve student learning. When implementing SLOs, teachers and administrators engage in a collaborative process that helps teachers analyze their own practice and adjust instructional strategies to better meet student needs.

Using SLOs, teachers must determine how the student can best demonstrate his or her abilities, whether through a group or individual exercise, in-class or take-home test, or another method. SLOs require teachers to do the thinking up front about the skills and conceptual understandings students need to exhibit. SLOs also allow students to participate as active, independent learners. Teachers communicate

clearly with students what they are expected to be able to do, so they can also assess their own progress and concentrate on their weaker skills. Using SLOs, teachers and students are better able to see how the learning activities relate and build upon each other.

SLOs can make teacher assessment of student performance and student self-assessment easier. Since learning objectives state a performance goal, teachers can more easily develop a method for assessing that performance. Student learning should include evidence of growth in knowledge and skills, based on multiple measures.

SLOs are annual targets for growth that a teacher sets at the beginning of the year and strives to attain by the end of the year (or at the end of a grading period, if appropriate). Classroom-based assessments, including those that are program-specific, are more instructionally-sensitive, thus creating a more direct link between the progress of the student and the ability of the teacher.

In addition, SLOs are:

- Aligned with the school's standards and curricula.
- Based on outcomes, not activities.
- Rigorous and measurable.
- A way to incorporate an individual teacher's contribution to student learning into an evaluation system, and are:
 - ✓ Teacher-created
 - ✓ Student-specific
 - ✓ Instructionally-sensitive
- Applicable to all content-area teachers, as well as support personnel.
- Applicable and appropriate for Special Education teachers and relate directly to IEP goals.

SLOs combine objectives, assessment and instructional strategies:

- Objectives articulate the knowledge and skills teachers want students to acquire by the end of the unit/course.
- Assessments allow the teacher to check the degree to which the students are meeting the learning objectives.
- Instructional strategies are chosen to foster student learning towards meeting the objectives.

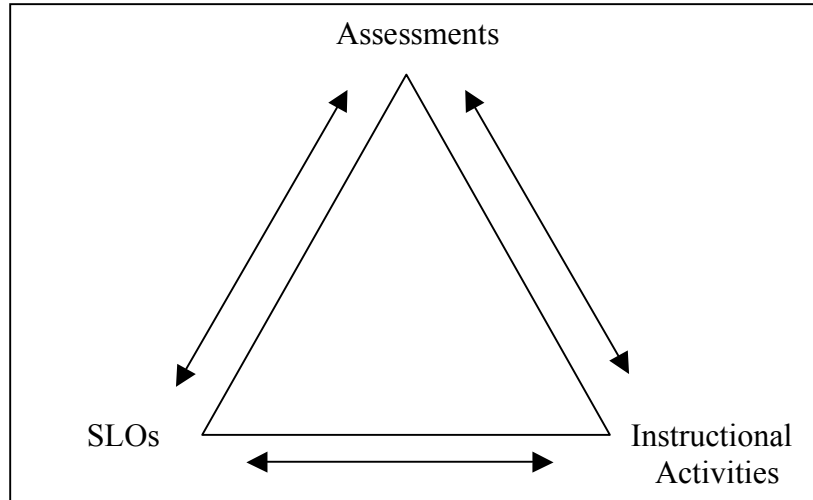


Figure1. Interrelatedness of SLO Components

Teachers would be provided a framework, or rubric, for creating rigorous SLOs. In the APS School Improvement Grant (SIG) pilot, we are recommending starting with the work of Charlotte Danielson as outlined in her book, *Framework for Teaching*. Her well-respected and often-used work in the domains of teaching includes a section on *Selecting Instructional Goals*.

Danielson includes the following four elements of instructional goals:

- **Value:** Goals represent high expectations for students; and reflect important learning and conceptual understanding, curriculum standards, and frameworks.
- **Clarity:** Goals are stated in terms of student learning and permit sound assessment.
- **Suitability for Diverse Students:** Goals reflect needs of all students in a class.
- **Balance:** Goals represent opportunities for different types of learning (for example, thinking as well as knowledge) and coordination or integration, either within or across disciplines.

The rubric below is an example of how SLOs can be assessed for appropriateness and rigor as they are created by teachers and reviewed by administrators.

Element of Instructional Goals	UNSATISFACTORY	BASIC	PROFICIENT	DISTINGUISHED
Value	Goals are not valuable, and represent low expectations or no conceptual understanding for students. Goals do not reflect important learning.	Goals are moderately valuable in either their expectations or conceptual understanding for students and in importance of learning.	Goals are valuable in their level of expectations, conceptual understanding, and importance of learning.	Not only are the goals valuable, but the teacher can also clearly articulate how goals establish high expectations and relate to curriculum frameworks and standards.
Clarity	Goals are either not clear or are stated as student activities. Goals do not permit viable methods of assessment.	Goals are only moderately clear or include a combination of goals and activities. Some goals do not permit viable methods of assessment.	Most of the goals are clear, but may include a few activities. Most permit viable methods of assessment.	All the goals are clearly written in terms of student learning, and permit viable methods of assessment.
Suitability for Diverse Students	Goals are not suitable for the class.	Most of the goals are suitable for most students in the class.	All the goals are suitable for most students in the class.	Goals take into account the varying learning needs of individual students or groups.
Balance	Goals reflect only one type of learning and one discipline or strand.	Goals reflect several types of learning but no effort at coordination or integration.	Goals reflect several different types of learning and opportunities for integration.	Goals reflect student initiative in establishing important learning.

Using student learning objectives will ensure that we are measuring what we value—deep learning and conceptual understanding. In addition, teacher collaboration becomes more meaningful and directly tied to classroom work as they co-create and debate SLOs that will be appropriate, rigorous, and are directly connected to their instruction.

Teachers create student learning objectives and assess student progress toward these goals as a normal part of their instructional program. SLOs would formalize this practice, deepen it, and attach it to a teacher evaluation, making their ability to assess learning an integral part of how they are assessed as teachers.

There is no evidence that tying teachers' evaluation to standardized test results improves practice, nor is there proof that it enhances learning. Using what we already know about effective practices would immediately and substantially improve student learning.

Measure 3: Collective Responsibility – 10%

The EPSS is currently a PED requirement for every school. By including this component as one of the multiple measures in an evaluation system, the EPSS moves from the shelf and into the hands of teachers, making the goals more meaningful and establishing the EPSS as an essential component of a school staff's collaborative efforts.

This measure addresses collaboration, recognizing that, in a comprehensive system, teacher development and evaluation must foster collective responsibility and accountability to the school's goals, as outlined by the EPSS, as well as individual responsibility and accountability on the part of each teacher. Indeed, collaborative teams of educators working to improve teaching and learning are doing some of the most effective reform. Collaboration can also be measured for individual educators through the school responsibilities they choose to take on.

Measure 4: Student Feedback – 20%

Based on the research of Ron Ferguson at Harvard, many states and districts are including student feedback as one of the measures in a comprehensive teacher evaluation system. Surveys have been developed and refined over ten years with more than 400,000 students completing surveys across more than 50 districts. Dr. Ferguson's work on student surveys is also included in the APS SIG pilot.

The Tripod refers to the three points of content knowledge, pedagogical skill and relationship-building skills. The student surveys measure multiple domains of teaching practice and student engagement. Data-collection instruments and procedures are tailored for grades K-2, 3-5 and 6-12.

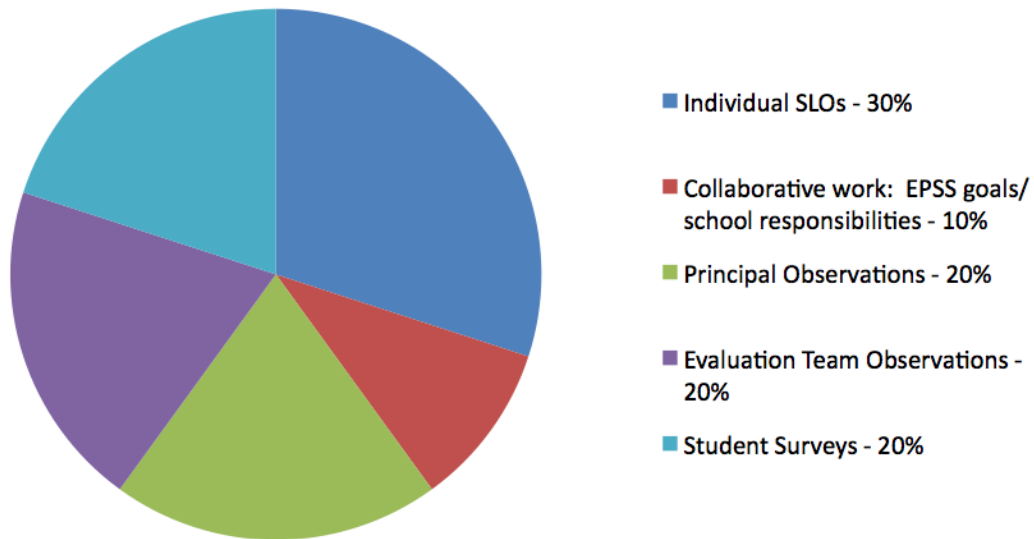
The questions are crafted carefully to get at each point of the Tripod while skillfully avoiding questions that might lead to popularity contents. For example, "My teacher explains difficult things clearly" or "Our class stays busy and does not waste time" are typical questions asked of students.

There are many ways that student perspectives can be used to provide useful information about patterns of teaching effectiveness, ideally in combination with other measures. Dr. Ferguson asserts the importance of using *multiple measures, multiple times*.

Based on the project's work, Ferguson is convinced that, "Student perceptions of classroom practice can help in predicting learning outcomes and should be taken seriously by policymakers and educators."

Tripod surveys provide reliable, detailed and validated insights on teaching and learning in any classroom, including classrooms not tested with standardized exams.

Evaluation Components



Peer Assistance and Review

Allowing ineffective teachers to stay in the classroom is detrimental to students, teachers, their unions and the profession. Teacher unions have a responsibility to teachers, the teaching profession and students to ensure that all teachers meet high professional standards of practice. ATF, and many other teacher unions, have negotiated a system known as Peer Assistance and Review (PAR).

The PAR Program is an intervention program designed to help improve the performance of experienced teachers who are having serious difficulties in the performance of their professional responsibilities. This help is provided through peer assistance from a Consulting Teacher (CT). The CT works directly with the struggling teacher to provide constructive and intensive intervention. The goal of the PAR Program is to develop and maintain the highest caliber teaching staff.

A Consulting Teacher (CT) is an experienced Mentor Teacher who is released part-time from the classroom. The CT is a professional colleague who provides non-evaluative support for teachers who are on an intensive evaluation plan or an improvement plan.

Unions that have negotiated PAR programs, including the Albuquerque Teachers Federation, have universally reported that teachers hold their peers to higher standards than the traditional evaluation system did.

ATF's Position on Compensation

Teachers should be paid well for teaching, and evaluation must have as its primary goal strengthening the individual and collective practices of teachers and schools to improve student learning. Teaching is broad in its scope of responsibilities, all of which must be taken into account within a compensation system. It is time to move away from the traditional single salary schedule to a differentiated compensation system that:

- Rewards both expertise and extra time.
- Values teachers for their expertise and not just for quasi-administrative work.
- Is tied directly to a legitimate evaluation system. All movement must be contingent upon satisfactory evaluations.
- Builds on the New Mexico 3-Tiered Licensure System.
- Includes a PAR program that provides intensive help for those struggling or, if necessary, counsels them out of the profession, which must be an integral part of any evaluation/compensation system.

Teachers and their unions have traditionally opposed merit pay plans. Merit pay is a compensation system that links the salaries of employees to evaluations of their performance. Although merit pay plans profess that they recognize and reward the best teachers, these plans usually require individual teachers to compete against each other for a limited pool of funds and the criteria for merit have either been subjective, based upon the principal's discretion, or are based on standardized testing that is not a reliable nor valid way to measure a teacher's expertise and worth.

We believe that by building on the 3-Tiered System, we will be able to:

- Emphasize student achievement based on the understanding that teachers' knowledge and skills are the key ingredients of teacher effectiveness.
- Appropriately emphasize classroom teaching by rewarding teachers who want to stay in the classroom.

We propose a compensation system that compensates an educator for job-related knowledge and skills, as well as for added responsibilities. We believe that a good compensation system will support teachers in their efforts to improve student achievement, as well as provide them with motivation for their own professional growth. Furthermore, we believe that this system will provide opportunities for teachers to take on additional responsibilities and to serve as instructional leaders.

Tying it All Together: Building on the 3-Tiered Licensure System with a Teacher-Weighted Formula

It is important to note that New Mexico's 3-Tiered System:

- Is unique
- Is a high-stakes system
- Currently includes both student learning and teacher evaluations

As teachers, we want to be recognized for experience, education, and additional responsibilities, but the current Training & Experience (T&E) Index is not adequate. We need a funding formula that ensures the minimum salaries in the 3-Tiered System do not also become maximum salaries. We propose adapting the T&E Index so that it's similar to the way we fund students—by assigning teacher attributes a weight according to a formula.

In a Teacher-Weighted Formula, each Level II and Level III teacher is weighted by a formula based on experience, additional degrees, state endorsements and additional responsibilities. This differentiates between new Level II and III teachers and experienced Level II and III teachers.

By changing the T&E Index to a Teacher-Weighted Formula, school districts will truly have a career ladder of opportunity within the licensure levels. This system will continue to honor education and experience for all teachers by adding a weighted formula for each Level II and III teacher. It will also provide incentives and pay growth potential if a teacher wants to stay at Level II for an entire career. In addition, it will allow pay increments for extra work or assignments for teachers if they choose to take on additional responsibilities.

The New Mexico School Finance Act must be modified to provide the financial under- girding for the salary structure created in the Education Reform Act of 2003—the 3-Tiered Licensure System. Specifically, changes in the T&E Index must be created or new factors, like the Teacher-Weighted Formula, added to support salary structures consistent with the 3-Tiered Licensure System. State funding must be provided so that local districts can honor experienced teachers and the credentials they hold. Ideally, changes would:

- Fund beyond the minimum salaries by attaching weight and compensation to experience and credentials for increases in salaries.
- Create a mechanism to get the additional funding to the employees and school districts.
- Allow employees and school districts to design changes that honor education and experience and incorporate the minimum salary levels required by law.

Changes in the School Finance Act must ensure that funding for minimum pay levels flows to school districts to provide differentiation between Level II and III. Each Level II and Level III teacher is weighted by a formula based on:

- Experience
- Education
- Credentials
- Added responsibilities
- Evaluations that include student growth
- School growth

The system we recommend must honor the induction years in the profession by ensuring that Level I teachers are asked to focus on three things:

- Learning to teach well.
- Participating in mentoring.
- Learning how to create and measure Student Learning Objectives.

After a teacher successfully completes his/her years in Level I, this system would differentiate between new Level II and III teachers and experienced Level II and III teachers, continuing to support their development on a career continuum. In a comprehensive system, teacher development and evaluation must foster *collective* responsibility and accountability. Examples of how factors can be associated with education and experience to augment the minimum pay levels and create a Teacher-Weighted Formula are shown within each of the following charts.

Level I Example Base pay \$30,000				
Years within Level	BA+30	MA	MA+30	Credential Differential (National Board Certification, ESL, Bilingual, etc.)
0-3	.05 / \$31,500	.15 / \$34,500	.20 / \$36,000	.10 in addition to current base
4-5	.10 / \$33,000	.20 / \$36,000	.25 / \$37,500	.10 in addition to current base

Level II Example Base Pay \$40,000							
Years within Level	Education			Credential Differentials	Added Responsibilities	Evaluation/ Student Learning Objectives (SLO)	School EPSS Goals
	BA	BA+30	BA+45 or MA				
1-5	\$40,000	.15 \$46,000	.20 \$48,000	.10 \$4,000	.10 \$4,000	.10 \$4,000	.10 \$4,000
6-10	.10 \$44,000	.20 \$48,000	.25 \$50,000	.10 \$4,000	.10 \$4,000	.10 \$4,000	.10 \$4,000
11+	.15 \$46,000	.25 \$50,000	.30 \$52,000	.10 \$4,000	.10 \$4,000	.10 \$4,000	.10 \$4,000

Level III Example Base Pay \$50,000							
Years within Level	Education			Credential Differentials	Added Responsibilities (Mentoring)	Evaluation/ Student Learning Objectives (SLO)	School EPSS Goals
	MA or NBC	MA+30	Doctorate				
1-5	\$50,000	.15 \$57,500	.20 \$60,000	.10 \$5,000	.10 \$5,000	.10 \$5,000	.10 \$5,000
6-10	.10 \$55,000	.20 \$60,000	.25 \$62,500	.10 \$5,000	.10 \$5,000	.10 \$5,000	.10 \$5,000
11+	.15 \$57,500	.25 \$62,500	.30 \$65,000	.10 \$5,000	.10 \$5,000	.10 \$5,000	.10 \$5,000

The Investment Needed

Attracting, retaining and supporting excellent teachers within a reformed evaluation/compensation system cannot be accomplished without an investment in our workforce.

Based on teachers only, and not yet including all the support and related personnel that are essential to our students' success, we estimate that creating a comprehensive evaluation and compensation system will require an investment of \$350 million dollars.

	APS	State-Wide
Current teacher payroll approximately	278,000,000	1,029,629,630
Projected Cost	319,724,000 additional cost=41,724,000	1,184,162,963 additional cost=154,533,333
Additional Responsibilities	TBD	TBD
Mentoring	2,000,000	7,407,407
Credentials <i>(such as Bilingual/ESL in APS. National Board differential not included in cost since already in formula.)</i>	1,512,500	5,601,852
SLO/EPSS maximum cost	47,900,000	177,407,407
EST TOTAL	\$93,136,500	\$344,950,000

Note: Figures are based on APS accounting for 27% of state public education budget.

Conclusion

A high quality comprehensive teacher evaluation/compensation system will improve schooling and will benefit teachers and their students. The Albuquerque Teachers Federation is recommending ways to accurately measure student learning in ALL content areas, so that it can be appropriately factored into teacher evaluation and compensation systems. The multiple measures we are recommending will enhance professional accountability in a way that increases benefits for all students. In the ATF system, we measure what is valued: for teachers, effective practices leading to significant student outcomes; for students, the ability to understand essential concepts, work in groups, think critically and solve problems.

In order to overhaul these systems, we need:

- State lawmakers to have the political will to raise revenue to fund public education appropriately.
- A truly collaborative relationship among all stakeholders in our public education system.
- To ensure that teachers work within a highly functional *system* — one that, at a minimum, provides meaningful performance evaluations, high-quality professional development, reasonable class sizes, reliable and stable leadership, and time for planning and collaboration.

We can have a perfect evaluation system, but if nothing else in public education changes, then revamping the teacher evaluation system will change nothing.

Our recommendations, based on current research and best practices avoid the pitfalls embedded in systems that have leaped to quick fixes and have serious unintended consequences. It's critical we get this right. With multiple measures we can link teacher evaluation, compensation and student outcomes together in a fair, reliable system that will change public education for the better.

Comparison: Governor’s Task Force Recommendations vs. ATF Plan

Component	Task Force	ATF
Teacher “Effectiveness”	5 levels, based on student achievement (test scores), observations and other measures TBD. Each teacher’s value-added contribution would be calculated by PED	4 Levels based on multiple measures: <ul style="list-style-type: none"> • Student Learning Objectives • Collaborative work • Observations by the Principal • Observations by the Evaluation Team • Student Surveys
Student Learning	50% of an evaluation, measured by test scores – SBA scores for those grades/subjects giving SBA; school letter grade for others until assessments are identified, developed and approved for non-tested grades/subjects (within 2 years)	30% on Student Learning Objectives (SLOs) - annual learning objectives that can be developed by any teacher for any subject in any grade, individualized to meet students’ needs
Compensation	Beginning 2013-14, merit pay based on effectiveness ratings, consisting of value-added contributions calculated by PED, observations and other measures	Teacher-weighted formula, additional compensation for experience, education, responsibilities, student outcomes (evidenced by SLOs), and achieving school EPSS goals
Observations	25% observations – at least 4 per year, minimum of 2 by principal. Others may be done by outside PED-approved evaluators, or other colleagues (formative)	20% - Observations by the Evaluation Team -every year for formative evaluations. 20% - Observations by the Principal -every three years during summative evaluations.
Other Multiple Measures	25% locally selected, PED-approved measures - e.g. portfolios of teacher and/or student work; parent and/or student surveys; and other research-based measures, <i>INCLUDING basing up to 10% on school letter grade.</i>	10% - Collaborative work 20% - Student Surveys
Collective Accountability	NONE-focus on competition.	School-wide EPSS, with compensation for whole school meeting EPSS goals
Shortage Areas	Incentives to teach in critical shortage subject areas, geographic	Completing and funding the 3-Tiered compensation system

	areas and other hard-to-staff areas	address shortage areas. Competitive salaries have been proven to address shortages.																				
Due process	Given to teachers after obtaining Level 2 license and receiving effective evaluations	Provided to teachers after 2 successful years of teaching and upon signing their third contract— as per state law																				
3-Tiered System & PDD	<ul style="list-style-type: none"> • Evaluate in 2011-12 to minimize administrative costs & determine effectiveness • Incorporate “effectiveness” into licensure => tie test scores to licensure • Restructure pay to create smaller jumps, more frequent opportunities to advance (?) and free up resources for merit pay 	<ul style="list-style-type: none"> • Advancement through PDD process (*perhaps revise PDD to include demonstration of ability to develop SLOs and evaluate student progress) • Adequately fund and expand the 3-Tiered System to include a teacher-weighted formula, and incorporate SLOs and school-wide goals 																				
Funding	Funding for merit pay for a few	Adequately fund education— Funding for programs and salaries that start but not end with the 3-Tiered minimums																				
Measures	<table border="1"> <caption>Measures Evaluation Components</caption> <thead> <tr> <th>Component</th> <th>Percentage</th> </tr> </thead> <tbody> <tr> <td>Student Achievement</td> <td>50%</td> </tr> <tr> <td>Observations</td> <td>25%</td> </tr> <tr> <td>Other Measures</td> <td>25%</td> </tr> </tbody> </table>	Component	Percentage	Student Achievement	50%	Observations	25%	Other Measures	25%	<table border="1"> <caption>Evaluation Components</caption> <thead> <tr> <th>Component</th> <th>Percentage</th> </tr> </thead> <tbody> <tr> <td>Individual SLOs</td> <td>30%</td> </tr> <tr> <td>Collaborative work: EPSS goals/school responsibilities</td> <td>10%</td> </tr> <tr> <td>Principal Observations</td> <td>20%</td> </tr> <tr> <td>Evaluation Team Observations</td> <td>20%</td> </tr> <tr> <td>Student Surveys</td> <td>20%</td> </tr> </tbody> </table>	Component	Percentage	Individual SLOs	30%	Collaborative work: EPSS goals/school responsibilities	10%	Principal Observations	20%	Evaluation Team Observations	20%	Student Surveys	20%
Component	Percentage																					
Student Achievement	50%																					
Observations	25%																					
Other Measures	25%																					
Component	Percentage																					
Individual SLOs	30%																					
Collaborative work: EPSS goals/school responsibilities	10%																					
Principal Observations	20%																					
Evaluation Team Observations	20%																					
Student Surveys	20%																					

Appendixes

1. APS-SIG Principal Observation Tool
2. APS-SIG Student Learning Goals Rubric
3. APS-SIG Sample Student Survey Questions

References

- American Federation of Teachers. (2011). *A Guide for Developing Multiple Measures for Teacher Development and Evaluation*. Washington, DC: Author.
- American Federation of Teachers. (2010). *Why Teachers Must Have an Effective Evaluation System*. Washington, DC: Susan Moore Johnson, Ed.D.
- American Federation of Teachers. (2011). *Evaluating Teaching with Multiple Measures*. Washington, DC: Laura Goe, Ph.D.
- Danielson, Charlotte (1996). *Enhancing Professional Practice: A Framework of Teaching*. ASCD.
- Futernick, Ken. (2010). *Incompetent Teachers or Dysfunctional Systems?* Phi Delta Kappan, October, pp. 59-64.
- Kerchner, Charles (2011). *New Analysis Shows Serious Flaws in LA Times Evaluation of Teacher Effectiveness*, Claremont Graduate University
<http://www.huffingtonpost.com/charles-kerchner/new-analysis-shows-seriou b 820755.html>
- National Comprehensive Center for Teacher Quality (2009). *Methods of Evaluating Teacher Effectiveness*. Washington DC, Laura Goe and Andrew Croft.
- New Mexico Effective Teaching Task Force *Final Report and Recommendations*, August 26, 2011.
- Popham, W. James. (2007). *Instructional Insensitivity of Tests: Accountability's Dire Drawback*. Phi Delta Kappan, pp146-155.
- The Economic Policy Institute (2010) *Problems with the Use of Student Test Scores to Evaluate Teachers*. Eva L. Baker, Paul E. Barton, Linda Darling-Hammond, Edward Haertel, Helen F. Ladd, Robert L. Linn, Diane Ravitch, Richard Rothstein, Richard J. Shavelson, and Lorrie A. Shepard. EPI Briefing Paper #278.

