Pursuing Excellence
Reforming Teacher and Principal Evaluation in Washington
Curriculum in Context

Pursuing Excellence – Reforming Teacher and Principal Evaluation in Washington

A message from the editor .................................................... 3
    David W. Denton

A message from the president ............................................. 4
    Tim Nootenboom

Using the Framework for Teaching as the Foundation of a District’s System of Teacher Evaluation .................................. 5
    Charlotte Danielson

The Marzano Model for Teacher Evaluation in Washington State ............................................................... 8
    Phil Warrick and David Livingston

Leading for Teacher Effectiveness ........................................ 11
    Stephen Fink, Annette Markholt and Sandy Austin

Effective Principal Evaluation Systems: Embracing Standards Avoiding Standardization .............................................. 14
    Thomas Alsbury

Toward a More Meaningful Evaluation System:
Scaling Implementation Statewide ......................................... 18
    Chris Korrimo and Heather Cope

The Art of Teacher and Principal Evaluations ............................ 20
    Paul Sturm

Learning from the Journey of the NEWESD 101 TPEP Consortium Pilot ................................................................. 23
    Helene Paroff, Pam Veltri, and Shauna Schmerer

The TPEP Process: A Teacher’s Perspective ............................... 25
    Amy Bumpus

Microwaves and other Teaching Tools ..................................... 27
    Michaela Miller

Executive directions:
A message from the executive director ................................... 29
    Kathy Clayton

Publisher

Washington State Association for Supervision and Curriculum Development
825 Fifth Avenue SE, Olympia, WA 98501

WSASCD publications do not necessarily reflect ASCD views and are not official publications of ASCD.

Curriculum in Context is published twice a year. Manuscripts should be addressed to David Denton, Seattle Pacific University, e-mail: dentod@spu.edu.
The editorial committee seeks articles that provide perspectives, research and practical information about the issues of and ways to improve learning and teaching in Washington State.
ISSN 2165-7882

Cover photo credit: Info
A message from the editor

By David W. Denton, Ed.D.

The theme of this edition of Curriculum in Context is teacher and principal evaluation. In 2010, Governor Gregoire signed Senate Bill 6696 into law, initiating reform of Washington’s teacher and administrator evaluation framework for educators across the state.

This issue begins with three descriptions of the evaluation models districts are adopting. Charlotte Danielson (Danielson Group), Phil Warrick and David Livingston (Marzano Research Institute), and Stephen Fink, Anneke Markholt, and Sandy Austin (Center for Educational Leadership) summarize their frameworks for evaluation and describe steps for implementation. These summaries are informative and useful for understanding similarities and differences between the systems. They also identify resources required for effective implementation, which is an area of significant challenge.

The article by Thomas Alsbury (Seattle Pacific University) transitions the issue with an analysis of research on effective practices in principal evaluation. An interesting conclusion from Dr. Alsbury is that effective leadership is situational and adaptive, given the variety of contextual factors that building administrators encounter. I would venture to speculate that effective leadership, whether at the building or classroom level, exhibits these qualities. This conclusion has implications for standardizing evaluation systems. Nevertheless, Chris Korso and Heather Cope (League of Education Voters) suggest that reform is necessary, and that Washington must lead the way toward ensuring that effective teachers and administrators are working in each school.

The latter half of this issue could aptly be titled Challenges and Opportunities. The authors provide insights about the resources and actions needed to effectively implement a new evaluation model. Readers of these articles will be reminded that even the most elegant and effective framework is rendered clunky and useless unless serious thought is given to implementation. These insights are told from the perspective of teachers, principals, and district level administrators. Some authors share similar understandings, such as the notion that more time will be required on the part of principals and teachers to engage in evaluation activities. These insights are sure to resonate with those who are currently reforming their evaluation systems.

If the articles in this issue are any indication, then the hard work of reforming teacher and principal evaluation has just begun and it offers tremendous potential and challenge. Nevertheless, one message that comes through loud and clear from all of the authors is that change is needed. Nevertheless, time will tell whether our efforts bring about substantive improvement not only in student achievement, but also in the social-moral fabric of the schoolhouse. Personally, I am cautiously optimistic. My optimism stems from the belief that the way forward has been identified, but I am cautious because reform efforts, of all kinds, have been layered onto a system which has not changed structurally for at least 100 years.

Let me close this note with thanks to Becky Cooke, Jim Howard, and Gene Sementi for their excellent work as editors of Curriculum in Context over the last two years. They set a high standard, one which the editorial staff at CiC will strive to match by assembling expert voices to discuss current topics, relevant to Washington educators.

David W. Denton, Ed.D., is an Assistant Professor at Seattle Pacific University. Before joining Seattle Pacific, David taught middle school students, in a variety of disciplines, for ten years. In 2005, David earned National Board Certification in early adolescent mathematics.
Greetings fellow educators across Washington State! With just a few weeks left in the current school year, we find ourselves in familiar territory - facing the pressures of getting everything done that we set out to do, striving to provide our very best for students so that they finish the year with as much success as possible. At the same time, we are making preparations for September to begin the cycle again.

How time flies, it seems like the current school year just began. Perhaps this feeling of decreased distance between beginning and ending is the result of our mindset, which tends to focus on reaching goals, creating success, and celebrating accomplishments. By the end of the year, we are looking back and reflecting on those activities, and we remain hopeful about the results for our students, our schools, and for ourselves.

What have you been contemplating and reflecting on as this school year draws to a close? What have you learned about instruction and assessment and how can you apply your new knowledge and experience across situations? How do you engage in self-assessment with regard to these ideas? As I consider possible answers to these questions, I think about the idea of “silos of success.” It seems that we experience victory in particular places and situations, or silos, and as a result, we do not always share common language or concepts for replicating outcomes.

Nevertheless, we overcome these obstacles through thoughtful reflection, professional action, and kindness. Another way we overcome is through systematic change, such as the kind directed through Senate Bill 5895 and 6696. These new regulations provide a foundation for increased opportunities to engage in effective teacher and principal assessment. I believe that one result of these reform efforts is to make assessment more constructive, collaborative, and meaningful. While some educators may be apprehensive about these changes, I believe they provide common language and concepts for penetrating the assessment silo, and generalizing results across classrooms and schools. If we, regardless of our roles, strive to engage in this reform and adopt these new ideas, then I believe this initiative can serve as an effective means in bettering our craft and bringing more focus to our work.

Although this legislative action can be daunting, we know that remaining proactive presents the most viable way forward. Striving for a collaborative and transparent process in this endeavor will help minimize some of the apprehension that we may feel. Ultimately, reforming teacher and principal evaluation gives us a common language and set of concepts for improving instruction, and increasing student achievement.

It is with these ideas in mind, and the call to full engagement, thoughtful reflection, and rigorous kindness, that I wish you a successful conclusion to the 2011-2012 school year. And, on behalf of WSASCD, I wish you a relaxing summer.

Tim Nootenboom is President of WSASCD, and Executive Director for Learning and Teaching, Central Valley School District, Spokane Valley.
We are, in the US (and around the globe), at a unique moment in public education; several research studies, and the wave of legislation that has occurred simultaneously, has focused the public and the nation’s educators on the quality of teaching as never before. It is now widely recognized that while teacher quality is not the only contributor to student learning, it is vitally important and is the most significant factor under the direct control of the school. Therefore, it is essential to use all the available policy levers, and the findings from research on both teacher quality and professional learning, to enhance the quality of classroom practice.

But there’s some distance to go; the Widget Effect report (Weisberg et al., 2009), released by the New Teacher Project in 2009 concluded that “A teacher’s effectiveness – the most significant factor for schools in improving student achievement – is not measured, recorded, or used to inform decision-making in any meaningful way” (p. 31). The report found that the performance of the vast majority of teachers, when they were evaluated at all, was judged to be at the highest level of whatever rating scale was used. And while this would be excellent news (our schools are filled with highly-effective teachers!) if students themselves were achieving at high levels, the fact that they’re not has forced educators to conclude that the procedures for identifying their most effective teachers, and for helping all teachers improve their practice, are inadequate.

This was a highly influential report, and contributed to efforts in many states, including Washington, to develop better procedures for evaluating teacher practice. At the same time, and extending for several years on either side, there have been a number of rigorous research studies exploring the relationships between teacher practice and student learning (Consortium on Chicago School Research, 2012; Measure of Effective Teaching, 2012; Wisconsin Center for Education Research, 2012).

As educators have engaged more deeply with the area of teacher evaluation, several things have become clear:

• Accurate assessment of teacher knowledge and skills is essential in order to make defensible personnel decisions and to help all teachers improve their practice.

• To argue for teacher evaluation being used to help teacher practice improve is not to suggest that teaching is of poor quality and has to be “fixed.” Rather, it’s in recognition of the fact that teaching is so hard that it can always be improved. No matter how good a lesson is, in other words, it could always be at least a little bit better.

• Therefore, an evaluation system should yield information that is diagnostic and useful to teachers in improving their practice in those aspects of teaching that are most important, and make the biggest contribution to important student learning.

As performance appraisal has become higher-stakes for individual teachers, it has also become high-stakes for the school districts (or states) as they implement new procedures. That is, if consequential decisions will be made based on assessment procedures, those procedures themselves must be rigorous, valid, and reliable. Furthermore, if one of the aims of an evaluation system is for teachers to be able to improve their practice, then the systems must incorporate what we know about professional learning.

Fortunately, recent research has pointed the way towards both of those goals. And, since my Framework for Teaching has been integral to several of these studies, I’ve been in the fortunate position to learn from these studies and to incorporate their findings into my approach to teacher evaluation.

So what have we learned from the research?

Quality Assurance: Ensuring that all Teachers are Meeting Rigorous Standards of Practice

In order to implement a defensible system of teacher evaluation, several components must be in place:
A validated instrument, one that demonstrates high correlations between teacher performance and student learning. I was delighted to learn that the Framework for Teaching, when subjected to rigorous, independent, scrutiny, proved to be a valid instrument. As part of this effort to identify a valid instrument, educators should also take care to use one that reflects the true complexity of teaching. It’s not difficult to devise a simple checklist of teacher “behaviors” which, while easy to train observers to recognize, don’t reflect what’s important about good teaching. It also suggests that the use of an instrument that reflects the complexity of teaching requires more time for training than a more simplistic one; this suggests that while there is certainly some urgency in implementing good systems of teacher evaluation, the project cannot be rushed so that educators do not have the opportunity to deeply understand what is required of them.

The Gates-funded research effort known as the MET (Measures of Effective Teaching, 2012) project is the largest-scale study to this point that has looked at this issue. It involved video recording approximately 23,000 lessons. Then, these lessons were evaluated using five observation protocols (of which my framework for teaching was one). Furthermore, the teachers shown in these lessons were evaluated using associated instruments and then scores were correlated with value-added student outcome data.

Another study was conducted by a research group affiliated with the University of Chicago, analyzing the correlations of teacher performance on the framework for teaching against measures of student learning, over a two-year period (Consortium on Chicago School Research, 2011).

Both of these large, independent, studies came to similar conclusions, that the framework for teaching is a valid instrument for evaluating teacher performance. It is impossible to overstate the significance of these findings; the first requirement for any district seeking to implement a defensible teacher evaluation system must start with a valid instrument.

Instruments and procedures through which teachers can demonstrate their skill. This includes classroom observation, of course, but for those “behind the scenes” aspects of teaching, other approaches (for example, teacher artifacts) are essential. The most thorough research to date, however, has concentrated on the observation of classroom practice.

Trained evaluators who can make accurate and consistent judgments based on evidence. At the Danielson Group, we have had years of experience in training evaluators. It is, to be sure, enormously challenging. In fact, as a consequence of participation in a precursor study to the MET project, we were obliged to make some revisions to the framework for teaching (tightening the language, creating “critical attributes” for each level of performance for each component) so observers could, after training, evaluate teacher performance accurately.

It is not sufficient for those who will be evaluating teachers (and making high-stakes personnel decisions) to simply attend a training. They should demonstrate that they have the skill to make accurate and consistent judgments based on evidence. They should, in other words, be certified to do this important work. Fortunately, this is an area where the Danielson Group has accumulated a good deal of experience.

In order to participate in the MET project, we were required to develop the capability of training hundreds of observers; because of the huge numbers involved, this training (and the proficiency test) had to be conducted entirely online. This approach to training was highly successful. Raters passed the certification test at better than 90%. For district implementations, the online training, practice, and proficiency test is now available, and it serves as a powerful supplement to some types of face-to-face training.

The training for the hundreds of observers in the MET study was conducted entirely online, with raters working independently, and at their own speed. This training was highly effective. For example, raters passed the certification test on the Framework for Teaching at over 90%. I believe that online training, in a school setting, is greatly enhanced, however, by group sessions in which participants discuss the video examples provided in the online resources, to learn from the insights of others, and to ensure that their judgments are accurate and consistent with nearby colleagues.

**Promoting Professional Learning**

It is not sufficient to evaluate teacher performance, unless those evaluations lead to enhanced practice. It is even better if the procedures used to evaluate teachers can, themselves, contribute to professional learning. The good news here, from several studies, is that this is possible.

The Chicago study, referenced earlier, included interview and survey data collected from hundreds of teachers and evaluators. This data, along with correlational analyses between evaluation performance scores and student achievement scores, suggest the following:

- Virtually everyone found the process to be worthwhile. Those who had reservations about it were concerned primarily with the time they were required to do the process well. This has enormous implications, of course, for practitioners in determining the appropriate work load, and perhaps in identifying individuals other than site administrators who can conduct some of the observations.

- The principal mechanism for teacher learning, as reported by both teachers and administrators, was the conversations that followed each observation. These were professionally rich, and teachers, in particular, said that they became more thoughtful and analytic about their practice as a result of these conversations.

- Principals reported that they wanted to improve their skills in conducting these professional conversations; they recognized their value, and wanted to further develop their capacity to do them well. Conducting professional conversations is clearly an important skill, one that combines aspects of providing feedback regarding an observed lesson, coaching skills to help teachers analyze their own practice, and knowing when different types of conversations should be used.

An additional study, published by Eric Taylor and John Tyler (2011), examined the effects of a rigorous evaluation system on teacher performance, during the year of the evaluation and in subsequent years. What
they learned has important implications for ongoing efforts in teacher evaluation. Specifically, these researchers found the following:

- The very act of going through a year-long evaluation process in Cincinnati strengthened teacher performance. While the research and statistical details are still being analyzed, the results suggest that the correlations are positive, and the effect sizes are large enough to be quite consequential.
- Not only does a teacher’s effectiveness increase in the year in which they are undergoing evaluation, but the effects of going through the evaluation cycle are even larger in the years after the evaluation.

Implementation Challenges

Any system of teacher evaluation is, ultimately, only as successful as the details of its implementation; it has to work, in other words, “on the ground.” So what needs to be done to ensure a successful implementation?

- **System Design.** The most important decision made during the design process is the instructional framework that will serve as the foundation of the system. Naturally, I hope that people will select my framework for teaching, but I know that there are other options. But in addition, as part of system design, it’s essential to think through questions such as:
  - Will the system be the same for probationary teachers as for teachers on a continuing contract?
  - How many classroom observations are there? How long will they be? Will they be announced or unannounced?
  - Will the system incorporate other evidence in addition to classroom observation? What about planning documents? Record of teachers’ communication with families? Records of teachers’ participation in the professional community, or engagement with professional learning opportunities?
  - What instruments, or forms, will be used?

- **Training of Evaluators.** It is difficult to overstate the importance of this element of implementation. Its importance has been discussed previously. An evaluation system cannot be rigorous or legally defensible if the judgments made by evaluators are not accurate and consistent. But the challenge for implementation is a little different. Namely, how to ensure that sufficient resources are devoted to training for the results to be reliable. When the Danielson Group conducts face-to-face training, we spend a full day on helping people understand the framework for teaching, followed by two days of training in observation skills. Some of this can be enhanced by the online tools we have helped to develop, but there must be a commitment to offer training (and certification) for anyone who will be making high-stakes personnel decisions.

- **Professional Development for Teachers.** The first rule of assessment (of any kind) is that people are not evaluated on things they don’t know about. Even a test for a driver’s license, for example, has clear guidelines as to what people should know and be able to do in order to pass. The situation is the same for teachers, only more complex, since teaching is such complex work. Professional development for teachers can take many forms; the lowest-tech, and cheapest, option, is simply a book study. Many schools purchase copies of the framework for teaching for all their teachers, and devote a meeting each month to discussing the book, one chapter at a time. As part of the book study, it’s enlightening for teachers to consider how the components of the framework for teaching apply to their own situations, what high quality questioning and discussion skills, for example, sound like in a 3rd grade social studies lesson, or in an AP Chemistry class.

Beyond a book study, many of the online resources developed to train evaluators are highly relevant to teachers as well. Video clips illustrating the different components of the framework for teaching, particularly those that demonstrate each level of performance, sharpen everyone’s lens in observing practice and enhancing teacher knowledge and skills.

**Summary**

Washington State is embarking on an exciting period of educational improvement by supporting districts in implementing new evaluation systems. Clearly, this work has implications for assisting teachers in new ways, such as encouraging new kinds of teacher reflection, delivering focused and accurate feedback about practice, and communicating guidelines for improving instruction. Whether or not this vision is realized depends on implementation. Serious attention must be given to multiple factors, such as designing an overall system that coordinates evaluation efforts, selecting a valid instrument, rigorous training for evaluators, and targeted professional development for teachers.

**References**


Charlotte Danielson is a former economist and educational consultant who has taught at all levels, kindergarten through college, and has worked as an administrator, a curriculum director, and a staff developer. In her consulting work, Ms. Danielson specializes in aspects of teacher quality and evaluation, curriculum planning, performance assessment, and professional development. More recently, she has advised state and national policy bodies on assessing teacher effectiveness.

Ms Danielson is the founder of the Danielson Group, whose 30+ consultants assist school districts and states in designing and implementing systems of teacher evaluation.
Effective teaching has the potential for playing an incredibly powerful role in the life of a learner. It is an art and science that takes place in the dynamic environments of our nation’s classrooms. This complex system of human interaction deserves a research-based evaluation model that clearly focuses on student achievement as the end result. Cultivating effective teaching practices occurs when instructional planning and instructional supervision share a common language. The Marzano Teacher Evaluation Framework is built on a common language for instruction. Through the common language teachers receive focused feedback that recognizes their current strengths and encourages their continued professional growth in four domains of professional practice.

This article presents the four domains of the Marzano Teacher Evaluation Framework, makes direct links to the Washington Teacher Evaluation Criteria, provides feedback from practitioners who have used the model, and introduces the concept of the Teacher Reflective Practice Scales which assist teachers in planning instruction in the same format with which they will be evaluated. Additionally, as this article is being written, Marzano Research Laboratory is working directly with State of Washington representatives and the Wenatchee School District to finalize a Washington version of the Marzano model. The Wenatchee school district has been working with the Marzano model for nearly 18 months and has been instrumental in the alignment to Washington’s evaluation criteria.

The Marzano Model for Teacher Evaluation in Washington State

“The quality of an education system cannot exceed the quality of its teachers.”

Barber and Moursched, quoted in Leaders of Learning, DuFour and Marzano, (2011), p. 4

Effective teaching has the potential for playing an incredibly powerful role in the life of a learner. It is an art and science that takes place in the dynamic environments of our nation's classrooms. This complex system of human interaction deserves a research-based evaluation model that clearly focuses on student achievement as the end result. Cultivating effective teaching practices occurs when instructional planning and instructional supervision share a common language. The Marzano Teacher Evaluation Framework is built on a common language for instruction. Through the common language teachers receive focused feedback that recognizes their current strengths and encourages their continued professional growth in four domains of professional practice.

This article presents the four domains of the Marzano Teacher Evaluation Framework, makes direct links to the Washington Teacher Evaluation Criteria, provides feedback from practitioners who have used the model, and introduces the concept of the Teacher Reflective Practice Scales which assist teachers in planning instruction in the same format with which they will be evaluated. Additionally, as this article is being written, Marzano Research Laboratory is working directly with State of Washington representatives and the Wenatchee School District to finalize a Washington version of the Marzano model. The Wenatchee school district has been working with the Marzano model for nearly 18 months and has been instrumental in the alignment to Washington’s evaluation criteria.

The Marzano Framework

At the center of Robert Marzano’s contributions to our profession have been his analysis of the research on effective instruction and the translation of those findings into concrete guidance for practitioners. The publication of The Art and Science of Teaching identified 41 classroom strategies and behaviors that are associated with gains in student achievement. These 41 elements are organized under nine lesson design questions in the Marzano Teacher Evaluation Framework and comprise Domain 1.

In addition, the remaining 19 elements of the model are contextualized in three interrelated domains: Planning and Preparing, Reflecting on Teaching, and Collegiality and Professionalism.

With four domains and 60 elements, the Marzano model mirrors the complex realities of teaching while revealing several of our fundamental biases. First, more than half of the model’s elements (41 of 60) are nested in Domain 1, evidence of our recognition of the importance of the direct, classroom-based actions of teachers. The causal link between strategies employed in a classroom and student learning is beyond dispute. Second, for a framework of teaching to be useful it must be sufficiently fine-grained to identify specific things that a teacher can focus on and refine. Which implies our third bias: it’s about growth… growth in the skills and behaviors of our teachers that pay dividends for our students.
**Domain 1: Classroom Strategies and Behaviors**

This is the domain that directly addresses what teachers do in classrooms. In the Marzano Framework the 9 design questions and 41 elements, drawn from *The Art and Science of Teaching*, are organized under three general types of lesson segments. A lesson segment is an event in the classroom with strategies employed that are designed to meet a specific learning purpose. We use the concept of lesson segments to serve as a unit of analysis and provide focused feedback to teachers. In the following we address each of the three lesson segments in turn, although it is important to note they are not linear.

“Routine segments” are strategies that include communicating learning goals, tracking student progress, celebrating the learning successes of individual students and whole classes, and the establishment and maintenance of productive routines including the physical layout of the classroom.

“Content specific segments” are precisely that; things that teachers do to help students successfully interact with content. The Marzano model contains three specific situations for teaching content: 1. Introducing new content, 2. Having students practice and deepen their knowledge, or 3. Having students generate and test hypotheses using what they have learned. Depending on which of these three specific situations is in play, teachers use different strategies that are clearly reflected in the model. These specific content situations represent *The Art and Science of Teaching’s* Design Questions 2, 3, and 4.

“Lesson segments enacted on the spot” is the final of the three segment categories. These are strategies that might not be a part of every lesson but when called for a teacher must attend to them or the learning environment is at risk. They are elements that teachers need to be prepared to use at a moment’s notice. When we look at the specifics of these elements they reflect research about student engagement; what we do to acknowledge when students are adhering, or not adhering, to established classroom rules and procedures; establishing and maintain effective relationships with students; and communicating high expectations for all students…especially those for whom school success is an elusive target.

**Domain 2: Planning and Preparing**

Central to an understanding of the Marzano Framework is the influential relationship that a teacher’s planning has on the classroom strategies employed when it’s time to teach. Impactful teaching has never been a matter of winging it. When we talk about planning we mean the planning it takes to ensure the effective scaffolding of information within lessons, the logical and meaningful progression of lessons within units, and the attention a teacher must pay to the established content standards of districts and states. Furthermore, that planning must be understood to include the thoughtful and effective use of available materials and technologies that can enhance students’ understanding of the content in a given lesson or unit.

Because the learning needs of students are as individual as the students themselves, it is essential to identify the support that three specific sub-populations of students may require. First, as the population of English language learners grows in many of our schools, teachers are called on to think ahead about the adaptations that must be made for those learners. Second, there are few if any classrooms in America that don’t need to serve the needs of special education students. The critical support that special educators provide does not supplant the instructional adaptations that regular education teachers must be prepared to make. And third, when students come from home environments that offer limited support for schooling, teachers must plan supportive approaches for students with few material or psychological resources to draw on outside of school.

These three elements of Domain 2, planning and preparing for lessons and

<table>
<thead>
<tr>
<th>The State of Washington’s Teacher Evaluation Criteria per E2SSB 6696</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 – Centering instruction on high expectations for student achievement</td>
</tr>
<tr>
<td>#2 – Demonstrating research-based instructional practices</td>
</tr>
<tr>
<td>#3 – Addressing the needs of individual students</td>
</tr>
<tr>
<td>#4 – Focusing on subject matter content and curriculum</td>
</tr>
<tr>
<td>#5 – Providing a safe, inclusive learning environment</td>
</tr>
<tr>
<td>#6 – Using student data to modify instruction and improve student learning</td>
</tr>
<tr>
<td>#7 – Communicating and collaborating with families and communities</td>
</tr>
<tr>
<td>#8 – Utilizing collegial collaboration to improve practice and advance student learning</td>
</tr>
</tbody>
</table>
units, for the use of materials and technologies, and for the special needs of students speak directly to WA Criterion #3: Addressing the needs of individual students; and WA Criterion #4: Focusing on subject matter content and curriculum.

Domain 3: Reflecting on Teaching
A case could be made that the State of Washington’s intent to reform current teacher evaluation practices is significantly realized in the fulfillment of the two elements of Domain 3: evaluating personal performance and developing and implementing a professional growth plan. When fulfilled in their logical sequence, the central role that a teacher will play in being responsible for his or her own professional growth could not be more clear. A teacher who is able to evaluate his or her performance has taken a giant first step toward ever-increasing effectiveness.

Marzano’s “Teacher Scales for Reflective Practice: Domain I” (marzanoresearch.com offers a free download) invites a teacher to examine the 41 elements of Domain I of the framework and then use a scale to rate his or her present skillfulness in the utilization of each one. The format of the scales for Domain 1 helps avoid the simplistic trap of some self-assessment and feedback protocols that devolve into checklists of strategies being used or not used. The values from the Marzano Teacher Reflective scales are:

- Innovating (4) – Adapts and creates new strategies for unique student needs and situations
- Applying (3) – Engages students in the strategy and monitors the extent to which it produces the desired outcomes
- Developing (2) – Engages students in the strategy with no significant errors or omissions
- Beginning (1) – Uses the strategy incorrectly or with parts missing
- Not Using (0) – Strategy was called for but not exhibited

The scales for Domains 2, 3, and 4 are in a format that parallel those of Domain I and can be found at marzanoresearch.com > Services > Teacher Evaluation Model – Resources, on pp. 40-48 of the “Marzano Teacher Evaluation Scales.”

Ultimately a teacher’s self-evaluation must translate into systemic action, and a vehicle for such action is a plan for professional growth and development. For a meaningful and manageable annual professional growth plan, we advocate that a teacher select two or three elements from Domain 1 and another several from Domains 2-4, as targets for focus and improvement. The selection process will, of course, draw heavily on the teacher’s own self-reflection, but done in consultation with that teacher’s supervisor.

Expectations for the creation of these kinds of plans are typical in current practice. But in our framework a teacher monitors the extent to which goals are being accomplished and charts his or her progress on both the identified instructional strategies and the accompanying student learning gains. It is in this consistent and frequent monitoring that both the purpose and profit of growth planning comes to life.

The elements of Domain 3 align with parts of three of Washington’s new criteria: WA Criterion #2: Demonstrating research-based instructional practices; WA Criterion #6: Using student data to modify instruction and improve student learning; and WA Criterion #8: Utilizing collegial collaboration to improve practice and advance student learning.

Domain 4: Collegiality and Professionalism
This domain is not directly related to growth in classroom strategies and behaviors, but it is the context in which the other domains function. In a school or district, high levels of collegiality and professionalism can significantly enhance Domains 1, 2, and 3.

Domain 4 includes three elements: promoting a positive environment, promoting the free and open exchange of ideas and strategies, and promoting district and school development. In the Marzano Framework, it is the elements and actions of Domain 4 that specifically address WA Criterion #7: Communicating and collaborating with parents and communities, and WA Criterion #8: Utilizing collegial collaboration to improve practice and advance student learning.

Changing the Conversation: Collaboration for Teacher Effectiveness
For years, instructional evaluators have been searching for better tools to help them assess, provide feedback, and foster teacher pedagogy. The Marzano model provides these tools within the framework of a research-based common language for instruction. By clearly defining 60 effective teaching behaviors across four domains of professional practice, the Marzano model offers a user-friendly, common language for all educators. Within the Marzano model, 41 of the 60 elements are specifically focused on classroom practices and come directly from the body of research presented in The Art and Science of Teaching. Additionally, numerous free resources to support the use of the Marzano model are available on the Marzano Research Laboratory website. Marzano Research Laboratory has also developed two new publications to support teacher growth. These are Becoming a Reflective Teacher and Coaching Teaching.

For more information about the Marzano Teacher Evaluation model in the state of Washington, please email either of the authors of this article: phil.warrick@marzanoresearch.com or david.livingston@marzanoresearch.com

References


Phil Warrick, Ed. D. joined Marzano Research Laboratory in July of 2011 as a national trainer and Associate Vice President. Dr. Warrick spent 24 years working in public education including 14 as an administrator. In 2005 he was selected as the Nebraska High School Principal of the Year. In 2008 he accepted the position of Campus Principal at Round Rock High School in Round Rock, TX where he led the implementation of The Art and Science of Teaching and the Marzano Observation Protocol for teacher evaluation.

David Livingston, Ph.D., joined Marzano Research Laboratory in the fall of 2008. His public school career spanned 37 years and included ten years of classroom teaching, 20 years serving as principal of four schools, and from 1998 to 2005 he was Executive Director of Elementary Education for the Cherry Creek Schools, CO. David is coauthor with Robert Marzano and Tony Frontier of Effective Supervision: Supporting the Art and Science of Teaching, published by ASCD in 2011.
Recently we celebrated the 10-year anniversary of the University of Washington Center for Educational Leadership (CEL). Our mission a decade ago, as it is today, was to support school and teacher leaders in the challenging work of eliminating the achievement gap among students. When we began this ambitious journey, high school seniors in the current 2012 graduating class were in second grade.

Today, our nation’s education system continues to face stagnating achievement gains, despite nearly 10 years of federal education reform efforts. The once-bright promise of those reform efforts has faded as we recognize that too many of the students who were in second grade in the spring of 2002 will not be graduating with their class of 2012 this year.

If we, as a nation, are to raise achievement for all students, we must be prepared to invest where it matters most: in leading for teacher effectiveness. The research is clear. Student achievement will not increase until the quality of teaching improves, and the quality of teaching will not improve until our school leaders can observe, analyze and support the continuous growth of teachers.

Quality Teaching Matters

Over the course of 10 years, CEL has worked with thousands of teachers, principals and school district leaders across America. Our experience and the research show that if we want students to learn, we need to provide them with powerful learning opportunities. More so than family income or education levels – two reasons widely cited by educators why students are not learning – it is quality teaching that matters most when it comes to student achievement (Haycock, 1998; Peske & Haycock, 2006).

A close examination of student achievement reveals that differences among students, and differences among schools, pale in comparison to the differences in the quality of teaching from classroom to classroom when it comes to variances in achievement (Rowan, Correnti, & Miller, 2002). Rather than searching for the best schools for their children, parents should be searching for the best teachers.

Teaching Is More Complex Than People Realize

Producing the best teachers is no simple feat. While high-quality teaching matters most for student achievement, high-quality teaching is much more complex than the general public and policymakers realize.

What is high-quality teaching? In practice, there is little consensus. Our front-line educators simply do not agree on what constitutes powerful instruction.

In a frequent experiment with groups of school leaders, we show a video of an actual classroom lesson and at the end we ask them to rate the quality of instruction on a scale of one to five. We have run this experiment dozens of times with hundreds of educators and, invariably, the ratings run the scale from low quality to high quality. The reason for the discrepancies is that most school leaders and educators do not share a vision and common language for quality instruction.

The Complexity of Teaching Can Be Defined

Years ago, CEL set out to define a common language for teaching, drawing on the abundant research in the learning sciences. In a multiyear effort, CEL experts created a framework for quality teaching through a process of combining a number of empirical studies, drawing from experiential research, and corroborating our findings with a panel of practitioners with expertise in observing classrooms and providing feedback to teachers.

The result was our 5 Dimensions of Teaching and Learning instructional framework which has been used by school districts throughout Washington state and across the nation to develop a common vision of high-quality instruction. In brief, the 5D™ instructional framework defines teaching and learning along five dimensions: purpose, student engagement, curriculum and pedagogy, assessment for student learning, and classroom environment and culture. These five dimensions are further defined by 13 subdimensions (Figure 1).

While teaching remains complex, the 5D framework provides a clear instructional roadmap for educators. The resulting common language of instruction provides a foundation for powerful discourse about effective teaching, instructional feedback, and the collection and use of formative and summative data. The 5D™ instructional framework is helping to redefine the practice of teaching and learning and is redefining the meaning of school leadership by providing a framework for powerful discourse about effective teaching and learning.
The odds of improving instruction are measurably higher when both teachers and leaders use observation tactics linked to a clear purpose, intended outcomes and a theory of action. But, shifting to a culture of open and frequent classroom observation requires the right support. Structures, processes and protocols are critical to helping instructional leaders and teachers learn collaboratively how to use a framework to observe and analyze instruction.

At CEL, we utilize guided classroom walkthroughs as one means to ground instructional leaders and practitioners in the foundation of the 5D framework. By capturing what an instructional leader notices and thinks about in relation to instructional practice, a process of collaborative analysis and discourse can occur in service of creating a common vision and common practices in the classroom.

The CEL 5D+ Teacher Evaluation Rubric

With the passage of teacher evaluation legislation across the country, it is clearer than ever that an instructional framework – coupled with protocols around observation and feedback – is central to improving teaching practice and learning for all students.

To meet the specific needs of teacher evaluation legislation, CEL developed the 5D+ Teacher Evaluation Rubric. The 5D+ rubric provides language around four tiers of instructional expertise ranging from unsatisfactory to distinguished. By providing specificity about improving teacher practice in stages and steps, the 5D+ rubric allows leaders and teachers to engage in conversations focused on continuously moving instruction to a higher level.

The 5D+ Teacher Evaluation Rubric extends the five dimensions of the 5D instructional framework with the addition of a dimension covering the professional collaboration and communication roles of teachers outside the classroom (Figure 2).

Developed from the same research base as the 5D framework, the 5D+ Teacher Evaluation Rubric is a sophisticated tool that enables consistent, reliable observations of instruction across classrooms, schools and districts.

But it bears repeating that as a tool, the use of any evaluation rubric will be limited by the expertise of those using it. Principals and evaluators must develop the empirical expertise to use the 5D+ Teacher Evaluation Rubric – or any evaluation rubric – well.

Teacher Evaluation – Proceed with Caution

Moving from the historical teacher evaluation systems (satisfactory or unsatisfactory) to multitier systems provides much promise for giving teachers the feedback they need to improve instruction.

But the ramifications for rushing into implementing revised evaluation systems inadequately prepared could be costly for school districts.

Policymakers and educational leaders would be wise to carefully consider both the initial professional development and ongoing professional development necessary to ensure that evaluations are carried out reliably and consistently. Ensuring evaluator reliability across classrooms and schools is critical. Failure to provide adequate initial and ongoing training to evaluators could lead to costly challenges of evaluation results down the road.

School districts should beware of developing their own evaluation systems or altering an existing one. Developing a
research-based evaluation system is a huge undertaking, one that is prohibitive for the vast majority of school districts in both cost and time. And the prospect of altering or adapting an existing evaluation system risks losing the research base on which its validity rests.

Without doubt, revising teacher evaluation systems is a high-stakes investment – for policymakers, for school districts, for teachers, and ultimately for our nation’s school children. It speaks to equity, social justice and our commitment to future generations. By raising teacher effectiveness across the nation, we will raise the educational bar for all students, moving us closer to eliminating the achievement gap. Every child, in every school and in every classroom, deserves nothing less.

References


Rowan, B., Correnti, R., & Miller, R. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools. Teachers College Record, 104(8), 1525-1567.

Stephen Fink, Ed.D, is executive director, Anneke Markholt, Ph.D., is an associate director, and Sandy Austin, Ed.D., is a project director at the University of Washington Center for Educational Leadership. Fink and Markholt are co-authors of Leading for Instructional Improvement: How Successful Leaders Develop Teaching and Learning Expertise.
Research has shown that principal quality is second in importance only to teacher quality, among school-related effects on student achievement. Principal leadership affects student achievement through its influence on school conditions such as expectations for staff and student learning, staff quality, and organizational structures and culture. Coupled with the importance of quality school leadership, increased accountability demands from state and federal policy makers have pressed for the development of improved performance evaluation systems for school principals. Most new principal evaluation systems purpose to provide:

- A grounding in leadership qualities or processes that can help (teachers) improve student achievement; and,
- Rubrics that specify multiple levels of performance in enough detail to clarify the behaviors or competencies required of a high performing principal.

To achieve these ends, most standards-based leadership models consist of three basic components:

- A listing of behaviors and competencies (the standards and rubrics); and
- Support systems to help principals improve (e.g., feedback on current performance, coaching, professional development).

Yet, despite the call of state-level policy makers to mandate new and more accountable principal evaluation systems, little empirical evidence supports their effectiveness, particularly those that incorporate standardized leadership performance descriptors, such as the Interstate School Leaders Licensure Consortium (ISLLC) Standards for School Leaders (Council of Chief State School Officers, 1996).

It is not surprising that developers of standards-based models for principal evaluation see the utility in incorporating features taken from standards-based teacher evaluation systems. This tactic, however, may be problematic, as some teacher evaluation approaches have been found wanting. For example, Linda Darling-Hammond and other researchers (2012) have concluded that the new value-added teacher improvement models, designed to evaluate student test score gains from one year to the next, are “fraught with inaccuracies and inconsistencies” (p. 8). Research findings on teacher evaluation models have determined that teacher effectiveness ratings are more influenced by the type of student test administered, the non-random nature of how students are assigned to each teacher, and the multitude of many influences on student progress not measured or remediated by high-stakes accountability teacher evaluation models.

Indeed, inconsistent findings in studies of teacher evaluation systems should compel state policy makers to take pause when considering the development of new principal evaluation systems. Namely, principal evaluation systems should first be piloted and researched to determine their effectiveness in producing performance improvement. This article will focus on one such study that provides insight and guidance in ensuring fidelity in the application of new principal evaluation models.

The Study

The Wisconsin Center for Education Research consortium for Policy Research in Education (WCER) conducted a two-year study on an ISSLC standards-based principal evaluation system (Kimball, Milanowski, & McKinney, 2007). The study was funded by the U. S. Department of Education and collected both quantitative and qualitative data.

Who was in the Study?

The school district in the study had 61 elementary schools, 15 middle schools, and 12 comprehensive high schools. In addition, there was a special education school for children with multiple and complex disabilities, and three alternative secondary schools. Nineteen of the elementary schools operated on a year-round schedule, with eight on multi-track schedules and 11 on single track schedules. There are over 62,000 students, with an ethnic breakdown of about 60% white, and 40%
non-white. About 30% of the non-white students were Hispanic. There were close
4,000 certified staff (including teachers)
and over 300 administrators. The district
was led by a superintendent and three as-
sistant superintendents. The two assistant
superintendents that oversee elementary
and secondary schools supervise principals at
these levels and work with five directors who
also directly supervise school and principal
performance. Among the 88 principals in
the district half were randomly assigned to a
new principal evaluation system, while the
rest were assigned to the old system.

What was the Purpose of the Study?
The study intended to determine if the new
principal evaluation system:
• Influenced perceptions of district ex-
  spectations for leadership performance;
• Influenced the perception of feedback
  quality;
• Compelled principals to spend more
time on job facets emphasized in the
new system (instructional leadership);
• Led to improved performance as per-
  ceived by the principal’s supervisor.

What Data Were Collected?
All principals were asked to complete a
survey evaluating of the principals’ per-
tions of the new evaluation system’s utility,
clearly of expectations, and feedback quality.
The study compared both groups to see if
principals in the new system were placing
more emphasis on standards emphasized
in the new system; namely instructional
improvement. Interviews were conducted
of both the principals and their supervi-
sors asking about principal perceptions and
experiences with the evaluation systems.
Archival data sources included the prin-
cipal’s most recent evaluation, their School
Improvement Plan (SIP), and their annual
performance report. This combination of
data provided multiple measures of principal
goal focus and performance.

The New Evaluation Program
The old principal evaluation system incor-
porated performance criteria including
leadership, communication skills, group
processes, curriculum, instruction, perfor-
mance, evaluation, organizational manage-
ment, field management, and policy man-
agement, with a total of 44 descriptors and
three-part rating system; Highly Proficient,
Proficient, and Targeted Areas of Growth.

The new evaluation system was
centered on eight dimensions of principal
performance that included (a) vision, (b)
student achievement, (c) political, social,
economic, legal, and cultural context of
learning, (d) communication/interpersonal,
(e) instructional leadership/supervision, (f)
organization/management, (g) decision-
making, and (h) professional development.
Each dimension was represented by from
to 11 standards. The design team drew
from a number of competency models and
sources from the literature about effective
leadership practice including the perfor-
mance standards from ISLLC, the work of
Douglas Reeves (2004), and the California
Leadership Academy Standards. Each stan-
ard included a rubric with four levels of
performance as shown in Figure 1 below:

<table>
<thead>
<tr>
<th>Performance: 2.1 Understands student requirements and academic standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsatisfactory</td>
</tr>
<tr>
<td>Progressing</td>
</tr>
<tr>
<td>Proficient</td>
</tr>
<tr>
<td>Exemplary</td>
</tr>
</tbody>
</table>

Figure 1: Example of Performance Dimension and Rubric

The design team was explicit in its
test to make the new principal evalu-
ation system “look and feel” similar to the
teacher evaluation system that was adopted
by the district and modeled on A Framework
for Teaching (Danielson, 1996), with clear
standards and rubrics differentiating per-
formance on the standards. As such, the team
developed three tracks for the system. Track
1 of the evaluation system served as an in-
duction process, providing support for new
principals learning the district’s performance
standards. Evaluation was based on all eight
performance dimensions over a one-year
probationary period. Once a principal was
granted tenure, they moved to track 2 of
the evaluation system. Track 2 included
annual evaluations over a three year cycle
that entails one major and two minor evalu-
ations with more intensive focus on a few
of the standards. The purpose of this track
was to provide a structured, supportive and
evaluation rubric was adopted and used by
principals. For example, more principals in
the new system spoke of the importance of
instructional leadership as a goal. However,
the difference appeared little more than
semantic as all principals focused on the
importance of raising student performance.
Additionally, all principals reported the
importance of developing a culture of col-
laboration and achievement in their schools.
Principals also cautioned that expectations
may vary across principals due to variations
in school context and need and thus raised
concerns about standardization of goals or
behaviors.

In addition, principals in both groups
discussed the importance of informal “just-
in-time” interactions with their supervisor
as a primary source for their understanding
of district expectations. Principals felt that,
regardless of the evaluation tool being used,
these informal communiques from their

What did the Study Find?
Did the new system influence perceptions of
district expectations for leadership perfor-
ance? Yes and No.

While the study found that many
principals on the new system reported being
more clear about expectations, due to the
explicit nature of the rubrics, no statistically
significant difference was found between the
expectations of the principals on the old and
new evaluation system. The study did find
that the language incorporated into the new

Page 15
supervisor had more weight and provided more clarity, even though they were informal and spoke to individual context; that is, not standardized. In fact, real-life events in a school district influence the cultural and political environment and call for a flexible, formative, and mostly informal feedback system for principal behavior. For example, one principal reported frequent and specific feedback from her supervisor, because she was experiencing issues with the teachers’ union as a result of curricular and professional development changes she had been making. The study indicated that the most successful school leader was one who could monitor and interpret these shifts of expectations from their supervisor who were unpredictably influenced by emergency, legal situations, or parent/teacher controversies.

**Did the new system influence the perception of feedback quality? Yes.**

There was a statistically significant improvement in regard to the quality of feedback, indicating that principals evaluated under the new system perceived both more specific and more useful feedback than those in the old system. Principals under the old system had to request more useful narrative feedback because the old rating system was not as specific or useful in delineating specific leader behaviors. Principals in the new system indicated that the rubric (See figure 1) helped focus the supervisor’s evaluation meeting conversations by providing a common language.

However, beyond the formal evaluation meetings, which were limited, most principals reported difficulty distinguishing between feedback tied directly to their evaluation and the general advice and recommendations they were given in the more typical and informal interactions with their supervisor. The reality is that principals receive a cacophony of feedback from multiple sources regarding their performance and that this informal feedback is understood as just as important, if not more important, to successful performance as the formal evaluation feedback.

**Did the new system compel principals to spend more time on job facets emphasized in the new system (e.g. instructional leadership)? No.**

The new evaluation system emphasized developing and communicating the school’s mission, analyzing student achievement results, understanding student standards, and using technology more explicitly. However, no more time or effort was reported on these facets of the job by principals in the new system. Principals in the new system reported giving more thought to these new job facets, and including them as goals in the SIP. However, administrators in both the old and new system reported that their daily focus was based on the immediate needs of the school and not on what was listed or not listed in their evaluation document. Indeed, the principals in the new system appeared to include the new language in their personal goals simply because the wording was embodied in the evaluation tools and thus expected to be used. As a result, principals under both the old and new system used different terminology, but performed in similar and varied ways based on perceived school needs.

**Did the new system lead to improved performance as perceived by the principal’s supervisor? No.**

Principals in the new system reported more beneficial and authentic dialogue in the evaluation meetings with their supervisor. In addition, they reported the common rubric language made it a little easier and more explicit when crafting a goal statement. However, supervisors and principals ascribed varying level of importance to using a common and explicit language directed by a rubric. Some believed the standardization of language was restrictive and may lead to diminished flexibility and responsiveness to needs unique to each school. Overall, while anecdotal comments seemed to favor the standardization of the new system, no statistically significant difference in principal performance was evident. As one principal noted, “it’s not the evaluation running us” rather the needs of the students, staff, and parents.

**Implications For Practice**

Among the overall findings, the following are critical elements for policy makers when discussing the adoption of a mandated, standardized performance-based evaluation system.

1. The principal’s situation (school type, school problems, level of experience) substantially influences the impact of performance evaluation through emphasis on different goals and different amounts, types and sources of feedback. This raises the issue of whether one set of standardized evaluation measures can and should fit all.

2. Principal evaluation systems are only one of many ways that principals are given direction and held accountable. The potential for principals to hear a cacophony of demands is great, as is the tendency for district expectation to shift with the political winds. Communities, cultures, and people are ever-changing and unpredictable, calling for adaptive leadership.

3. Under any system, there can be substantial differences between how supervisors apply a principal evaluation system, regardless of how much it is standardized. This variability prevents true implementation standardization and thus evaluation fidelity. Just because a standardized evaluation system treats everyone equal, it does not necessary make it equitable or equally effective.

Policy makers may need to accept that frequently changing school and community environments require more situational or adaptive leadership goals and behaviors complicating the effectiveness of standardized educational reform models. Indeed, a system that standardizes goals may weaken, not strengthen leader effectiveness. Another thought is that evaluation systems themselves may not be the most critical solution to improving school leadership. Indeed, Jim Collins in his book *Good to Great* emphasized the importance of leader selection verses leader performance evaluation. Researchers (Creighton & Young, 2003) have suggested the need to distinguish *standards from standardization*, appealing to the need to maintain contextually fluid leadership and embrace the art versus the science of leading. Administrators have only a limited amount of time and energy. Because little evidence exists that standardized evaluation systems improve leader performance, districts and policymakers may want to focus their energies on the promotion and use of formative models to improve performance, and results-driven measures to determine effective leadership. In the end, these approaches, coupled with improved hiring practices, are likely to have much more influence over leader performance than increasingly standardized, detailed, and inflexible principal evaluation systems.

**References**


Thomas L. Alsbury is Professor of Educational Leadership at Seattle Pacific University and serves as Director of the national University Council for Educational Administration (UCEA) Center for Research on the Superintendency and District Governance. Dr. Alsbury also directs the Innovative Leadership Academy working with district leaders in high minority, high poverty rural North Carolina districts to build capacity and sustainability for student-directed innovations.

---

**Research-Based Instructional Strategies for Educators to Increase Student Achievement**

**Part I**

**Thursday, September 20, 2012 • Renton Community Center**

Registration information available in May (check www.wsascd.org)

Focus: Setting Objectives and Providing Feedback, Reinforcing Effort and Providing Recognition

This session is the first of a five-part series. Participants do not have to attend all sessions. Research shows that an individual teacher can have a powerful effect on his/her students even if the school doesn’t.

Dr. Stone will share the research related to what those highly effective teachers do, as presented in the McREL publication *Classroom Instruction that Works, 2nd Edition.*

**Subsequent Sessions – Save the dates!**

Part II – Tuesday, December 4, 2012 – Brief review of session one PLUS Cooperative learning; Cues, Questions and Advance Organizers; and Nonlinguistic Representations

Part III – Tuesday, January 22, 2013 – Brief review of sessions one and two AND Summarizing and Note Taking; and Assigning Homework and Providing Practice

Part IV – Wednesday, March 6, 2013 – Brief review of sessions one, two and three AND Identifying Similarities and Differences; and Generating and Testing Hypothesis

Part V – Thursday, May 9, 2013 – Putting it all together: How can school leaders facilitate good instruction throughout the district, school and classroom?
Two of the most influential in-school factors on student success are the effectiveness of teachers and principals. But how do we measure effectiveness? In the current system, educators receive an “unsatisfactory” or a “satisfactory,” telling them they are either not suited for the profession, or some degree of acceptable — with no clear indication of whether they are at the top of the field, need to improve certain skills, or fall somewhere in between. The passage of SB 6696 in 2010 started the ball rolling on an evaluation system designed to provide meaningful feedback to teachers and principals, support improvement when necessary, and give districts better insight into the effectiveness of their instructional staff. After two years of development and piloting, it became time to take the next step and begin the process for implementation statewide.

Satisfactory, or not

Today, most educators are told they are either “satisfactory” or not, with little feedback on where they really excel, and where they need to improve. In the current system, it is unreasonable to expect teachers and principals to regularly reflect on and alter their practice if they do not receive meaningful feedback on their practice. As former educators, we understand the lack of clarity that comes with receipt of an “S.” Does this mean I’m using data well? Am I using appropriate assessments for my subject matter? Am I an expert at differentiating my instruction, or a novice? Is my classroom well managed or just not chaotic? Our teachers and principals deserve more detailed feedback so they can truly develop, grow and improve their skills.

On the flip side, administrators need more detailed information on the staff they manage. If everyone in the district is an “S,” how can any manager be expected to know who works most effectively with specific colleagues, supervisors, or, most importantly, students? When staffing decisions are to be made, including those around placements and layoffs, staff strengths and weaknesses must be considerations in those decisions. Administrators and managers cannot make informed decisions in the best interest of students without a more comprehensive understanding of their human talent pool.

This, along with the (ultimately unsuccessful) chance of winning federal Race to the Top funding, was largely the reasoning behind creation and passage of SB 6696 in 2010. This legislation put Washington on the path toward a meaningful evaluation system for teachers and principals by:

• Moving from the two-tier system of “satisfactory” and “unsatisfactory” to a four-tier system;
• Establishing specific, relevant and uniform performance criteria for teachers and principals;
• Allowing the use of student data in evaluations; and
• Piloting evaluation frameworks prior to statewide implementation.

Piloting toward progress

The state selected 16 districts, located across the state and of varying size, to participate in the pilot program, beginning in the 2010-11 school year. These districts chose from leading instructional frameworks (e.g. Danielson, Marzano) to help develop evaluation tools and frameworks. Piloting of the models began in the 2011-12 school year, with full statewide implementation to begin in 2013-14.

Initial progress reports were encouraging, with educators and administrators in the pilot districts supportive of their progress, as well as cautionary on the amount of work necessary to develop the frameworks. These lessons were good examples of why piloting was a good first step in implementation. However, they also highlighted the need for additional changes to the system prior to statewide implementation.

First and foremost, the need for professional development and training around developing and implementing the new evaluation system became quite clear. This was followed by the need for a robust professional development system to encourage and support continuous growth by principals and teachers. Our system of professional development needs to continue to move away from one-size-fits-all to a more meaningful system for individual educators informed by identified strengths and skills.
in need of improvement.

Further, as a response to lessons learned, policymakers saw the need to tighten up some of the direction around requirements for the new evaluation system. This need prompted the governor and legislators to propose legislation in the 2012 legislative session to better prepare the state for full implementation.

Preparing for the next step

Based on the strong work of the pilot districts, state leaders recognized further direction was needed as the remaining 279 districts would soon begin their own transition to the new four-tier system. Despite different initial approaches, the governor and legislative leaders crafted compromise legislation, SB 5895, to support the next phase of implementation by:

• Creating defined performance ratings of (1) Unsatisfactory, (2) Basic, (3) Proficient, and (4) Distinguished;
• Requiring districts to adopt one of three frameworks chosen by the state;
• Including multiple measures of student growth data in three of the eight criteria for evaluation;
• Providing training for teachers, principals, and administrators on the new evaluation system prior to implementation;
• Aligning professional development with performance evaluation criteria;
• Requiring evaluations be a factor in assignment and layoff decisions, beginning 2015-16; and
• Phasing in the evaluation system over three years between 2013-14 and 2015-16 to make the process more manageable for districts.

To some, SB 5895 seemed premature, preferring to wait until closer to statewide implementation. While this approach may have worked in the end, policymakers chose to proactively enact changes to better guide statewide implementation. The adage of “what gets measured gets done” is certainly relevant here. The pilots — through no fault of their own, they were following the direction of the law — were not, in the eyes of state leaders, including all of the factors policymakers hoped to see included. Additionally, it was unclear if evaluations would be used for purposes other than staff feedback.

Districts and staff cannot be expected to intuitively follow policymakers’ hopes and wishes; they may be many things but they are not psychic. The only way the state can hold districts, schools, and staff accountable for desired results is to specify in law what the expectation is. Without speaking for them, it is clear that this is largely why the governor and Legislature took that next step and passed SB 5895. There can be no mistaking the intent for evaluations to inform professional development and staffing decisions.

From here to there

So, now what? Is it time to sit back and wait for statewide implementation? Not exactly. A compromise included in SB 5895 allows districts, not the state, to determine how evaluation results will be used in staffing decisions.

Washington operates under a “local control” model of school governance, where school districts are responsible for the majority of decisions affecting school buildings, staff, and students. Certainly the state must respect this governance model; however, the state, per the state constitution, has the responsibility of ensuring every child in the state has access to a “general and uniform” education system and provides the majority of funding. This is not to say state policymakers should angle for centralized management of the more than 2,200 schools in the state, rather that the state has a role in and responsibility for how systems are implemented across districts.

To ensure a uniform education system across Washington, the state should set the baseline or minimum requirements in the new evaluation system — not only for what is being evaluated, but also for how evaluations will be used. Most agree the evaluation criteria set by the state provides clear direction, sets a baseline, and allows for localization; it is yet to be seen if the same can be said for the language around use of evaluations in staffing decisions. To avoid 295 different interpretations of the law, tracking is required at all levels, especially by state policymakers, to ensure implementation follows intent.

Every student deserves to attend schools led by effective educators, and every educator deserves the support to be his or her very best. One of the key tools in accomplishing this is a meaningful evaluation system for teachers and principals that not only supports and improves individual educators’ practice, but also helps districts decide how to staff buildings in ways that best support and improve student learning. This is why including student learning — using multiple and varied measures, not just standardized test scores (but those too) — in the evaluation criteria is important, not only to provide additional insight into individual educators’ impact on student achievement but also ensure student learning is kept at the forefront.

And really, that is what is at the heart of our education system — students and their opportunity to compete in today’s workforce. If districts, schools, educators, and the state are not constantly evaluating whether their policies, decisions, and daily-to-day practices (instructional or otherwise) are driving the highest levels of achievement possible for each student, we are failing our kids. And no one wants to do that.

Chris Korsmo is CEO of the League of Education Voters. A first-generation college graduate, Chris knows first-hand the transformative power of education. Chris began her career as a high school teacher and coach before spending 15 years working in reproductive rights advocacy.

Heather Cope is policy director at the League of Education Voters. A Washington native, public school graduate and recovering journalist, Heather began her career as a middle school teacher. Heather spent some time in the education policy world of the other Washington before returning home.
My good friend is a visual artist who creates oil paintings of various community scenes from urban to rural settings. I recently visited my friend while he was preparing for an exhibition. He had committed to complete about forty new original paintings within about six weeks. I later attended the exhibition and was astonished. Beautifully displayed, my friend’s paintings varied from large six-foot canvases to 5x7 inch gems. While each had a recognizable style, every one exhibited unique and compelling features, combinations of colors, and conveyed a wide range of moods and emotions.

I asked, “How much time does it take you to create one of these painting?” His answer took me by surprise. He said, “Each painting takes about thirty years.” When my puzzlement wore off I realized what he was revealing. Had my friend not studied the craft of painting, had he not practiced the skill of seeing and the art of portraying images on canvas for the past thirty years or more, he would not be proficient enough to create such a display of work in such a short time.

It is not that my friend has not made very good art continuously over the previous thirty years. However, the adept and efficient manner in which he can produce a painting is much different now than when he began years ago.

Teaching and leading in schools, as well as evaluating these practices, can be compared with the work of such an artist. A teacher can be proficient on certain levels early in his or her career. However, the art of teaching can and must develop and improve over time. The same level of proficiency cannot be expected from an early career professional as expected from an experienced one. Likewise, just as it is difficult to evaluate a painting, with various criteria emanating from various judges and purposes, it is difficult to evaluate teaching. No two people observe the products of the painter or the teacher from exactly the same perspective, nor do they judge those results with exactly the same set of personal criteria, no matter how “trained” an observer he or she may become.

Judging teaching against a set of standardized criteria and using student performance on standardized tests is like making an accomplished artist or even an emerging painter create paint-by-number replicas of someone else’s design. It will be flat, emotionless, and lack the inspiration desired and needed to evoke student’s passion to learn. Learning as we know from brain research is dynamic, individual, and inspired. It is not programmable like a robot. A strictly systematic approach to teaching might be useful as an early training technique to master the use of the tools of the profession, but it will never be effective to create an inspirational work of art, an inspirational teacher, or an inspired young learner.

A few weeks ago, a principal and I observed a writing lesson taught to young students. The observations were through the lens of our new pilot evaluation model, based on the Charlotte Danielson, Framework for Teaching. In this case, the principal and I are practicing with and learning how to use the evaluation tool. That is, how to observe and judge the applied techniques and the expressed art of teaching taking place in the classroom. Mind you, observing only one lesson is like watching my friend the artist put on the first layer of paint that will become the background for a myriad of artistic moves to come. So it is the teacher’s task, to artfully build upon each lesson until students master a concept or skill.

This classroom observation was a culmination of nearly a year and a half of conceiving, planning, and learning until we reached this first step in implementing the Teacher Principal Evaluation Pilot (TPEP). Bear in mind, we have had limited training or practice with this new system. It quickly became apparent that our observations, while holding commonalities, bore striking differences in the evidence and perceptions we gathered. Consider that the two of us, while observing the same lesson, have different content backgrounds, different family and cultural experiences, and have had very different training from different colleges of education. Furthermore, we began our college and educational careers more than 20 years apart. We have encountered different mentors, and likely possess different attitudes and dispositions about what we judge to be good teaching. It is truly amazing that our conclusions about the quality of what we observed are so close. We believe it is in no small part due to the fact that we have adopted a common framework that describes in detail what we know about and can agree is good teaching.
Observing and judging good teaching is a bit of an art itself. It is comprised of knowledge and skills and tools. However, proficient, reliable, and useful the tool, judgment still comes from knowledge and experience and the ability to extrapolate from the point of observation into the past and into the future. It requires skill to discern intent and motive from empirical observations that are simply a snapshot of complex and extended processes about teaching and learning. It involves collecting observational evidence that is not laden with value judgments and that contains only the facts of what occurred.

Thinking back to how we arrived at this observation on this day, I recall sitting in a regional WASA meeting in the spring of 2010 listening to Alan Burke, Deputy Superintendent at OSPI. Dr. Burke gave an update about the recently ended legislative session. He was speaking about the reform bill E2SSB 6696, and in particular the new teacher and principal evaluation criteria and the pilot projects it created. I vividly remember him saying it would be useful if a group of small schools would join together in a Teacher and Principal Evaluation Pilot (TPEP). I looked across the room and caught the eye of a colleague who I knew was, as we were in Pullman, considering applying to be a pilot school district. We later talked and decided to join with NEWESD 101 to submit a consortium grant for TPEP. This launched a journey along a path that was yet unknown. Eight school districts joined the consortium ranging from under 100 to over 2300 FTE student enrollment.

In the beginning, each of the consortium districts was in a different place in defining best professional practice for teachers and principals. In addition, we found that within our own district there was a wide range of knowledge and understanding of the research about best professional practices. Different instructional frameworks had been introduced in different schools. In some, the only experience with instructional frameworks was what individual teachers had read, or learned about in workshops and at conferences. Even principals had diverse knowledge and understanding of instructional frameworks and rubrics.

There were times when it was tempting to turn back and abandon the journey. Balancing the work of learning with our own district’s challenges to learn together and then turn to integrate that work in concert with seven other districts was at times frustrating and gruelingly slow. Not having a developed path to follow, sometimes we would set out in a direction and later retreat and turn in a new direction until we could find the best way forward.

We began by trying to “unpack” the eight criteria listed in state law, E2SSB 6696. Our purpose was to identify elements of these criteria that were essential to each. We studied rubrics and rubric writing more deeply. We attempted to devise our own language to describe our collective knowledge and understanding of the evaluation criteria and the professional practices that contribute to successful and effective teacher and leadership results. It gradually became clear that there had been individuals and institutes that had spent a decade doing exactly what we were trying to do in several hours throughout a single school-year. We considered the work of Charlotte Danielson, Robert Marzano, and the Center for Education Leadership, among others. It was when the State contracted with Charlotte Danielson to review and advise the State system that we concluded that her Framework for Teaching would become the framework for the consortium.

Pullman Public Schools consist of about 2500 students (2300 FTE) with pre-school, three elementary schools, one middle school, and one high school. At the district level we began planning for the pilot with a steering committee that would interface with the consortium. This committee consisted of an elementary teacher (the association bargaining chair), a high school principal, and the superintendent. In addition, we identified a district work group made up of the steering committee, a teacher from each school, and each principal and assistant principal, and the assistant superintendent. This later group came together to plan and lead the District’s communications and professional development for the implementation of TPEP. Some of them ultimately became participants in the pilot.

In the spring of 2011 the teachers’ association and the school board agreed to a memorandum of understanding (MoU) for teachers participating in the pilot. The MoU outlined participation expectations, employee rights, and the number and demographics of participants in each school. It was important to limit risk to teachers who would participate to assure that they are able to trust that we are evaluating the tools and processes, and that their individual employment status was not put in jeopardy simply by their participation in the pilot.

Three teachers were identified in each school teamed with each principal and assistant principal. These became the first year pilot participants. Early, mid, and late career teachers were identified for each pilot team in order to help surface issues unique to evaluating each of these types of teachers.

The pilot work so far has led us to a number of important conclusions. The process of viewing professional practice through the lens of evidence of student learning is complex and difficult. Our past practice has been more focused on teacher behaviors and practices with minimal checking on the learning that is or is not taking place. Deciding what evidence is available, what it means, and how to respond to it demands new thinking about the relationship between teachers and principals. To be effective, the principal and teachers must become much more collaborative. The principal must be much more engaged in knowing the learning needs of students and the instructional practices and curriculum used by teachers. Conversations are more reflective and involve a more comprehensive examination of teaching and learning in the classroom on the part of both the teacher and the principal. How to shift time from other administrative demands placed on principals and central office administrators has to be considered and resolved. Furthermore, if the process is to produce desired results, more engagement in the evaluation process by teachers and more time on the part of teachers, principals and central office administrators will be required.

The emerging process is like coaching a professional who seeks to improve techniques that will lead to improved performance or produce an exceptional work of art. By reflecting on objective observation data, teachers and principals become motivated to examine their own practice and often to further exploration and learning. The process has developed to be less about judging from a narrow set of criteria whether teaching is good or poor. Just as a painter may elicit critique and even invite observation and comment about a certain technique, so should the interactions between the principal and teacher evolve. Judgments about teaching must encompass not only the components of teaching, teacher behaviors, but also the results, the learning and development of individual students.
For us, creating and using reliable evidence of student learning is emerging as a significant system challenge. Standardized tests are not very useful to inform instruction and to monitor student learning progress. They may be useful in providing some evidence of system-wide progress, but are not very useful when trying to discover how to help Susie and Johnny master learning on a day to day basis. What is needed are classroom based formative assessments closely aligned with the classroom and district curriculum, narrowly focused on the significant mastery of “big ideas” students need so they can be successful future learners. Of course, the time and skill required to reflect on the evidence of student learning, both individual and groups of students are significant. In the end, we believe the improved evaluation tools and processes will lead to steadily improving and more artful teaching. If we expect that a new evaluation system will suddenly create masterful teaching, and immediately result in every student becoming a masterpiece of learning and achievement, there is potential for considerable disappointment.

The new evaluation criteria, a new system of evaluation based on solid research about best instructional and leadership practices, and a system more reliably and equitably deployed across our state is needed and brings great promise. Simply centering the focus on individual student learning and the relationship between that goal and the art and practice of teaching and leading has already brought more professionalized conversations among educators. In order to develop the art of teaching and leading among teachers and principals, the new model of evaluation needs to foster collaborative interaction and learning, cooperative problem solving, and the shared ownership of learning for each student. The evaluation system must be a means to assist professionals to continuously gain and develop knowledge and skill, and like an artist, improve their ability to apply techniques creatively, innovatively, and inspirationally.

Paul Sturm has been the superintendent for Pullman Public Schools for the past six years. Before becoming superintendent, he also served as assistant superintendent in curriculum and instruction. Paul began his career in the Central Valley School District where he worked as a high school principal, assistant principal, and high school math and science teacher for over 20 years. He currently serves as the president-elect for Washington Association of School Administrators (WSA), and he has previously served on the WSASCD board.
Learning from the Journey of the NEWESD 101 TPEP Consortium Pilot

“The real voyage of discovery lies not in seeking new landscapes but in seeing with new eyes.”
—Marcel Proust

In the spring of 2010, eight school districts embarked on a journey to pilot the new teacher and principal evaluation system. Almira, Davenport, Liberty, Medical Lake, Pullman, Reardan, Wellpinit and Wilbur made the decision to form a consortium and collaborate with their regional Educational Service District to coach and support them through the process. The discoveries along the way have provided invaluable opportunities to “see with new eyes.”

Northeast Washington Educational Service District 101 is comprised of 59 public school districts, ranging in size from the smallest with 8 students to the largest near 30,000 students. This article will focus on how member districts - specifically one smaller rural district and one mid-sized district - embraced the work and operated from the premise that this revised system is truly about professional growth - not a “gotcha” model.

Medical Lake School District is mid-sized with 1,910 students in the K-12 district. An instructional framework was in place at the time that the pilot began; however, the evaluation rubric was outdated and focused on the general criteria of “satisfactory,” “unsatisfactory” and “needs improvement.” It did not provide for the dynamic process of reflection on teaching practices nor the opportunity for imbedded professional growth.

Almira School District’s Journey

Over the decades, educators in Washington state have been asked to provide quality instruction and ensure that learning takes place for each and every student without requiring an evaluation tool that helps guide their success through a professional development model. When E2SSB 6696 for Washington state was passed, a grant was provided to allow pilot districts to experiment and have a voice in the new teacher and principal evaluation system. Our small, rural school district in eastern Washington decided to take part in that quest, along with seven other school districts, using Northeast Washington Education Service District 101 as our support. In the Almira School District, students and their success are front and center in every decision we make. To ensure our actions are purposeful, our leaders must possess tools that will assist them in giving meaningful, daily feedback to teachers on quality instruction and student learning.

Almira’s story began with introducing the idea of participating in the pilot grant to our teachers at a staff meeting. The superintendent/principal’s first task was to disseminate the new state law, E2SSB 6696. Throughout the 2010-2011 school year, the superintendent/principal continually shared with staff information garnered through the TPEP meetings and trainings. In addition, teachers’ comments, concerns and feedback were taken to each TPEP planning meeting. The staff was given each draft version of a teacher evaluation rubric as it was created. Teachers’ likes and dislikes, concerns and questions were noted. Many of the teachers had the same concern; the new process would require teachers to do more when their plates were already full with adopting new curriculum aligned with the common core standards. They also were somewhat suspicious about what the “real agenda” was for this new evaluation process. After a lengthy discussion of the importance of their voice in this new process and digging into Charlotte Danielson’s Enhancing Professional Practice: A Framework for Teaching (2007), teachers were asked to join together and embrace this as a positive change for their ongoing professional development.

Each full-time teacher in the district agreed to a memorandum of understanding to become a part of the pilot for the 2011-2012 school year.

The pilot is not only delving into the teacher evaluation process but also the principal evaluation process. Principal evaluation is problematic because Almira School District has a dual position of superintendent/principal. To tackle this issue, many ideas were tossed around and we finally settled on dissecting the principal evaluation tool, putting it out on a survey and allowing staff to anonymously provide feedback. The new process is a paradigm shift, and administration and teachers have to trust one another to implement the process with integrity. When quality instruction occurs in the classroom and all children are learning, their futures become brighter.
The Almira staff would admit to you that this is very hard work and sometimes difficult to wrap their minds around. But, in Almira, we are learning new processes for evaluating the art of teaching and learning in ways to improve together. The belief is that the district/building leader is a learner first with administrative duties being secondary. The changes throughout this journey have ignited feelings of both enthusiasm and fear. Optimistically, using a four-tiered evaluation tool will help to promote professional growth with the knowledge that not all teachers can live in the distinguished tier all the time. We need to continue to reassure our teachers that this is not a "gotcha" system – rather one about professional growth. Will this take time? Yes! Are we thankful that we are a part of the pilot? Why? Because we have been enriched with what we have learned, are learning currently, and will be learning along this road. Our journey continues. We have not yet reached our destination as we continue to learn and refine our process. We are truly operating as a dynamic professional learning community.

Medical Lake School District’s Journey

Our journey to reinvent our process for evaluating teaching practice began over ten years ago. Little did we know the amount of time and effort that would go into this process. Our problem was age old – the process we were using was a dinosaur. It provided teachers with simplistic evaluative comments without consistency as to what it meant and provided no guidance on where teachers should focus their improvement efforts.

We convened a steering committee, conducted a book study on the most recent research in teacher evaluation and developed a rubric from Charlotte Danielson’s Professional Practice: A Framework for Teaching (2007). We thought we had climbed the mountain, made the big change, moved into meaningful evaluation. We implemented the new rubric and set about observing and providing feedback. We made numerous mistakes. (1) The evaluation ratings were still comprised of: unsatisfactory, needs improvement, satisfactory, and outstanding. We did not have a consistent and common definition or understanding of these ratings. Teachers began to compare ratings with each other and would say to their principal, “Mrs. Jones received an “outstanding” in the area of classroom management, yet I only received a satisfactory. I think I do just as much as she does.” Principals and legal counsel advised dropping the “outstanding” rating. (2) We neglected to provide training to the evaluators. Evaluators viewed teaching through different lenses. Inter-rater agreement was in the low 25%. (3) We did not provide staff development nor an opportunity for reflection on our practice. We continued with this “new” evaluation rubric and realized from the beginning that we still did not have a process for continuous improvement of our teaching practice.

In 2003, we began to operate as a “professional learning community.” We offered and encouraged teachers and administrators the opportunity to receive staff development from the “experts,” Richard DuFour, Becky DuFour and Robert Eaker. We assimilated Rick DuFour’s quote, “a good school is not a collection of good teachers working independently, but a team of skilled educators working together to implement a coherent instructional plan, to identify the learning needs of every student, and to meet those needs” into our work.

In 2004, we trained and implemented an instructional framework using the STAR protocol from the BERC group. Now we have teacher leaders that lead staff through the process of observing peers and debriefing lessons.

We have come full circle by being a member of a consortium of eight school districts involved in the Teacher and Principal Evaluation Pilot.

What have we learned from this experience?

We are continuing to use Enhancing Professional Practice: A Framework for Teaching (2007), developed by Charlotte Danielson to evaluate our teachers. Our work as a consortium has provided us with a consistent definition of good teaching. Through our training and practice we know good teaching when we see it. Everyone involved in the pilot - teachers and principals - have a shared understanding of a common language to describe quality teaching. This shared understanding allows us to have rich conversations around effective professional practice. Our discussions center on Danielson’s four domains: planning and preparation, classroom environment, instruction and professional responsibilities.

We are becoming “expert” evaluators of teaching and learning, able to recognize examples in each of the four domains. We do not make judgment calls, but rather collect and interpret evidence against four specific levels of performance: unsatisfactory, basic, proficient and distinguished. By having a consistent definition and a shared understanding of good instruction we are able to have fair, reliable and valid evaluations of teaching and learning.

Evaluation of teachers and principals has a legal purpose. The laws in our state require this. However, the legal aspects were not the only reasons for wanting a change in the evaluation system. As dedicated professionals, we wanted an evaluation system that was also a professional growth model. We wanted quality assurance while promoting a professional learning community where all are constantly improving their practice. It has been a challenge merging the two purposes. We are making progress but still need further clarification and fine-tuning of the system. We still have formal observation but also include the elements of personal self-assessment, structured pre-conferencing that allows the teacher to “set the stage” with the evaluator, and a post-conferencing process that provides the opportunity for rich dialogue around what the evidence is saying, steps to improve our practice and plans, and goals for continued professional development.

Certainly the journey has resulted in experiencing evaluation through new eyes. The benefits have been worth the investment.

References

Helene Paroff serves as Assistant Superintendent of Student Learning and Support Services at NorthEast Washington Educational Service District 101; a position she has held for 14 years. This year, she has been assisting OSPI with the rollout of the ESD-coordinated TPEP Regional Implementation Grants. Helene has held positions as a special education teacher, department chairperson, and curriculum development director in Maryland and New York prior to moving to Washington.

Shauna L Schmerer serves as Superintendent/Principal for the Almira School District, a position she has held for five years. Before joining Almira, Shauna worked as a special education teacher, a middle school teacher, a building LAP coordinator, and a counselor for the Central Valley School District for 11 years.

Dr. Pam Veltri is the superintendent at Medical Lake School District. She has held that position for 12 years. Previously, she served as a classroom teacher, special education teacher, special education director, director of teaching and learning and assistant superintendent.
I've been a part of the work sparked by SB 6696 for the past year and a half. I work within a consortium made up of teachers, principals, and superintendents from small districts all over Eastern Washington, supported by EDS 101. Our work on the Teacher, Principal Evaluation Pilot (TPEP) has been an interesting, growth producing process.

When we first began, all we had to work with were the eight criteria established in SB 6696 and the mandate that teachers be rated on a four-point scale. Neither the criteria nor the four-point scale came with any description or vocabulary. The task of choosing labels for the different levels was more difficult than I would have imagined. Obviously, it was easy to identify the bottom, below standard, level, but from there the wording was tricky. What do you call the level that is just one up from below standard? How do you define the level that is at the very top of the scale? These were questions our group wrestled with. In addition to those decisions, we needed to write descriptions that elaborated on the eight criteria. As a group, we labored to design a scoring rubric that contained word labels for the four levels and began to write descriptions of what each criterion looked like at each of the four levels. It was a time-intensive, conversation-intensive process.

A New Process

This past December, I was the first teacher in our district to be observed and evaluated using the new evaluation system. It was a process far different than anything I had been through since being hired as a full time teacher. It was actually reminiscent of the work I did with my student teaching supervisor, where I submitted a plan for my lesson and then was observed while teaching, and then had a conversation following the observation; these conversations after the observation were focused on improving my teaching practice and student learning.

Over the past 12 years of full-time teaching in two different districts, evaluations of my teaching have involved a principal’s observation followed by my receiving a typed form in which the principal summarized the lesson and gave me a score of satisfactory – the highest possible score. Often my evaluations would also include a conversation with my principal about the lesson, which mostly revolved around my administrator sharing his or her opinions about things seen in my classroom.

In sharp contrast, the observation and evaluation process that we are piloting is evidence-based. In the past my administrator may have written something like, “The students were involved in an interesting lesson.” Whereas my most recent observation included pieces of evidence stating things like, “Student was writing in response to prompt and asked peer for help in identifying the location of a quote from the text. The two students worked together to identify the quote.” The evidence gathered was completely unfiltered by a global view of who I am as a teacher; instead it was focused on my practice that day and the student learning that was produced.

The observation was simply a process of gathering concrete, factual pieces of evidence - what I said, how I said it, how the students responded, etc. In the conversation that followed, I was asked “why” and “how did you know” kinds of questions, and I had to look at the evidence to shape my response. The evidence, and evidence alone, was used to evaluate my lesson. Throughout the process, I needed to reflect on my decisions and processes at a level that I hadn’t really done since working on my National Board Certification. Reflecting on my practices and making changes based on evidence is a significant part of my growth and development as a teacher, the evidence-based conversations empower self-reflection because they are rooted in fact and not loosely based on opinion and conjecture.
Time to Implement

   The tool has the potential to be powerful, and training of all involved will be a lengthy process. However, if we don't provide adequate time once people are trained and are trying to implement the tool, it will be useless. Evaluating teaching at this level of thoughtfulness and preciseness is a very time intensive endeavor. This evaluation tool requires a time commitment from both teacher and administrator like nothing else I have ever seen.

   From start to finish, the process of this goal setting and one observation (one of two that I will be required to have this year) took over 5 hours. The time commitment was similar for my administrator. This year a team of 3 teachers are piloting the new evaluation tool.

   To date, with 3 of the 6 observations completed, my administrator has had 3 goal setting meetings (1 hour each), 3 pre-observation meetings (45 minutes each), and 3 post-observation meetings (45 minutes each). He is half-way through the process, already having spent 2 ½ hours in meetings for one 45 minute observation. And the scary reality is that we are only piloting 2 of the 4 domains.

   To put this into perspective, my principal will ultimately be responsible to evaluate 20 staff members. That means 20 meetings for goal setting, 40 pre-observation and 40 post-observation meetings (1 for each observation), 40 observations, and then 20 end-of-the-year final evaluation meetings. That means a total of 160 meetings in a 180-day school year; it means on average, if each meeting only lasts 45 minutes, my administrator will have at least 120 hours of meetings during the year. On top of that is the time spent preparing for each meeting – a very conservative estimate would put each of the goal setting, pre/post observation, and end of the year evaluation meetings at 30 minutes each or a total of 3 hours additional time spent preparing for each teacher. That added to the actual meeting time puts the total time requirement at approximately 180 hours. There is not an administrator in the state who has 180 extra hours. If this system is to be successful, and if we truly see our principals as the educational leaders of our schools, then something must be done to empower them to be effective. The process is a good thing and worth the effort to rethink how we do school. If we want this new evaluation tool to be successful we will:

   •   Re-define the role of administrator as educational leader,

   •   Create new positions,

   •   Provide intensive training for all involved parties.

Amy Bumpus is a National Board Certified teacher of 7th grade Language Arts at Lincoln Middle School in Pullman. She loves watching her students grow and develop as readers and writers.
As you walk into Mrs. Cliff’s 4th grade class the first thing you notice is order. The room looks lively and active, but everything operates in harmony and with a purpose. Mrs. Cliff is not fazed one bit that the six of us (four teachers, one union president and one OSPI director) troop into her classroom holding clipboards. She effortlessly moves her students from one part of the lesson to the next. Her transitions are flawless. I think back to my days with my 9th grade language arts students and feel embarrassed that I could not quite muster the kind of seamless fluidity that she has with these 10 year olds. We watch, and take notes with our instructional framework rubrics in hand as she leads the class through a read-aloud. Not one student is distracted or talking. They are fully engaged as she employs several techniques, such as accessing prior knowledge, questioning the text, and interpreting the content.

The bell sounds and another transition ensues. The students line up at the door ready for the afternoon assembly. As they are filing out of class, Mrs. Cliff leans over to me and whispers, “My microwave is unplugged.” I can only assume this comment is meant for someone else, but I am the only one standing in the area. I reply “O.K. thanks” and she takes off like a flash after the last student turns the corner and down the hall. I am puzzled by this comment, but when I want something fast and filling, I can throw something in it and in a matter of minutes feel full. It may not be as tasty or as satisfying as taking the time to prepare a meal in the oven, but it will get the job done. I have had this rushed feeling again and again in the TPEP project. We are all under tremendous pressure to “nuke” the work, instead of carefully and thoughtfully “baking” the process. We, as educators, need to resist this urge. From the classroom teacher to the superintendent, we must put cynicism aside and make the most of this change by embracing it early, rolling up our sleeves and digging into the critical question: What does quality teaching and leading look like?

Please don’t mistake my wish to slow down as a cry for delaying the legislative timeline. I feel a great sense of urgency about this work and know that it can be the single greatest change that will have the most long-lasting impact on our system for this and future generations of students. I know this work can change practice and thus impact student learning in the most meaningful ways. I spent 10 years working with teachers in their beginning, middle and veteran years around induction, and professional and national board certification. This work forced me to analyze teaching practice, provide feedback, ask probing questions and ultimately forced me to critique my own teaching. That work took time. In many cases it took the entire year working on discrete areas of a teachers practice. The ongoing conversations, feedback and support we give to both teachers and principals related to the new definitions of educator effectiveness will constitute the heart of the new teacher and principal evaluation system.

I love to cook. In building our new house, we designed a kitchen with double ovens. Our microwave sits in our laundry room, a little white Emerson shunned because it does not match our new stainless steel appliances. I rarely use it, but I know when I want something fast and filling, I can throw something in it and in a matter of minutes feel full. It may not be as tasty or as satisfying as taking the time to prepare a meal in the oven, but it will get the job done. I have had this rushed feeling again and again in the TPEP project. We are all under tremendous pressure to “nuke” the work, instead of carefully and thoughtfully “baking” the process. We, as educators, need to resist this urge. From the classroom teacher to the superintendent, we must put cynicism aside and make the most of this change by embracing it early, rolling up our sleeves and digging into the critical question: What does quality teaching and leading look like?

Please don’t mistake my wish to slow down as a cry for delaying the legislative timeline. I feel a great sense of urgency about this work and know that it can be the single greatest change that will have the most long-lasting impact on our system for this and future generations of students. I know this work can change practice and thus impact student learning in the most meaningful ways. I spent 10 years working with teachers in their beginning, middle and veteran years around induction, and professional and national board certification. This work forced me to analyze teaching practice, provide feedback, ask probing questions and ultimately forced me to critique my own teaching. That work took time. In many cases it took the entire year working on discrete areas of a teachers practice. The ongoing conversations, feedback and support we give to both teachers and principals related to the new definitions of educator effectiveness will constitute the heart of the new teacher and principal evaluation system.

After completing our discussion of Mrs. Cliff’s lesson, I sat with the workshop facilitator, assistant superintendents and union president and discussed the latest from Olympia and how it may affect their district’s ongoing work. Everyone from the instructional rounds and calibration work that day had gone home and it is now just the four of us discussing the road ahead. I mention the microwave comment and
they all break into howls of laughter. Just two days prior to my trip, the district had announced a ban on all microwaves in classrooms because of a two-year old lawsuit filed against the school district. I just chuckle as I think of Mrs. Cliff and her concern over the contraband microwave. During my drive home, I reflected on the push to microwave the process of changing school cultures, building professional learning for educators and supporting the complex task of impacting student learning. All three of these things take a tremendous amount of time, resources and unfailing dedication. While we cannot expect the political process to slow, every district must begin now in order to bake in the new culture and transform the support for teacher and principal professional growth.

Michaela Miller is a National Board Certified Teacher who is currently working at the Office of Superintendent of Public Instruction. She is the director of the teacher and principal evaluation project, the national board certification program and the beginning educator support team. Michaela attained her Washington State principal certification in 2009 and renewed her national board certification in 2011.

---

**Become a Washington ASCD Member!**

**Over 2,000 Educators Have Joined WSASCD!**

Washington State ASCD is the only educational organization in Washington State that reaches practitioners at all levels of education and your involvement is one of the keys to our success. As a member of WSASCD, you can take advantage of our member benefits.

**Receive:**

- *Curriculum in Context*, an award-winning journal addressing timely and challenging issues, plus feature articles from Washington educators.
- On-line newsflashes (through MailChimp) furnishing current information about WSASCD events and activity, awards and board action.
- Reduced registration fees for all professional development programs.
- *Practitioner’s Best Friend* tips, strategies, and exchanges via the Web.

**Connect:**

- WSASCD Website offering 24-access to information on professional development, association activities and valuable resources.
- Leadership and diverse networking opportunities enabling members to share resources, face challenges together and explore new ideas.

**Attend:**

- Regional workshops covering topics relevant to the needs of local educators.
- The statewide Annual Conference, mixing national experts with statewide presenters and practitioners to share ideas and promising practices – one of the state’s premier education conferences.

Visit [www.wsascd.org](http://www.wsascd.org) and click on Member Benefits for more information.
A message from the executive director

Executive Directions

On behalf of the Washington State ASCD Board of Directors, I want to say thank you to all the educators who have seen value in learning and networking with us this year. The success of our affiliate can be directly attributed to some important decisions made by our Board of Directors, with you and the education of our children in mind. Among those decisions, was the choice to delve deeper into the ASCD Whole Child Initiative. Our Board completely revamped our awards program to reflect the five tenets of the Whole Child. At our Annual Conference in October 11-12 in Seattle, we will honor the recipients of our new State Recognition Awards Program with the Healthy Schools Award, Safe Schools Award, Student Engagement Award, Supported Students Award, and the Challenged Students Award.

Other decisions made by the Board of Directors were related to the professional development opportunities provided by WSASCD. Knowing that educators need quality professional development and understanding the financial challenges of schools and districts, our Board made the decision to continue to provide our regional workshop series, related to Professional Learning Community Concepts, at a price that was reasonable for educators. In addition, WSASCD partnered with OSPI to present two symposiums to provide information on the Common Core State Standards. On behalf of our wonderful Board of Directors, I would like to thank Federal Way Public Schools, NEWESD 101 (Spokane), ESD 112 (Vancouver), and Central Valley School District for providing sites for our series of workshops this year. For the 2011-12 school year, we plan to offer a new series of five workshops related to Research-Based Instructional Strategies for Educators to Increase Student Achievement. This series, presented by Dr. BJ Stone, co-author of Classroom Instruction that Works, 2nd edition, is intended to provide support to teachers and administrators for the Teacher Principal Evaluation Pilot (TPEP). Dr. Stone's information will benefit school districts that have adopted the Danielson Framework for Teaching, the CEL 5 Dimensions of Teaching and Learning, as well as the Marzano Teacher Evaluation Model.

The decision was made by our Board of Directors to continue our partnership with OSPI for our 2012 Annual Conference. We have been pleased that OSPI has utilized our conference to provide updates on new standards and curriculum alignment issues. This year, we will bring in a third partner, as we work with the Washington Association for School Administrators (WASA) to provide a rich, balanced program for teachers and administrators. This year’s conference theme, Empowered Learners: Teaching and Leading through Standards, will involve experts on all facets of education at the P-12 and university levels, presenting a wide array of information relative to Common Core State Standards (CCSS), Teacher Principal Evaluation Pilot (TPEP), Assessment, Grading and Reporting, and so much more. We hope you will join us at the Seattle Doubletree Airport Hotel on October 11-12, 2012 for this great learning and networking opportunity.

The journal you are currently reading is the result of yet another important decision made by our Board of Directors, as we continue to publish Curriculum in Context, online. The Board would like to thank our departing co-editors, Dr. Becky Cooke, Superintendent, Deer Park School District; Dr. James Howard, Associate Professor, Washington State University; Dr. Gene Semeni, Assistant Superintendent, West Valley School District, Spokane, for pioneering the eJournal. WSASCD greatly appreciates the fact that Washington State University hosted our journal for the past six years. We are pleased to have our journal now hosted by Seattle Pacific University, with the co-editorship of Dr. David Denton, Assistant Professor, SPU School of Education, and Dr. Michael Dunn, Superintendent, Northwest ESD 101, Spokane.

Board members contribute countless hours of conference calls and meetings, providing a critical level of insight into issues that guide our Association forward. I would like to thank the following members of the 2011-12 Board of Directors, who will be completing their term of service, for their perseverance and involvement in so many critical decisions.

by Kathy Clayton
The contributions of Josh, Andrew, Barry, and Jose, as well as our entire Board of Directors have been centered on supporting you, as educators, so you can best serve the children of our state. In these challenging financial times, remember that you count! As you go about your daily life as an educator, remember that the work you do with children and your community does make a difference. On behalf of WSASCD, thank you for your caring, your strength and your belief in children.
Common Core Standards

In 2011, Superintendent Dorn formally adopted the Common Core Standards for Washington. The new standards will be implemented during the 2013-2014 academic year. What are the Common Core Standards? How will the new standards affect curriculum, instruction, and assessment? What are the implications of these new standards on teaching and learning? How are district administrators, principals, and teachers preparing to implement the new standards? What steps need to be taken to prepare for 2013? These are some of the questions under consideration in the next issue of Curriculum in Context.

The editorial staff invites you to submit a manuscript on this topic by September 10th to David Denton (dentod@spu.edu). Final manuscripts are typically 1000-2500 words.

SUBMISSION DEADLINE
Fall 2012 – September 10, 2012