



Center for  
**Educator Compensation Reform**  
**OBSERVATIONS OF TEACHERS'**  
**CLASSROOM PERFORMANCE**

Anthony T. Milanowski, University of Wisconsin-Madison

Cynthia D. Prince, Vanderbilt University

Julia Koppich, J. Koppich and Associates

The Center for Educator Compensation Reform (CECR) was awarded to Westat—in partnership with Learning Point Associates, Synergy Enterprises Inc., Vanderbilt University, and the University of Wisconsin—by the U.S. Department of Education in October 2006.

The primary purpose of CECR is to support the Teacher Incentive Fund (TIF) grantees with their implementation efforts through the provision of ongoing technical assistance and the development and dissemination of timely resources. CECR also is charged with raising national awareness of alternative and effective strategies for educator compensation through a newsletter, a Web-based clearinghouse, and other outreach activities.

This work was originally produced in whole or in part by CECR with funds from the U.S. Department of Education under contract number ED-06-CO-0110. The content does not necessarily reflect the position or policy of CECR or the Department of Education, nor does mention or visual representation of trade names, commercial products, or organizations imply endorsement by CECR or the federal government.

This report is in the public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, the suggested citation is: Milanowski, A., Prince, C., & Koppich, J., *Observations of Teachers' Classroom Performance*. Center for Educator Compensation Reform. U.S. Department of Education, Office of Elementary and Secondary Education, Washington, D.C., 2007.

# OBSERVATIONS OF TEACHERS' CLASSROOM PERFORMANCE

*Many organizations in both the private and public sectors use observations or evaluations of employee performance along with (or instead of) measures of performance based on outcomes. Although not as easily quantifiable as test scores, these performance evaluations reflect the judgment of an evaluator or set of evaluators against a set of standards. They often include employee behaviors and attitudes, as well as outcomes, as the evidence base for making a rating of performance. They also frequently use rating scales that attempt to capture the range of performance on a set of pre-defined performance dimensions.*

## Advantages of using classroom observation as another measure of teacher performance

The strengths of this type of measurement are that:

1. It is applicable to jobs where performance measures based on outcomes are hard to develop or where outcomes cannot be assigned to an individual person;
2. It ensures that important aspects of performance that go beyond measured outcomes, such as how the outcomes are achieved, are taken into account;
3. It focuses on aspects of performance most likely to be in employees' control – their own behavior – which helps teachers understand the connection between their performance and their pay;
4. It gives employees credit for their efforts when circumstances outside their control prevent achieving success, as defined by student test scores or other outcome measures; and
5. It can provide formative feedback to employees on what they can do to achieve important outcomes (e.g., behaviors, task strategies).

Because of these strengths, most organizations, including educational organizations, will want to use both outcome-based performance measures and classroom observations in their efforts to improve educator performance and hold educators accountable. Most organizations already have a formal performance evaluation system, and therefore this module does not cover the basics of setting up these systems – books by Danielson and McGreal and by Stronge and Tucker provide basic information on system design.<sup>1</sup> Rather, this article is a guide for using observations of educator performance as a basis for educator compensation.

## Typical state and district measures of the effectiveness of teachers' classroom performance: Advantages and disadvantages of these approaches

Educational organizations are likely to take two approaches toward non-test score evaluations for educator pay systems:

1. Build on the existing performance evaluation system; or
2. Develop a special-purpose measurement process.

Each approach has its advantages and disadvantages. If they use existing systems, districts avoid additional measurement overhead and keep the focus on one set of performance measures. Educators and administrators are already familiar with these systems and using them also avoids the perception of adding a lot of extra work. One problem, however, is that performance evaluation for pay and for educator development may have some conflicting requirements. The former makes reliability and validity paramount, but the latter prioritizes feedback and assistance.<sup>i</sup> It may be hard to find the time and resources to do both. Another problem is that many performance evaluation systems are not designed to do much more than weed out the poorest performers, and many school districts do not implement the systems in a way that reliably and validly differentiates between different levels of performance. Since performance assessment for compensation will likely subject the assessment system to close scrutiny, program designers need to examine the quality of their assessment systems critically. Is the performance measurement good enough to help district officials determine pay? The material presented below will help you decide how to answer that question.

If examination of the current educator performance evaluation system suggests deficiencies, program designers will need to decide whether the existing system should be modified or whether they will need to develop a completely new system for pay purposes. The advantage of a new, separate system is that it can be designed to be rigorous and reliable and to focus on measuring the most important aspects of performance. The National Board for Professional Teaching Standards uses a special purpose performance measurement system that involves a teacher portfolio, videos of teaching practice, and a series of written exercises. It concentrates on assessing just the competencies that the National Board has identified as important and does so with good reliability and efficiency. In the private sector, companies often use special purpose performance measurement for employee selection, in the form of assessment centers, where employees must complete a variety of exercises to demonstrate specific competencies. Again, the advantage is focus and efficiency of measurement. Potential disadvantages, of course, are the costs required to develop a completely new system for pay purposes and to train evaluators to use it to assess teachers' classroom performance.

<sup>i</sup> Reliability refers to the consistency of performance ratings over time or agreement among different evaluators, while validity refers to the degree to which a rating actually measures what it claims to measure (e.g., effective teaching practices).

## Basic requirements for a system that links observations of teachers' classroom performance to pay

Whichever direction the district takes, the basic requirements for a system that links non-test score evaluations to pay are the same. That is, the system:

1. Measures the right things;
2. Produces valid and reliable measurements;
3. Provides tools to help educators improve performance in response to the measurement; and
4. Is accepted by those whose performance is being measured and by those doing the measuring.

### Is the system measuring the right things?

This question is important because, if the assumptions of pay-for-performance are correct, incentives will direct educators' efforts toward the measured and rewarded performance. One challenge in education is that because there are so many conceptions of good performance, it is often hard to decide which should be emphasized. Additionally, using performance evaluation for tenure decisions and to hold tenured teachers accountable tends to encourage systems to aspire to comprehensive coverage of all job responsibilities. For example, systems based on Charlotte Danielson's Framework for Teaching define about 68 performance dimensions within four broad domains of planning and preparation, creating an environment for learning, teaching for learning, and professionalism.<sup>2</sup> But in order to maintain focus and reduce measurement overhead, it may not be desirable for the system to measure and reward all possible aspects of performance. Instead, the system should measure and reward only the aspects of performance that are to be the key drivers of important outcomes such as student learning. The performances that the system measures and rewards should directly reflect what educators need to do to carry out the organization's strategies for achieving its goals.

The system also needs to measure those aspects of performance that distinguish outstanding performers from those who are just acceptable. Many performance evaluation systems, by design or in practice, serve simply to distinguish unsatisfactory from minimally acceptable behavior. This type of system will not likely be useful in motivating substantially improved performance because outstanding performance is not defined and distinguished for reward.

If the answer to any of the following key questions is no, a redesign will be needed. The key step is to identify what aspects of educator performance really drive organizational performance. One approach would be for the

*The performance evaluation system should measure and reward only the aspects of performance that are to be the key drivers of important outcomes, such as student learning.*

**Key questions for program designers:****Does the system measure the correct things?**

1. Are the dimensions of performance that the system measures the drivers of important outcomes such as student learning, attendance, or graduation rates?
2. Are important drivers missing?
3. Does the system have so many dimensions that the key drivers get lost?
4. Does the system include what truly distinguishes an outstanding performer from an average performer?
5. Does the system provide a clear distinction between satisfactory and outstanding performance?

school system to review the programs or strategies that the organization is deploying to meet its key goals and then make inferences about what teachers and principals have to do well in order to carry out the programs or strategies. Another approach is to go back to the research on teaching and learning to identify the most important contributors to student achievement and then decide which ones are most in need of improvement. These will become the dimensions of performance that the system will assess. For example, differentiation of instruction may be a key part of a district's strategy to close achievement gaps and is identified in research as a likely contributor to student learning.<sup>3</sup>

**Do the system's procedures support valid and reliable measurement?**

Reliability and validity are important not only because linking pay to performance may motivate increased scrutiny of the measurement process, but also because they affect the motivational impact of rewards linked to performance. If educators believe that favoritism or measurement error determines how well they do on the assessment, rather than their true performance level, they will be less likely to expend effort to perform as desired. If an evaluation system cannot validly measure, rewards will be less contingent on performance, and when educators realize this, they will be less motivated to perform.

### Key questions for program designers:

#### Does the system support valid and reliable measurement?

1. Are the system's procedures uniform or standardized with respect to the educator groups with which they are used?
2. Are types and sources of evidence, and methods of gathering evidence, specified so that educators and evaluators know what to look for?
3. Are evaluators trained to apply the system consistently?
4. Is there any evidence that evaluators can — or do — apply the system consistently?
5. Is there any evidence that evaluators' judgments are related to other measures of educator performance?

Program designers need to build reliability and validity into performance assessment systems. The recommendations that follow discuss some basic design features that program designers should consider. Many of these recommendations also address the concerns that teachers often have about principals being biased or not knowledgeable enough to do fair evaluations. While these concerns may not be born out by research (which tends to suggest that principals are lenient in their evaluations and that few teachers are ever rated as unsatisfactory), it is important for teacher acceptance that the evaluation system's processes and procedures minimize bias and are perceived as fair.

## Recommendation 1

### Use relatively detailed rating scales (“rubrics”) that define a set of levels for each performance dimension.

Rating scales, or rubrics, provide guidance to evaluators in making decisions about performance. While they cannot completely define each performance level (to do so often requires too many words to be practical), rubrics can provide the structure needed to develop consistency among evaluators and reduce the impact of idiosyncratic evaluator beliefs and attitudes on evaluation results. District officials should share these rating scales with those being evaluated so that educators know what the

performance expectations are, rather than wondering what the evaluator thinks is good performance. The descriptions or examples of performance in the rating scales need to be good exemplars of the performance dimension the scale is attempting to capture. Developers should write the rating scale descriptions and examples clearly, minimize the use of vague quantifiers like “consistently” or “frequently,” and make clear distinctions between performance levels. Generally, educator performance rating scales define between three and five levels of performance. It is quite difficult to develop rating scales that divide the performance range into more than five levels, and fewer than three do not allow the definition of a truly high or advanced level of performance. The use of four levels seems common in practice.

## Recommendation 2

### **Specify what counts as evidence for performance, and how it is to be collected, in a performance measurement handbook or manual.**

Specifying the evidence up front helps to structure the evaluators' decision process, discourage consideration of irrelevant factors, and reassure those being evaluated that they are being measured on observable evidence rather than on evaluators' biases. Program developers should also consider and specify the amount of evidence or the timing of collection. Because of the complexity and variability of most educators' jobs, observation of a single lesson or a single staff meeting is not likely to be a reliable basis to make a judgment. Instead, evaluators should conduct multiple observations and collect evidence at multiple points in time. Evaluators should also supplement their observations with other kinds of “artifacts,” such as unit and lesson plans, tests and assignments, examples of student work, parent contact logs, or classroom procedures. Danielson discusses the use of such artifacts in teacher evaluation.<sup>4</sup> Districts can substitute videos for live observations, but should establish procedures for videotaping to promote consistency and ensure that evaluators collect artifacts to round out the evidence.

## Recommendation 3

### **Use an analytic assessment process that separates observation, interpretation, and judgment.**

Often, evaluators begin the evaluation process with preconceived notions of who is and who is not a good performer and with their own idiosyncratic “gut feelings” about what good performance looks like. Evaluators also tend to form judgments about whether the educator being evaluated is a good performer or a poor one based on evidence collected early in the evaluation



process, then notice only confirming evidence and disregard contrary evidence. They may also fail to distinguish differences in the level of performance on different dimensions because they are concentrating on what they believe is most important or because they are overloaded by trying to observe and evaluate at the same time. These tendencies can produce inaccurate ratings that may also lack credibility to those being evaluated.

The use of an assessment process that separates observation, interpretation, and judgment can reduce these tendencies. Districts should start by defining what counts as evidence of performance and train evaluators to collect it – both positive and negative. Evaluators should then collect the evidence by observation or review of materials (e.g., lesson plans, student work), but withhold judgment about the level of performance represented. When they have finished collecting relevant evidence (for example, after an observation session has been completed), the evaluators should review and interpret the evidence collected, decide which evidence is relevant to the pre-defined dimensions of performance, and compare the relevant evidence to the rating scale or rubric. Only then should the evaluator make a judgment, one dimension at a time, about which rating scale level is supported by the preponderance of the relevant evidence. Evaluators should stick to the recorded evidence and compare each educator's performance to the rating scale, not to other educators. These steps maximize evaluator accuracy and are used in high-quality assessment systems such as those that have been developed for National Board for Professional Teaching Standards certification.

*The use of an assessment process that separates observation, interpretation, and judgment can reduce an evaluator's tendency to follow his or her "gut feelings" about what good performance looks like.*

## Recommendation 4

### Use multiple evaluators.

Research suggests that many principals have a difficult time evaluating teachers, for reasons ranging from lack of knowledge of the subject being taught to disinclination to upset working relationships.<sup>5</sup> Thus, principal evaluations are frequently lenient, and most teachers end up with "satisfactory" ratings or higher. A 2007 audit of Wake County, North Carolina, teacher evaluations, for example, found that only one of 363 teachers whose evaluations were examined had received an unsatisfactory rating from his or her principal.<sup>6</sup> Another recent study of teacher evaluations conducted in Chicago between 2003 and 2006 found that the majority of veteran principals in the district admitted to inflating performance ratings for some of their teachers.<sup>7</sup> Over the four-year period, 93 percent of Chicago teachers earned the two highest ratings ("superior" or "excellent"), and only 3 in 1,000 received "unsatisfactory" ratings. Even in 87 schools that had been identified as failing, 79 percent did not award a single unsatisfactory rating to teachers between 2003 and 2005.

Using multiple evaluators can help diffuse the “blame” for a less-than-glowing evaluation and reduce tendencies to be lenient. It can also reassure teachers that their evaluations are not based on the principal’s biases or subjective opinions. Some schools and school systems use peer evaluators or evaluators from outside the building (e.g., a central office program specialist) for this purpose. Examples include Columbus, Toledo, Cincinnati, Rochester, San Francisco, Minneapolis, Montgomery County Public Schools in Maryland, Poway Unified School District in San Diego County, and most recently, some low-performing schools in Chicago.<sup>8</sup>

Examples of teacher observation instruments developed in two of these districts are included in the exhibits shown left; links to these forms can be found in exhibits 1–3, pages 14–16. In

Montgomery County, Maryland, evaluators complete a narrative description of teacher performance based on seven performance standards (e.g., teachers know the subjects they teach and how to teach those subjects to students).<sup>9</sup> The narrative description includes multiple sources of evidence, such as “classroom observations, analysis and review of student results as described in the shared accountability system, contributions to overall school mission and environment, review of student and parent surveys, review of professional growth plans and implementation results, and any other documents collected by the evaluator and/or the teacher during the full length of the cycle.” As another example, teacher performance in Poway, California, is assessed by “reflection, observation, documentation, and conference,” and peer evaluators are a key source of evidence. Evaluators assess teachers’ performance on 18 dimensions within five broad domains of professional responsibility, such as using appropriate learning materials, effectively managing instructional time, and monitoring student performance.<sup>10</sup>

Districts might also consider the use of “360 degree” or multi-source evaluation, in which supervisors, peers, and parents, as well as the teacher, contribute evidence and/or assessment judgments.<sup>11</sup> A special purpose assessment system, separate from routine performance evaluation, could use a panel of assessors drawn from both inside and outside the organization. This approach would minimize tendencies toward lenience and maximize the ability to recognize technically good performance since districts would choose these assessors for their subject expertise and their probable lack of contact with those being assessed.

The image displays two overlapping teacher observation forms. The top form is the 'PROFESSIONAL GROWTH SYSTEM FINAL EVALUATION REPORT' from Montgomery County Public Schools, Rockville, Maryland. It includes fields for teacher and principal information, a narrative description box, and a 'Final Rating' section. The bottom form is the 'CERTIFICATED CLASSROOM OBSERVATION' from Poway Unified School District. It features a header with teacher and evaluator details, a 'PART I: DOMAINS OF PROFESSIONAL RESPONSIBILITIES' section with five domains (Planning and Designing Instruction, Instruction, Classroom Management, Assessment, and Professional Responsibilities), and a 'PART II: SUMMARY OF GOALS CONFERENCE' section at the bottom.

## Recommendation 5

### Train evaluators for consistency.

An effective training program for performance evaluators should follow 10 steps:

1. Start with an overview of the process — give evaluators the big picture.
2. Familiarize evaluators with the performance dimensions and rating scales.
3. Help evaluators recognize their own biases (e.g., for certain teaching styles).
4. Describe common judgment errors to be aware of and avoided.
5. Describe how to recognize and record evidence of performance.
6. Practice collecting evidence and observing.
7. Practice connecting evidence to performance dimensions.
8. Practice interpreting evidence relative to the rating scale.
9. Discuss rating conventions (e.g., What do you do if you do not have any evidence for a performance dimension? How do you interpret words like “consistently” or “frequently” when used in the rating scale?).
10. Practice rating samples of performance (e.g., videos, written scenarios, or cases).

Consider following the training with a “certification exercise” in which evaluators must match the ratings of an expert jury in order to “graduate.” Set a threshold, such as 75 percent absolute agreement with the “experts.” Districts could then provide additional training to those who do not meet this standard.

## Recommendation 6

### Monitor evaluators' performance, and hold evaluators accountable for doing a good job.

One way to monitor evaluator quality is to review a sample of the written evaluations that they complete and the supporting evidence that they use to make their judgments. Another way to monitor evaluators' performance is to assign two evaluators to work independently on the same sample of evaluations and then compare the quality of their work. Districts should articulate clear standards for following the assessment process and evaluate the evaluators themselves on their adherence to them.

It may be useful to conduct analyses of inter-rater agreement (reliability) and validity to help assure all stakeholders that evaluation scores are sound measures of teacher performance. One way to do this is to ask two evaluators to rate performance independently and then compare the results for agreement. Where multiple evaluators are used, this kind of analysis can also reassure teachers that the evaluation is based on observable evidence about which different evaluators can agree. Validity evidence could include the degree to which the performance ratings match measures such as value-added student achievement scores or even consistency with parent or student satisfaction surveys. Your program evaluator or staff at the Center for Educator Compensation Reform can help you design these studies.

#### Key principles of an alternative compensation plan

1. Awards are based on objective criteria.
2. A clear, transparent system is created.
3. Rewards are attainable.
4. Implementation is feasible.
5. Plan is affordable for the district.
6. System is sustainable in the long term.

### The evaluation system provides tools that the educators being evaluated can use to improve their performance

It is critical that the evaluation system provide tools and feedback to the educators being evaluated that can be used to improve their performance. Because a primary purpose of incentives is to motivate people to improve performance, teachers must be able to use the results from evaluation as a stepping stone. In many cases, improving performance is a matter of skill and focus, not just will and effort. Behavioral science research suggests that performers need to know what the specific performance expectations are, how well they are doing, and how to get better.<sup>12</sup> Research on teacher evaluation also suggests that receiving high-quality feedback and coaching improves the acceptance of the evaluation process by those being evaluated.<sup>13</sup>

**Key questions for program designers:****Does the system provide tools to the educators that are evaluated to improve their performance?**

1. Does the system make clear, in specific terms, what performances or behaviors are needed to be rated a good performer and receive the incentive?
2. Does the system provide educators with specific feedback that tells them why they were rated as they were?
3. Are evaluators knowledgeable enough about good performance to give quality feedback?
4. Is some form of coaching or assistance available to help those who are not doing well?
5. Are professional development opportunities available that directly address performance deficiencies, so that those who want to improve can get help?

*Feedback of results alone, without suggestions for how to improve, does not lead to improved performance.*

To help maximize motivation and acceptance, those who do not measure up need to perceive a clear path for improvement. Performance feedback is the first step. Current thinking on performance feedback suggests that feedback of results alone, without suggestions for how to improve, does not lead to improved performance.<sup>14</sup> Support and coaching are needed to turn the information provided by feedback – especially negative feedback – into improved performance, rather than rejection, resentment, or lower self-efficacy. Principals and mentor teachers should provide specific, concrete feedback that tells an educator not only his or her rating, but also exactly what prevented him or her from getting a higher score and what specific behaviors or results would raise the score. This could be followed by information about relevant training, suggestions about techniques to try, whom to observe to see good performance exemplified, and even modeling aspects of desired performance. This, in turn, requires that districts train evaluators to provide feedback, that there is a coach or mentor to go to for help, and that training and development programs are available to provide the skills needed to do well on the next performance assessment.

## Helping those whose performance is being measured and those doing the measuring accept the system

Acceptance of the system is a key to its long-term survival. Research shows that understanding how the system works, perceptions of procedural fairness, and workload are all influences on acceptance by both those administering the system and those covered by it.<sup>15</sup>

### Key questions for program designers:

#### Have we developed a system that will be accepted and perceived as fair and workable?

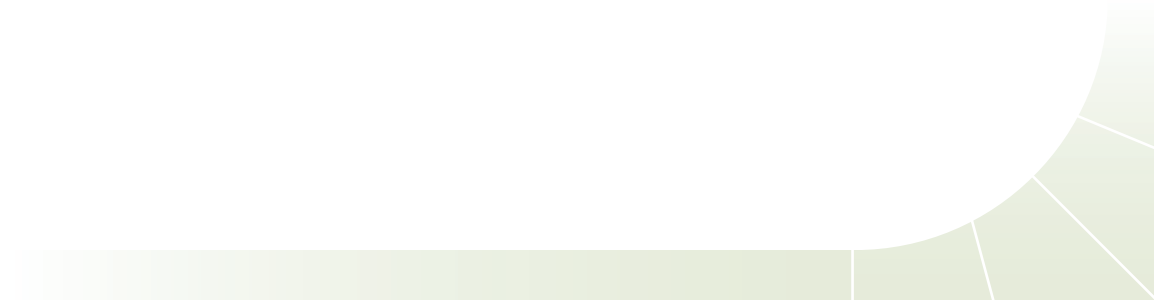
1. Are the system procedures fully worked out?
2. Does the system have streamlined procedures to minimize additional workload?
3. Are those procedures clearly communicated to those involved?
4. Has a communications plan been put in place to make sure that those involved understand how the system works?
5. Have the performance expectations been clearly outlined beforehand to those being evaluated?
6. Are the procedures designed to make sure that educators and evaluators have the time to do a good job?
7. Are safeguards in place to assure confidentiality where appropriate?
8. Is there an appeals process in place?

One way to maximize the possibility of acceptance and get a feeling for whether the system will be perceived as fair and workable is to involve those who will be evaluated in the design or re-design of the assessment process. Many consultants working with private sector firms routinely advise the companies to seek employee input and involvement in program design. In a collective bargaining state, involving teachers in the design of the assessment process is not simply a good idea; it is a statutory requirement and therefore must be bargained.

Whether you are considering using your current performance evaluation system as a basis for performance pay or designing a new system, you should answer most of the questions above in the affirmative. This will optimize the chance for long-term success. But the questions also illustrate the complexity of using classroom observations to assess teacher performance, and so organizations need to ensure that they have the resources and the will to make a commitment to operate a quality system.

One consideration that may make the effort worthwhile is that the performance dimensions underlying the performance assessment system can be used in other parts of the human resource management system. For example, providing teacher candidates with information about the performance requirements allows those who do not think they can meet them to screen themselves out of the hiring process. Another example is teacher selection. If interview questions or other selection measures are based on the same performance dimensions, it is more likely that new hires will have the competence to become good performers and will not need to be dismissed at the end of the probationary period or require remedial training. Aligning the human resource management system (recruitment, selection, induction, mentoring, professional development, performance management, compensation, and school leadership) to the performance assessment system reinforces the latter's importance and helps create a shared conception of good performance.<sup>16</sup>

Finally, we recommend that districts conduct some sort of pilot test or dry run when implementing a new performance measurement system to identify and correct problems that will inevitably occur. Previous research on standards-based teacher evaluation systems found that implementation problems reduced teacher acceptance of the new systems and that at least one pilot year was needed to work out the glitches.<sup>17</sup>



**Exhibit 1**

Montgomery County Public Schools, Rockville, Maryland. "Professional Growth System Final Evaluation Report: Teacher." Available at: <http://www.montgomeryschoolsmd.org/departments/forms/pdf/425-39.pdf>

<b>Office of Human Resources</b> <b>MONTGOMERY COUNTY PUBLIC SCHOOLS</b> <b>Rockville, Maryland 20855</b>	<b>PROFESSIONAL GROWTH SYSTEM</b> <b>FINAL EVALUATION REPORT:</b> <b>Teacher</b>
<b>INSTRUCTIONS:</b> Evaluators complete a narrative description based on the following performance standards. The description includes classroom observations, analysis and review of student results as described in the shared accountability system, contributions to overall school mission and environment, review of student and parent surveys, and review of professional growth plans and implementation results, and any other documents collected by the evaluator and/or the teacher during the full length of the cycle.	
Teacher: _____ Employee Number: _____ Years of MCPS Experience: _____ Principal: _____ Type: _____ First-year Probationary _____ Tenured (3-year cycle) _____ with CT _____ without CT _____ Tenured (4-year cycle) _____ Second-year Probationary _____ Tenured (5-year cycle) _____ Third-year Probationary _____ Special Evaluation School: _____ Subject or Grade Level: _____	
<b>Performance Standards:</b> I. Teachers are committed to students and their learning II. Teachers know the subjects they teach and how to teach those subjects to students III. Teachers are responsible for establishing and managing student learning in a positive learning environment IV. Teachers continually assess student progress, analyze the results, and adapt instruction to improve student achievement V. Teachers are committed to continuous improvement and professional development VI. Teachers exhibit a high degree of professionalism	
<b>Dates of Observations:</b> _____ <b>Dates of Conferences:</b> _____	
<b>Final Rating:</b> ( ) Meets Standard                      ( ) Below Standard	
Evaluator's Signature _____ Date _____ Principal's Signature _____ Date _____ Teacher's Signature _____ Date _____ <i>(Teacher's signature indicates that the teacher has read and reviewed the final evaluation summary, not necessarily that the teacher concurs with the contents. Teachers may choose to attach comments.)</i>	
MCPS Form 425-39 Rev. 9/02                      DISTRIBUTION: COPY 1/Employee; COPY 2/Principal; COPY 3/Office of Human Resources This form has been reprinted with permission from Montgomery County Public Schools	



## Exhibit 2

Poway Unified School District, Poway, California. "Certificated Classroom Observation."

Available at: <http://www.powayschools.com/projects/literacy/SSTTL/PPAP/PDFs/CertClassObserv.pdf>

<b>CERTIFICATED CLASSROOM OBSERVATION</b>							
POWAY UNIFIED SCHOOL DISTRICT 13626 Twin Peaks Road, Poway, California 92064							
TEACHER'S NAME			COURSE/SUBJECT/GRADE LEVEL			SCHOOL/DEPARTMENT	DATE
						MCHS / Science	
TIME:	(FROM)	(TO)	CLASS SIZE	<input checked="" type="checkbox"/> ANNOUNCED	TENURED	<input type="checkbox"/>	PAGE OF
				<input type="checkbox"/> UNANNOUNCED	STATUS: PROBATIONARY	<input checked="" type="checkbox"/> 1st Yr <input type="checkbox"/> 2nd Yr	
<b>EVALUATOR'S OBSERVATION</b>							
1. Instructional Activities:							
2. Reactions/Suggestions:							
EVALUATOR'S SIGNATURE				POSITION		DATE	
				Principal			
<b>TEACHER'S COMMENTS</b>							
All written summaries of observations shall be delivered to the teacher within three (3) working days following the observations, and signed by the teacher within five (5) working days following the observation. The teacher has the right to respond to the observation in writing, and the response shall be attached to the formalized observation.							
Signature of teacher does not constitute agreement, but merely acknowledgment of the observation and this record. A follow-up conference, if requested, will provide opportunity for discussion and comments on the observation.				TEACHER'S SIGNATURE		DATE	
PUSD P-65A (Rev. 1-97) Order from Inventory Control				Distribution: ORIGINAL-Principal COPY-Teacher			

**Exhibit 3**

Poway Unified School District, Poway, California. "Certificated Evaluation."

Available at: <http://www.powayusd.com/projects/literacy/SSTL/PPAP/PDFs/CertEval.pdf>

POWAY UNIFIED SCHOOL DISTRICT <b>CERTIFICATED EVALUATION - Page 1</b>			
Teacher's Name		Course/Subject/Grade Level	School
Social Security Number		TEMPORARY <input type="checkbox"/>	PROBATIONARY <input type="checkbox"/> 1 <sup>ST</sup> Year <input type="checkbox"/> 2 <sup>ND</sup> Year
		TENURED <input type="checkbox"/>	Date
DOES NOT MEET STANDARDS	MEETS STANDARDS	<b>PART I: DOMAINS OF PROFESSIONAL RESPONSIBILITIES</b> (Assessed by reflection, observation, documentation, conference)	
		<b>DOMAIN I: PLANNING AND DESIGNING INSTRUCTION</b>	
<input type="checkbox"/>	<input type="checkbox"/>	Designs Long-Range Plans to Accomplish Yearly Learning Goals	
<input type="checkbox"/>	<input type="checkbox"/>	Selects Appropriate Lesson Objectives	
<input type="checkbox"/>	<input type="checkbox"/>	Designs Lessons that include Elements Essential for Learning	
		<b>DOMAIN II: INSTRUCTION</b>	
<input type="checkbox"/>	<input type="checkbox"/>	Delivers Effective Instruction	
<input type="checkbox"/>	<input type="checkbox"/>	Utilizes Appropriate Learning Materials	
<input type="checkbox"/>	<input type="checkbox"/>	Utilizes a Variety of Instructional Strategies to Meet the Needs of the Students	
<input type="checkbox"/>	<input type="checkbox"/>	Creates Learning Opportunities for all Students	
<input type="checkbox"/>	<input type="checkbox"/>	Demonstrates Subject Matter Competence	
		<b>DOMAIN III: CLASSROOM MANAGEMENT</b>	
<input type="checkbox"/>	<input type="checkbox"/>	Provides a Positive Learning Environment which Promotes Appropriate Student Behavior	
<input type="checkbox"/>	<input type="checkbox"/>	Effectively manages Instructional Time	
<input type="checkbox"/>	<input type="checkbox"/>	Organizes Physical space	
		<b>DOMAIN IV: ASSESSMENT</b>	
<input type="checkbox"/>	<input type="checkbox"/>	Establishes Clear Academic Standards	
<input type="checkbox"/>	<input type="checkbox"/>	Monitors Student Learning	
<input type="checkbox"/>	<input type="checkbox"/>	Adjusts Teaching and Learning Based on Assessment	
<input type="checkbox"/>	<input type="checkbox"/>	Uses Assessment Results to Give Students and/or Parents Timely, Accurate and Constructive Feedback	
		<b>DOMAIN V: PROFESSIONAL RESPONSIBILITIES</b>	
<input type="checkbox"/>	<input type="checkbox"/>	Grows and Develops Professionally	
<input type="checkbox"/>	<input type="checkbox"/>	Shares in the Responsibility for the Smooth Operation of the School	
<input type="checkbox"/>	<input type="checkbox"/>	Complies with Established Rules, Regulations, Policies and Laws	
<b>Part II: SUMMARY OF GOALS CONFERENCE</b>			
Date of Meeting _____		<input type="checkbox"/> Goals Attached	<input type="checkbox"/> Goals Not Attached
ADMINISTRATOR'S INITIALS	TEACHER'S INITIALS	SUMMARY ON PAGE 2 SHALL INCLUDE WRITTEN RECOMMENDATIONS FOR ALL ITEMS CHECKED UNSATISFACTORY.	

## End Notes

- <sup>1</sup> Stronge, J.H., & Tucker, P.D. (2003). *Handbook on teacher evaluation: Assessing and improving performance*. Larchmont, NY: Eye on Education, Inc.; Danielson, C., & McGreal, T.L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- <sup>2</sup> Danielson, C. (2007). *Enhancing professional practice: A framework for teaching (2nd ed.)*. Alexandria, VA: Association for Supervision and Curriculum Development.
- <sup>3</sup> Center for Comprehensive School Reform and Improvement. (2007). *A teachers' guide to differentiating instruction*. Washington, DC: Author; Kline, L. (1995). A baker's dozen: Effective instructional strategies. In R. Cole (Ed.), *Educating everybody's children: Diverse teaching strategies for diverse learners*, 21-43. Alexandria, VA: Association for Supervision and Curriculum Development.
- <sup>4</sup> Danielson, C. (2007). *Enhancing professional practice: A framework for teaching (2nd ed.)*. Alexandria, VA: Association for Supervision and Curriculum Development.
- <sup>5</sup> Halverson, R., Kelley, C., & Kimball, S. (2004). Implementing teacher evaluation systems: How principals make sense of complex artifacts to shape local instructional practice. In W. Hoy & C. Miskel (Eds.), *Research and theory in educational administration*. (Vol. 3). Greenwich, CT: Information Age Publishing; Nelson, B.S., & Sassi, A. (2000). Shifting approaches to supervision: The case of mathematics supervision. *Educational Administration Quarterly*, 36(4), 553-584; Peterson, K.D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practice. (2nd ed.)*. Thousand Oaks, CA: Corwin Press.
- <sup>6</sup> Keung Hui, T. "Easy A's for Wake teachers?" *The (Raleigh, NC) News & Observer*, September 6, 2007. <http://www.newsobserver.com/news/education/wake/story/693999.html>
- <sup>7</sup> The New Teacher Project. (2007, July). *Teacher hiring, transfer and assignment in Chicago Public Schools*. New York: Author. <http://www.tntp.org/ourresearch/documents/TNTPAnalysis-Chicago.pdf>; Rossi, R. "Teachers' grades inflated," *Chicago Sun-Times*, July 30, 2007. <http://www.suntimes.com/>
- <sup>8</sup> Escamilla, P., Clarke, T., & Linn, D. (2000, January). *Exploring teacher peer review*. Washington, DC: National Governors' Association. <http://www.nga.org/cda/files/000125PEERREVIEW.pdf>; Keller, B. "Teachers to conduct peer reviews in Chicago," *Education Week*, January 25, 2006. <http://www.edweek.org/ew/articles/2006/01/25/20chicago.h25.html>
- <sup>9</sup> Montgomery County Public Schools, Rockville, Maryland. "Professional growth system final evaluation report: Teacher." <http://www.montgomeryschoolsmd.org/departments/forms/pdf/425-39.pdf>
- <sup>10</sup> Poway Unified School District, Poway, California. "Certificated classroom observation." <http://www.powayschools.com/projects/literacy/SSTTL/PPAP/PDFs/CertClassObserv.pdf>; Poway Unified School District, Poway, California. "Certificated evaluation." <http://www.powayusd.com/projects/literacy/SSTTL/PPAP/PDFs/CertEval.pdf>

- <sup>11</sup> Bracken, D.W., Timmreck, C.W., & Church, A.H. (Eds). (2001). *The handbook of multisource feedback*. San Francisco: Jossey-Bass; Dyer, K.M. (2001, February). The power of 360-degree feedback. *Educational Leadership*, 58(5), 35-38; Wilkerson, D.J., Manatt, R.P., Rogers, M.A., & Maughan, R. (2000, June). Validation of student, principal, and self-ratings in 360 degree feedback for teacher evaluation. *Journal of Personnel Evaluation in Education*, 14(2), 179-192.
- <sup>12</sup> Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284; Locke, E.A., & Latham, G.P. (1994). Goal setting theory. In H.F. O'Neil, Jr., & M. Drillings (Eds.), *Motivation: Theory and research*, 13-29. Hillsdale, NJ: Lawrence Erlbaum Associates; London, M. (1997). *Job feedback: Giving, seeking, and using feedback for performance improvement*. Mahwah, NJ: Lawrence Erlbaum Associates.
- <sup>13</sup> Heneman, H.G., III, & Milanowski, A.T. (2003). Continuing assessment of teacher reactions to a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, 17(3), 171-195; Heneman, H.G., III, Milanowski, A., Kimball, S.M., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay*. CPRE Policy Brief RB-45. Philadelphia: Consortium for Policy Research in Education.
- <sup>14</sup> London, M., & Smither, J.W. (2002, Spring). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review*, 12(1), 81-100; Ilgen, D.R., & Davis, C.A. (2000, July). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology*, 49(3), 550-565; Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- <sup>15</sup> Heneman, H.G., III, & Milanowski, A.T. (2003). Continuing assessment of teacher reactions to a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, 17(3), 171-195; Heneman, H.G., III, Milanowski, A., Kimball, S.M., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay*. CPRE Policy Brief RB-45. Philadelphia: Consortium for Policy Research In Education.
- <sup>16</sup> For more information on the alignment of the major human resource program areas around a model of performance competencies, see Heneman, H.G., III, & Milanowski, A.T. (2004). Alignment of human resource practices and teacher performance competency. *Peabody Journal of Education*, 79(4), 108-125.
- <sup>17</sup> Heneman, H., Milanowski, A., & Kimball, S. (2007, February). *Teacher performance pay: Synthesis of plans, research and guidelines for practice*. CPRE Policy Brief RB-46. Philadelphia: University of Pennsylvania, Graduate School of Education, Consortium for Policy Research in Education.