Appendix C. Summary of Measures

Measure	Description	Research	Strengths	Cautions
Value-Added Models	Statistical models used to determine teachers' contributions to students' test score gains. May also be used as a research tool (e.g., determining the distribution of "effective" teachers by student or school characteristics).	Little is known about the validity of value-added scores for identifying effective teaching, though research using value-added models suggests that teachers differ markedly in their contributions to students' test score gains. However, correlating value-added scores with teacher qualifications, characteristics, or practices has yielded mixed results and few significant findings. Teachers vary in effectiveness, but research has not determined why.	• Provides a way to evaluate teachers on their contribution to student learning, which most measures do not. • Requires no classroom visits because linked student/teacher data can be analyzed at a distance. • Entails little burden at the classroom or school level because most data are already collected for NCLB purposes. • May be useful for identifying outstanding teachers whose classrooms can serve as "learning labs" as well as struggling teachers in need of support.	• Models are not able to sort out *teacher* effects from *classroom* effects. • Vertical test alignment is assumed (i.e., tests are measuring essentially the same thing from grade to grade). • Value-added scores are not useful for formative purposes because teachers learn nothing about how their practices contributed to (or impeded) student learning. • Value-added measures are controversial because they measure *only* teachers' contributions to student achievement gains on standardized tests.
Classroom Observation	Classroom observations are used to measure observable classroom processes, including specific teacher practices, holistic aspects of instruction, and interactions between teachers and students. They can measure broad, overarching aspects of teaching, or subject-specific or context-specific aspects of practice.	Some highly researched protocols have been linked to student achievement, though associations are sometimes modest. Research and validity findings are highly dependent on the instrument used, sampling procedures, and the training of raters. There is a lack of research on observation protocols as used in context for teacher evaluation.	• Provides rich information about classroom behaviors and activities. • Is credible—generally considered a fair and direct measure by stakeholders. • Depending on the protocol, can be used in various subjects, grades, and contexts. • Can provide information useful for both formative and summative purposes.	• Choosing or creating a valid and reliable protocol and training and calibrating raters are essential to obtaining valid results. • Expensive due to cost of observers' time; intensive training and calibrating of observers adds to expense but is necessary for validity. • Assesses observable classroom behaviors, but not as useful for assessing beliefs, feelings, intentions, or out-of-classroom activities.

From: A Practical Guide to Evaluating Teacher Effectiveness, April 2009
National Comprehensive Center for Teacher Quality.

Measure	Description	Research	Strengths	Cautions
Principal Evaluation	Are generally based on classroom observation, may be structured or unstructured; procedures and uses vary widely by district. Generally used for summative purposes, most commonly for tenure or dismissal decisions for beginning teachers.	Studies comparing subjective principal ratings to student achievement find mixed results. Little evidence exists on validity of evaluations as they occur in schools, but evidence indicates that training for principals is limited and rare, which would impair validity of their evaluations.	• Represents a useful perspective based on principals' knowledge of their school and context. • Is generally feasible and can be one useful component in a system used to make summative judgments and provide formative feedback.	• Evaluation instruments used without proper training or regard for their intended purpose will impair validity. • Principals may not be qualified to evaluate teachers on measures highly specialized for certain subjects or contexts.
Analysis of Classroom Artifacts	Structured protocols used to analyze classroom artifacts in order to determine the quality of instruction in a classroom. Artifact examples: lesson plans, teacher assignments, assessments, scoring rubrics, and student work.	Pilot research has linked artifact ratings to observed measures of practice, quality of student work, and student achievement gains. More work is needed to establish scoring reliability and determine the ideal amount of work to sample. Lack of research exists on the use of structured artifact analysis in practice.	• Can be a useful measure of instructional quality if a validated protocol is used, if raters are well-trained for reliability, and if assignments show sufficient variation in quality. • Is practical and feasible because artifacts have already been created for the classroom.	• More validity and reliability research is needed. • Training knowledgeable scorers can be costly but is necessary to ensure validity. • This measure may be a compromise in terms of feasibility and validity between full observation and less direct measures such as self-report.

Measure	Description	Research	Strengths	Cautions
Portfolios	A collection of teaching materials and artifacts assembled by the teacher to document a large range of teaching behaviors and responsibilities. Has been used widely in teacher education programs and in states for assessing the performance of teacher candidates and beginning teachers.	Research on validity and reliability is ongoing, and concerns have been raised about consistency of scoring. There is a lack of research linking portfolios to observed changes in teaching practice or student achievement. Some studies have linked NBPTS certification (which includes a portfolio) to student achievement, but other studies have found no relationship.	• Is comprehensive; can measure aspects of teaching that are not readily observable in the classroom. • Can be used with teachers of all fields. • Has a high level of credibility among stakeholders. • Is a useful tool for teacher reflection and improvement.	• This measure is time-consuming for teachers and scorers; scorers should have content knowledge of the portfolios they score. • Stability of scores may not be high enough to use for high-stakes assessment. • Portfolios are difficult to standardize (compare across teachers or schools). • Portfolios represent teachers' exemplary work but may not reflect everyday classroom activities.
Self-Report of Practice	Teacher reports of their practices, techniques, intentions, beliefs, and other teaching elements assessed through surveys, instructional logs, or interviews. Measures cover a broad spectrum and can vary widely in focus and level of detail.	Studies on the validity of teacher self-report measures present mixed results. Highly detailed measures of practice may be better able to capture actual teaching practices but may be more difficult to establish reliability or may result in very narrowly focused measures.	• Can measure unobservable factors that may affect teaching, such as knowledge, intentions, expectations, and beliefs. • Provides the unique perspective of the teacher. • Is feasible and cost-efficient; can collect large amounts of information at once.	• Reliability and validity of self-report has not been fully established and depends on instrument used. • Using or creating a well-developed and validated instrument will decrease cost-efficiency but will increase accuracy of findings. • This measure should not be used as the sole or primary measure in teacher evaluation.

Measure	Description	Research	Strengths	Cautions
Student Evaluation	Surveys or rating scales used to gather student opinions or judgments about teaching practice and to provide information about teaching as perceived by students. Measures can vary widely in focus and level of detail.	Several studies show that student ratings of teachers may be as valid as judgments made by college students and other groups and, in some cases, may correlate with measures of student achievement; thus students can provide useful information about teaching. Validity is dependent on the instrument used and its administration and is generally recommended for formative use only.	• Provides perspective of students, who have the most experience with teachers. • Can provide formative information to help teachers improve practice in a way that will connect with students. • Can potentially provide ratings as accurate as those provided by adult raters.	• Student ratings have not been validated for use in summative assessment and should not be used as the sole or primary measure of teacher evaluation. • Students cannot provide information on aspects of teaching such as a teacher's content knowledge, curriculum fulfillment, or professional activities.

Education Commission of the States

Issue Paper

June 2007

Teacher Evaluation in Diversified Teacher Compensation Systems

The Joyce Foundation

by Angela Baber

The second paper in a series of four

Introduction to the Project

Changing the way teachers are paid to include outcomes, such as student performance, or incentives for teaching in at-risk schools is gaining support in districts and states across the country. The policy process for moving away from the traditional compensation structure is a complex one, however. As is true in all sound policymaking, those designing and seeking to implement **diversified teacher pay systems** would benefit from reviewing what has been learned by both the research and policymaking communities in order to design programs with a better chance at succeeding. With the generous support of the Joyce Foundation, the Education Commission of the States has created a series of resources to provide policymakers and leaders with information on redesigned compensation systems. The resources include:

- An issue site on the ECS Web site with current resources
- A redesigned teacher compensation database with information on state-, district- and local-level redesigned compensation programs
- A series of four issue papers:
  - *Funding Issues in Diversified Teacher Compensation Systems*
  - *Teacher Evaluation in Diversified Teacher Compensation Systems*
  - *Student Performance Assessment in Diversified Teacher Compensation Systems*
  - *The Use of Diversified Compensation Systems to Address Equitable Teacher Distribution.*

We hope these resources are of value and relevance to policymakers and practitioners who are considering redesigning teacher compensation systems in their states, districts and schools.

Table of Contents

Note: all URLs in this paper are live links.

INTRODUCTION TO DIVERSIFIED TEACHER COMPENSATION 3

TEACHER EVALUATION IN DIVERSIFIED TEACHER COMPENSATION PROGRAMS 4

Knowledge- And Skill-Based Evaluation Criteria 4

The Danielson Framework for Teaching (1996) 5

Performance- or Standards-Based Evaluation Criteria 5

Student Achievement 6

Classroom Observation and Teacher Evaluations 6

PROVEN PROGRAM EXAMPLES 7

Cincinnati Public Schools Teacher Evaluation System (TES) 7

Development of the Knowledge- and Skills-Based Teacher Evaluation System in Cincinnati 7

Teacher Evaluation Component 7

The Relationship between Teacher Performance Evaluation Scores and Student Achievement 8

Vaughn Next Century Learning Center 8

Development of the Knowledge-, Skills- and Performance-Based Teacher Evaluation System for Vaughn 8

Teacher Evaluation Component 9

The Relationship between Teacher Performance Evaluation Scores and Student Achievement 9

Teacher Advancement Program (TAP) 10

Development of TAP Teacher Evaluation System 10

Teacher Evaluation Component 10

Results 10

PROMISING PROGRAM EXAMPLES 11

Denver Professional Compensation Program (ProComp) 11

Development of a Teacher Evaluation System for ProComp 11

Teacher Evaluation Component 12

Quality Compensation (Q Comp) 12

Development of Instructional Observation and Standards-Based Assessment Program 12

Teacher Evaluation Component 12

CONCLUSION 14

POLICY RECOMMENDATIONS 15

BIBLIOGRAPHY 16

ENDNOTES 18

© 2007 by the Education Commission of the States (ECS). All rights reserved.
The Education Commission of the States is a nationwide, nonprofit organization that helps state leaders shape education policy. ECS is pleased to have other organizations or individuals share its materials with their constituents. To request permission to excerpt part of this publication, either in print or electronically, please e-mail ecs@ecs.org.

Document No. TQ-07-02.

INTRODUCTION TO DIVERSIFIED TEACHER COMPENSATION

Teacher quality is one of the greatest determinants of student achievement. It follows, therefore, that ensuring all students are taught by quality teachers is a priority, one that has been the subject of increasing focus with the passage of the No Child Left Behind Act in 2001 and its emphasis on establishing a minimum standard for highly qualified teachers and accountability for student performance. One of the ways in which policymakers are attempting to improve teacher quality and ensure all students are taught by a high-quality teacher is through changes in the system by which teachers are compensated.

Attempts to move teacher compensation systems away from the single salary schedule in which teachers are compensated based on years of service and educational attainment to one more reflective of teacher performance are not new. Earlier attempts at diversification fell into two basic categories: experimental merit pay and career-ladder systems; neither enjoyed uniform success.[1]

Experimental merit pay systems were limited in several ways. First, they tended to rely solely on subjective evaluation of the teacher by a school administrator as the means of determining bonus distribution. Additionally, these were zero-sum systems, meaning the number and amount of bonuses were limited by the lump sum given to a school for this purpose. These limitations contributed to the claim these systems created competition among teachers. Moreover, these programs showed no evidence they improved overall teacher quality or student success.[2]

Career-ladder systems were also tried as an attempt to eliminate the flat career structure of the teaching profession. These systems provided additional salary and advancement opportunities for teachers who assumed additional roles such as mentoring and administrative responsibilities. While these programs showed promise through some improvement in student achievement, many programs were not able to obtain sustainable funding. However, certain aspects of career-ladder systems exist today within diversified teacher compensation programs. For a more complete discussion of teacher compensation reform efforts please see the ECS issue paper, *Diversifying Teacher Compensation* available at: **http://www.ecs.org/clearinghouse/65/83/6583.pdf**.

Modern reform attempts are more sophisticated in their design and tend to include multiple methods of evaluation, rewards for taking on leadership roles and links to outcome-based assessment such as student performance. Many programs also reflect the goals of the schools, districts and states by offering focused incentives to address high-need or challenging areas. Further, it is important to note that these programs are likely to be most effective as part of a larger system of teacher support.

This is the second in a series of four issue papers that highlight and discuss various aspects of diversified teacher compensation systems.

The four papers in the series are:

- *Funding Diversified Teacher Compensation Systems* (**http://www.ecs.org/clearinghouse/74/75/7475.pdf**)
- *Teacher Evaluation in Diversified Teacher Compensation Systems* (**http://www.ecs.org/clearinghouse/74/78/7478.pdf**)
- *Student Performance Assessment in Diversified Teacher Compensation Systems* (**http://www.ecs.org/clearinghouse/74/76/7476.pdf**)
- *The Use of Diversified Compensation Systems to Address Equitable Teacher Distribution* (**http://www.ecs.org/clearinghouse/74/77/7477.pdf**)

These issue papers were created with the generous support of the Joyce Foundation as part of a larger project on redesigned teacher compensation systems. Other resources produced through this project include an issue site on teacher compensation (available through the ECS Education Issues site) and a database containing information on state-, district- and local-level diversified compensation systems (available at: **http://www.ecs.org/html/t_comp.htm**).

TEACHER EVALUATION IN DIVERSIFIED TEACHER COMPENSATION PROGRAMS

Central to diversified teacher compensation systems is the ability to effectively assess teacher performance with the goal of identifying and rewarding effective teachers. Performance-based pay is typically dependent on the evaluation of student achievement, teacher classroom performance and performance rubrics.[3] Knowledge- and skills-based pay is a more recent evolution in evaluating teachers and involves complex teacher evaluation instruments, such as the Danielson Rubric, to assess the attainment and application of new and relevant abilities of individual teachers. School-based performance awards, or group performance incentives, are based on a school or a group of teachers meeting different types of pre-set goals. Career ladders usually combine some or all of the pervious forms of evaluating teachers and reward teachers by assessing their level of performance, knowledge and skills, goal attainment, professional development, collaboration, leadership and additional duties by placing them on higher and higher levels of the career ladder which coincide with higher and higher levels of compensation.

Programs that have successfully implemented teacher evaluation systems with proven correlations between teacher evaluation results and student learning, as well as programs that show promise to evolve this link, are of great interest to policymakers in designing alternative evaluation systems for teachers. The **Vaughn Next Century Learning Center, Cincinnati Public Schools Teacher Evaluation System, Teacher Advancement Program (TAP), Minnesota Quality Compensation (Q Comp)** and **Denver Professional Compensation (ProComp)** programs have successful and promising teacher evaluation methodologies and systems. The designers of these programs realized the importance of incorporating most or all of the multiple types of evaluation criteria mentioned above. Performance-based evaluations and techniques of determining a teacher's knowledge and skills rely most heavily on evaluation systems designed to effectively determine an individual teacher's abilities and are the focus of this policy brief.

KNOWLEDGE- AND SKILLS-BASED EVALUATION CRITERIA

Before knowledge- and skills-based evaluation systems can be integrated into a teacher salary structure, the types of knowledge and skills to be assessed need to be identified and clear articulation of how incentives will be tied to these criteria communicated to stakeholders. There are multiple methods of knowledge and skill identification. Some formal methods include the **1996 Charlotte Danielson Framework for Teaching** and the **National Board for Professional Teaching Standards (NBPTS)**. Other more localized methods are implemented through research based on determining local educational goals. Many programs started with existing standards or definitions of good teaching and then adapted evaluation criteria to better suit local objectives. Multiple case studies of knowledge- and skills-based evaluation systems show that adapting existing standards or definitions of good teaching allowed design and implementation in a relatively short amount of time.[4]

Valid and reliable methods of assessment that are recognized as such by teachers are a requirement of successful knowledge and skills based systems. Teacher input is also integral to informing the selection of criteria so teachers feel what they view as important evaluation measures are included. This in turn increases teacher buy-in to the system. Additionally, the more substantial the incentives, the more effective the model is in motivating knowledge and skill acquisition by teachers.[5] Policymakers need to recognize that a pay structure dependent on the attainment of knowledge and skills requires that quality professional development linked to these knowledge and skills must be made available to the teachers. Not doing so could diminish teacher motivation and compromise the evaluation structure thereby undercutting program objectives.

Programs in This Report:

- Vaughn Next Century Learning Center
- Cincinnati Public Schools Teacher Evaluation System
- Teacher Advancement Program (TAP)
- Minnesota Quality Compensation (Q Comp)
- Denver Professional Compensation (ProComp)

Program	Evaluation Frameworks and Rubrics
Cincinnati Public Schools (OH)	Adapted from the Danielson Framework for Teaching.
Vaughn Next Century Learning Center (CA)	Adapted from the Danielson Framework for Teaching, but not primarily based on external standards.
TAP (multi-district and multi-state)	TAP Teaching Skills, Knowledge and Responsibility Standards based on research in educational psychology and cognitive science, as well as best practices in the field including: • Interstate New Teacher Assessment and Support Consortium (INTASC) • The National Board for Professional Teacher Standards • Massachusetts' Principles for Effective Teaching • California's Standards for the Teaching Profession • Connecticut's Beginning Educator Support Program, and the New Teacher Center's Developmental Continuum of Teacher Abilities • Danielson Framework for Teaching.
Denver ProComp (CO)	The Comprehensive Professional Evaluation (CPE) was locally designed but has similarities to existing frameworks.
Minnesota Q Comp	Based on the TAP model

The Danielson Framework for Teaching (1996)

Of the five programs discussed in this policy brief, two evaluation systems are adopted directly from the Danielson model, two rely on the Danielson model and the last has similarities consistent with the Danielson Framework. The Framework originated from Charlotte Danielson's work with Educational Testing Service (ETS) and the National Board for Professional Teaching Standards. The Framework articulates an effective means for communicating the nature of an excellent teacher to students enrolled in teacher education programs, while also suggesting an array of indicators of a successful teaching experience. Danielson argues that excellence in teaching is organized around four domains: Planning and Preparation, Classroom Environment, Instruction and Professional Responsibilities. Each domain consists of several components that comprise superior, assessable professional practices.

Researchers have reported some shortcomings of the Danielson Framework. According to Allen Odden in *Lessons Learned About Standards-Based Teacher Evaluation Systems* (2004), the Danielson system, or at least the versions used by the programs studied in his research, does not address the following key aspects of instruction:

- Assessments teachers used to measure student learning
- Feedback teachers gave to students on these assessments
- How teachers scored student work to district or state student performance standards
- Teacher reflection on the effectiveness of their instructional practice and how that reflection would lead to changes in instructional practice
- Data on actual student achievement.

For detailed examples of the Danielson rubric see: **Charlotte Danielson Enhancing Professional Practice: a Framework for Teaching 1996 (http://www.cesa11.k12.wi.us/Content/ProfessionalDevelopment/Initiatives/PI34/Pages/Danielson%20Rubric.pdf).**

Performance- or Standards-Based Evaluation Criteria

Evaluation criteria based on teacher performance include student achievement; classroom observations designed to assess a teacher's abilities; parent, peer and student surveys; teacher interviews; attendance rates; graduation rates; and goals and objectives set independently by the teacher or with guidance from administrators.

Student Achievement

Student achievement is perhaps the most widely used performance-based evaluation component in diversified teacher compensation programs. Although researchers argue that a teacher's impact on measured student achievement should be a substantial factor in evaluating teacher effectiveness, they also warn against relying too heavily on this measure alone.[6]

Student achievement is gauged through the use of standardized state tests, value-added models that attempt to isolate the effect the teacher has on his or her students, post- and pre-tests intended to measure individual student learning gains over the school year, and knowledge benchmarks set by teachers and administrators. This issue paper focuses on those evaluation criteria other than student achievement that are used in evaluating a teacher's performance. For a review of student performance assessment see the third issue paper in this series, **Student Performance Assessment in Diversified Compensation Systems (http://www.ecs.org/LINK).**

Classroom Observation and Teacher Evaluations

There are many concerns to address before evaluating teachers based on classroom observations. Teachers need to have a good understanding of the evaluation criteria and how they will be rated and compensated based on the evaluation outcomes. If teachers are involved in the design of evaluation criteria, there is a better chance they will trust the criteria are relevant and objective. Evaluations should be performed at multiple times throughout the year by a team of trained evaluators. This approach is a requirement in each of the five successful and promising programs highlighted in this paper. An appeals process is also an important component of a successful evaluation system, as well as the creation of an action plan for improvement, through professional development or other means, for teachers who receive low scores.

Programs that use multiple evaluators who report similar scores (high inter-rater reliability) report high degrees of teacher trust in the evaluation system. Cincinnati Public Schools, Vaughn Next Century Learning Center, the Teacher Advancement Program (TAP), Denver ProComp and Minnesota Q Comp all base their classroom observations on knowledge- and skills-based rubrics. Additionally, they incorporate other performance-based evaluation criteria such as teacher portfolios, which include units and lesson plans, attendance records, student work, family contact logs and documentation of professional development activities.

Program	Evaluations per Year	Evaluators	Evaluator Requirements
Cincinnati Public Schools (OH)	6	Four evaluations are made by a teacher evaluator from outside the school. Two observations are performed by building administrators.	Teacher evaluators have subject matter and grade-level expertise similar to that of the teacher being evaluated. All evaluators are required to use the same evaluation criteria.
Vaughn Next Century Learning Center (CA)	2	An administrator and a peer perform the classroom observations. The teacher also performs a self-evaluation.	All evaluators are required to use the same evaluation criteria.
TAP (multi-district, multi-state)	4-6	Multiple trained and certified evaluators	Evaluators must be trained and certified.
Denver ProComp (CO)	1 or every 3rd year depending on seniority	Peers and Principals	There are three types of evaluation; probationary, non-probationary and special evaluations. All evaluators are required to use the same evaluation criteria.
Minnesota Q Comp	Multiple	Peers and Principals	Evaluators must be trained. Peer reviews are performed by master and mentor teachers. All evaluators are required to use the same evaluation criteria.

PROVEN PROGRAM EXAMPLES

Successfully Implemented Teacher Evaluation Systems with Proven Correlations between Teacher Evaluation Results and Student Learning

Cincinnati Public Schools Teacher Evaluation System (TES)

Cincinnati Public Schools implemented a comprehensive system for evaluating teachers known as the Teacher Evaluation System (TES). This evaluation method is used on an annual basis to determine teacher movement on a traditional salary schedule. The original plan was to have two phases of implementation; the second phase intended to tie compensation to a teacher's TES ranking. However, in May 2002, the teachers' union voted by 96.3% – 1,892 to 73 – against the second phase of this plan.[7] Although the performance-based pay component was not passed in 2002, research shows a link between TES scores and student achievement, making the Cincinnati program a valuable case study for policymakers interested in teacher evaluation systems connected to student learning gains.[8]

Development of the Knowledge- and Skills-Based Teacher Evaluation System in Cincinnati

The Teacher Evaluation Committee, one of three committees established to form a design structure for the Cincinnati model, was responsible for the design of a new teacher evaluation system. The three committees, referred to as The Committee of the Whole, were comprised of 24 union representatives, 12 administrative representatives which included seven teachers certified by the National Board for Professional Teaching Standards, a National Board Member and several teachers who had completed training to be observers for the Praxis III assessment program. After studying multiple teaching standards models (National Board for Professional Teaching Standards, the Interstate New Teacher Assessment and Support Consortium, the Praxis III assessment and the Danielson Framework for Teaching) the committee chose to use the Danielson Framework for Teaching to define quality teaching because is was more closely aligned with what good teaching meant for Cincinnati.[9] In modifying these teaching standards to better fit the needs of Cincinnati Public Schools, the committee addressed appropriate definitions of domains and standards, shortcomings of the Danielson model and how to weight standards.

For a detailed account of the planning process in Cincinnati see the Consortium for Policy Research in Education (CPRE) research paper, **How Cincinnati Developed a Knowledge- and Skills-Based Salary Structure (http://www.wcer.wisc.edu/cpre/papers/Cincinnati%20KSBP%203-00.pdf)**.

Teacher Evaluation Component

The Teacher Evaluation System (TES) in Cincinnati is based on 16 standards divided into four domains. A teacher's performance is measured against each of these standards. The standards are aggregated into four scores, one for each of the domains. Teachers can earn from three to 24 points in each of the following:

- Planning and Preparing for Student Learning
- Creating an Environment for Learning
- Teaching for Learning
- Professionalism.

There are five teaching levels based on a teacher's evaluation scores: Apprentice, Novice, Career, Advanced and Accomplished Teacher. Increases in salary are associated with each of these levels and teachers move up the salary schedule through evaluation of TES scores. The first two levels, Apprentice and Novice, have a time limit associated with them. A teacher must progress from one category to the next within a specified period of time to continue his/her contract. If a teacher receives an evaluation that would place him or her in a lower category, the teacher's salary increase is withheld and he or she must undergo a second comprehensive evaluation the following year.

For two of the TES domains – creating an environment for learning and teaching for learning – evaluations are performed six times a year. Four of these evaluations are performed by a teacher from another school with equivalent subject-matter and grade-level expertise to the teacher being evaluated. The remaining two evaluations are performed by school administrators, either the principal or vice-principal of the school. Final summative ratings for these two domains are made from these six observations. Teachers are rated on the remaining domains – planning for student learning and professionalism – by administrators. Portfolios including units and lesson plans, attendance records, student work, family contact logs, and documentation of professional development activities are used to rate teachers on these two domains.

For more information on the Cincinnati TES domains and standards see the **CPS Standards and Domains Rubric (http://www.cps-k12.org/employment/tchreval/stndsrubrics.pdf)**.

The Relationship between Teacher Performance Evaluation Scores and Student Achievement

Research shows the teacher assessment system in Cincinnati is able to identify which teachers had students with higher-than-expected levels of achievement. These results support using the TES for teacher evaluation and pay differentiation, as well as using teacher evaluation scores to determine teaching practices that affect student learning. For a detailed account of the research supporting the correlation between TES scores and student achievement, see the Consortium for Policy Research in Education (CPRE) research article, **The Relationship Between Standards-Based Teacher Evaluation Scores and Student Achievement: Replication and Extension at Three Sites** (**http://www.wcer.wisc.edu/cpre/papers/3site_long_TE_SA_AERA04TE.pdf**).

Vaughn Next Century Learning Center

Vaughn Next Century Learning Center, located in Pacoima, is a large urban public school within the Los Angeles Unified School District (LAUSD). Since the early 1970's, low student achievement had been a pattern in this school. In 1993 Vaughn became the first conversion charter school in the nation and was authorized by LAUSD. The charter was renewed in 1998 and again in 2003. Vaughn is now a full-service, community-based PK-12 charter school serving almost 2,000 neighborhood children, almost 100% of whom are eligible for free or reduced lunch and over 50% of whom are English language learners. As a result of their efforts, Vaughn was awarded the California Distinguished Schools Award in 1996 and the National Blue Ribbon Schools Award in 1997. In addition to diversifying teacher pay based on alternative evaluation methods, Vaughn also defines teaching environment, teacher training, and professional growth and teacher leadership program components.

In addition to a base pay and extra compensation for certification and advance degrees, Vaughn pays teachers based on knowledge and skills; contingency-based awards (for achieving certain goals in the areas of student attendance, discipline, parental involvement and for working in teams); schoolwide student achievement bonuses; expertise compensation (teachers in leadership roles including grade-level chairs, committee chairs, peer reviewers, mentors, faculty representatives, etc.); and gain-sharing (unused sick days can accrue for monetary reimbursement). Added benefits include a long-term disability insurance policy for every teacher that provides 60% of their full pay until age 65. In addition, an account with $500,000 in the Los Angeles Teachers' Credit Union was established to guarantee health benefits after retirement.

Development of the Knowledge-, Skills- and Performance-Based Teacher Evaluation System for Vaughn

The driving force for restructuring the Vaughn school was to increase student achievement. Professional development programs are mandatory, but teacher choice is also a part of the process, with outside experts and dedicated administration support time for training, evaluation and professional development for their assigned staff. Vaughn has a curriculum committee responsible for planning staff development. Effective with the 1999-2000 school year all teachers are evaluated through a process based on knowledge and skills performance assessments. This evaluation includes a self evaluation by the employee using the same evaluation tool the peer reviewer and administrator use; several classroom observations each semester; and consideration of relevant staff development training taken by the employee.

The Consortium for Policy Research in Education (CPRE) conducted surveys and site visits and used the information gathered to report on the first, second and third year of the knowledge- and skills-based pay program at Vaughn. During the first year Vaughn addressed concerns about the heavy burden on evaluators conducting three annual teacher observations that may have been too long, sometimes spanning a full day. Although the first year was viewed as a difficult year, the primary reason was not due to the implementation of a new knowledge- and skills-based pay plan. Tension and friction between new and veteran staff developed. New staff were required to participate in the redesigned compensation system whereas a plan for expanding the new salary structure to incorporate veteran teachers was not yet implemented. New teachers felt they were working harder than the senior staff for less compensation. Veteran teachers were skeptical about plans to implement a similar knowledge- and skills-based pay system that would apply to all staff.

The second year at Vaughn saw the development of the expanded pay structure to teachers and administrators. One of the major challenges in developing new pay principles for teachers was that teachers near the top of the traditional salary schedule could be seriously disadvantaged by the new pay structure. Another challenge involved was deciding the relative value of these bonuses, because a set bonus would always be a higher percentage of a starting teacher's salary than an experienced teacher's salary. This led program de-

signers to implement a three-tiered knowledge, skills and/or proficiency system based, in part, on seniority. A commitment to develop rubrics was made to address concerns regarding the subjectivity of evaluators and evaluation scores for teachers.

CPRE reported that the third year of the knowledge- and skills-based pay program at Vaughn saw an improved sense of trust in the evaluation system and an overall positive climate. By this time the basic rubric was constructed and debugged, and the majority of staff had a good understanding of the evaluation system. A committee was created to guide ongoing work on the evaluation system; the time demand the system placed on evaluators was addressed; and previous teacher requests for more feedback were met by the creation of a Peer Assistance Review (PAR) committee. The creation of the PAR committee was reported as the most important change for year three. Some of the more significant issues that the PAR dealt with in year three include the following:

- An appeals process
- How to evaluate team teachers who do not typically teach all content areas
- How to evaluate teachers on technology when the computers in their classrooms are not working
- The establishment of a pre- and post-evaluation conference system
- Feedback and other opportunities for professional growth
- Inter-rater reliability among evaluators.

For more information on the development of the Vaughn knowledge and skills-based evaluation system see the Consortium for Policy Research in Education (CPRE) paper, **How Vaughn Next Century Learning Center Developed a Knowledge- and Skills-Based Program (http://www.wcer.wisc.edu/cpre/papers/Vaughn%20KSBP%208-01.pdf)**.

Teacher Evaluation Component

The Peer Assistance and Review System is the current system of evaluation and takes place three times per year. Teachers reflect on their own performance and rate themselves using established teaching standards and scoring rubrics. The scoring rubric was adopted from the Danielson Framework for teaching and selected peer reviewers (peers from Vaughn and/or objective outside peer reviewers) observe their colleagues and provide feedback as well as assistance. Instructional coordinators also conduct classroom visits and conference with teachers on an ongoing basis. Scoring from self, peer and instructional coordinators are averaged.

The evaluation system recognizes three levels of teachers – Apprentice, Level II and Level III – that are evaluated using a four-domain, multiple-component scoring rubric. There are two assistance visitations and two formal evaluations per year. During the formal evaluations, Apprentice-level teachers are scored on apprentice skill areas only. Level II and III teachers are scored on all four of the domains. Teachers are scored by a PAR member and an administrator, and are responsible for performing a self evaluation based on the same criteria. During the assistance visitation, a checklist is used as feedback for the teacher. The peer evaluator has a pre-visit conference with the peer teacher at the beginning of the year, and after each assistance visit the peer evaluator meets with the teacher to discuss the assistance checklist. Only the administrator, however, discusses the formal evaluation with the teacher.

For more information on the Vaughn teacher evaluation system components and for the full scoring rubric see **Vaughn Teacher Quality and Professional Growth (http://vaughncharter.com/s2/images/stories/miscfiles/PeerReview.pdf)**.

The Relationship between Teacher Performance Evaluation Scores and Student Achievement

Hierarchical Linear Modeling (HLM) was used to isolate classroom effects in order to ascertain the teacher's effect on student learning. According to a 2004 article by Alix Gallagher, results indicate a strong, positive and statistically significant relationship between teacher evaluation scores and student achievement in reading; a composite measure of teacher and student performance and a positive, although not statistically significant, relationship in mathematics.[10] Additionally, Vaughn has met the growth target established by the California Academic Performance Index (API) for seven years running. API increased by 263 points from 1999 to 2006.[11]

For a detailed account of the research supporting the correlation between Vaughn TES scores and student achievement, see the Consortium for Policy Research in Education (CPRE) research article, **The Relationship Between Standards-Based Teacher Evaluation Scores and Student Achievement: Replication and Extension at Three Sites (http://www.wcer.wisc.edu/cpre/papers/3site_long_TE_SA_AERA04TE.pdf)**.

Teacher Advancement Program (TAP)

The Teacher Advancement Program (TAP) was developed in 1998 and refined over the years based upon scientific research and supported by expert practice in the field. The TAP is funded by the Milken Family Foundation and is intended to attract, retain, develop and motivate talented people in the teaching profession. There are currently TAP programs in 32 districts across 13 states with charter school locations in 10 districts across eight states. TAP is based on the following four elements:

- Multiple career paths
- Ongoing, applied professional growth
- Instructionally focused accountability
- Performance-based compensation.

Teachers are compensated according to their roles and responsibilities, their performance in the classroom and the performance of their students. Districts are also being encouraged to offer competitive salaries to teachers in "hard-to-staff" subjects and schools.[12]

Development of the TAP Teacher Evaluation System

TAP teaching skills, knowledge and responsibility performance standards were developed based on education psychology research focusing on learning and instruction. Instructional guidelines and standards developed by numerous national and state teacher standards organizations were reviewed, and from this review a set of standards for teacher accountability was developed. The work reviewed included guidelines and standards developed by the Interstate New Teacher Assessment and Support Consortium (INTASC), the National Board for Professional Teacher Standards, Massachusetts' Principles for Effective Teaching, California's Standards for the Teaching Profession, Connecticut's Beginning Educator Support Program, and the New Teacher Center's Developmental Continuum of Teacher Abilities. The work of Danielson (1996) served as a valuable resource for defining the teaching competencies at each level of teacher performance.

TAP defined the skills of a quality teacher and determined how they are demonstrated at different levels of performance in order to measure a teacher's skills, knowledge and responsibilities. The TAP performance standards define the expected teaching skills, knowledge and responsibilities for each level teacher in the career path. The teacher responsibility rubrics were designed based on research in educational psychology and cognitive science, as well as best practices in the field.[13] These accountability systems include: Rochester Career in Teaching Program (NY), Douglas County Teacher's Performance Pay Plan (CO), Vaughn Next Century Charter School Performance Pay Plan (CA), and Rolla School District Professional Based Teacher Evaluation (MO).

Teacher Evaluation Component

TAP has a comprehensive system for teacher evaluation based on a combination of classroom observations and student achievement gains. A teacher's performance is measured against the TAP teaching skills, knowledge and responsibility standards. The standards involve teaching processes and outcomes, and are the basis for classroom observation evaluations. Teachers are evaluated four to six times each year by multiple trained and certified evaluators.[14]

Results

Many TAP locations report increased student achievement as well as higher rates of teacher retention. In 2002, three of four schools in Arizona using performance-pay plans under TAP performed significantly better than control schools, with 14- to 46-point percentile rank differences. Results in 2003 for South Carolina reported four of the six TAP schools performed significantly better in math than the control schools with 14- to 27-point percentile rank differences.

TAP was piloted in Minneapolis and Waseca, Minnesota, where promising results were reported based on preliminary data such as increased state and local student assessment results. Specifically, after the first year of TAP, Andersen Open, a K-8 school in Minneapolis, increased the number of 8th graders passing the Basic Skills Test in reading from 39% in 2004 to 62% in 2005.[15] Student achievement is thought to be improved because the teachers were provided with three-pronged job-embedded professional development:

1. Ongoing feedback and support from mentors and instructional coaches
2. Time during the school day to collaborate in professional development teams
3. A school improvement goal that focused on relevant and meaningful instructional strategies demonstrated by mentors and instructional coaches who had field tested the strategy with students at the school.

The successes of the Minnesota TAP program prompted policymakers to establish a statewide diversified teacher compensation program based on the TAP program principles and components. This program, Minnesota Q Comp, is also been highlighted in this brief issue paper as a promising teacher evaluation system.

For more information on the Teacher Advancement Program, visit the **National Institute for Excellence in Teaching (NIET)** site (**http://www.talentedteachers.org/tap.taf**).

PROMISING PROGRAM EXAMPLES

Denver Professional Compensation Program (ProComp)

Denver ProComp is the evolution of a pilot program initiated in 1999 through a partnership between the Denver Classroom Teachers Association and Denver Public Schools. A study of this pilot program was conducted by the Community Training and Assistance Center of Boston. One result of the initial findings was that a new teacher compensation agreement could not be based on student objectives alone. The ProComp system is a results-based pay program that uses multiple criteria to assess teachers' performance, with teachers receiving increases in pay and bonuses for documented results. A new teacher evaluation system was field-tested during the 2004-05 school year. The current ProComp salary system went into effect in January 2006 after Denver voters approved a $25 million mill levy to fund the compensation plan.

Denver ProComp provides incentives to teachers based on nine criteria divided into four categories. Teachers receive base-pay salary increases and bonuses and are eligible to receive tuition reimbursement for specific course work if they meet the evaluation criteria specified for each objective.

Development of a Teacher Evaluation System for ProComp

The new Comprehensive Professional Evaluation (CPE) system was designed collaboratively by teachers, administrators, specialists and parent representatives. The ultimate goal in this evaluation process is to improve student achievement by defining quality instruction and increasing the use of effective, research-based practices in a collaborative process.

Knowledge and Skills	
Professional development unit	2% salary increase
Graduate degree/national board certificate	9% salary increase
Course work	$1000 tuition reimbursement
Professional Evaluation	
Successful evaluation	3% salary increase
Student Growth	
Teachers/specialists set two annual objectives	1% salary increase
Teacher only meets one objective	1% index bonus
Meet CSAP accepted growth	3% salary increase
Distinguished school (student growth data, school climate, attendance and graduation rates)	2% index bonus
Market Incentives	
Hard-to-staff	3% index bonus
Hard-to-serve (% free/reduced lunches)	3% index bonus

Teacher Evaluation Standards[16]	Distinct Steps for the Formal Evaluation[17]
• Instruction • Assessment • Curriculum and Planning • Learning Environment • Professional Responsibilities.	• Orientation and notification of evaluation • Pre-observation conference with the evaluator • Formal observation • Post-observation conference • Mid-evaluation structured conference – educator provides the evaluator with records of teaching/service that address all five standards • Additional observation (probationary teachers) • Comprehensive evaluation – educator completes an action plan to be shared with the evaluator at the final meeting to sign documents.

Teacher Evaluation Component

There are three types of evaluation; probationary, non-probationary and special evaluations. Probationary evaluations are conducted annually during the educator's probationary employment; non-probationary evaluations are conducted every three years for teachers who have successfully completed their probationary period; and special evaluations are conducted when a supervisor determines that a teacher requires assistance in a non-evaluation year.

Teachers are evaluated on five standards. Each standard has multiple criteria with a number of indicators that can be used to evaluate whether the educator is meeting the standards. A criterion for assessment, for example, is that the educator uses and interprets a variety of assessments to monitor and evaluate students. Criteria for professional responsibilities include collaborating with identified teams on expectations, strategies and use of data; engaging families in the learning process; and demonstrating integrity, and professional and ethical standards.

Pay raises are based on a performance rating of "satisfactory." Evaluators use the performance evaluation process to identify how well educators are meeting the five performance standards and corresponding criteria. The ratings include: Exceeding (E), Meeting, (M), Developing (D) and Not Meeting (NM) expectations. A rating is given for each performance standard and each criterion on the comprehensive performance form. Evaluators rate a teacher's comprehensive performance as "unsatisfactory" if he or she determines a "Not Meeting" expectation for one or more performance standards or for a total of five or more criteria across all performance standards.

For more information on the **Denver ProComp** program, visit the ProComp site: **http://denverprocomp.org**.

Quality Compensation (Q Comp)

Quality Compensation, or Q Comp, is a performance-pay program adopted by the state of Minnesota. Performance pay is not a requirement in Minnesota; rather, districts apply to participate in Q Comp. Since the program was enacted by the Minnesota Legislature in July 2005, 22 Minnesota school districts were approved for Q Comp and received funding for implementation, and another 134 districts indicated that they plan to submit an application for the 2006-07 or 2007-08 school years.[18]

The Q Comp program is based on the Teacher Advancement Program (TAP) and has five components:

- Career ladders for teachers
- Job-embedded professional development
- Instructional observations and standards-based assessments
- Measures to determine student growth
- Alternative teacher compensation or performance pay.

Development of Instructional Observation and Standards-Based Assessment Program

The Minnesota Department of Education lists the basic steps that districts need to take in order to create a good instructional observation and standards-based assessment program under Q Comp. Evaluation standards are ultimately a district-level decision, but for a district plan to be approved for funding they must address the steps listed in the table on the next page.

Basic Steps Needed to Create a Good Instructional Observation and Standards-based Assessment Program under Q Comp[19]

1. Identify the standards by which the evaluation system is based (e.g. The Minnesota Standards of Effective Practice)
   a. This is recommended by the Minnesota Department of Education so the teacher evaluation program and the requirements for initial and continuing licensure are integrated and little additional bureaucracy or paperwork are added to the district.
   b. The teacher evaluation program should have processes and artifacts that are directly linked to the ongoing professional development teachers are expected to have to maintain their license.
2. Develop criteria for each standard that includes a description of the specific expectations for each performance level
   a. The standards need to be converted into a workable document that is easy for both the teacher and the evaluator to understand).
3. Determine the categories for a rubric of performance levels for each standard.
4. Fill the rubric with descriptions of each performance level for each criterion
   a. This should include the type of evidence necessary to prove that this is the level being demonstrated by the teacher's performance during the evaluation.
5. Select and train the evaluation team
   a. Q Comp requires that multiple evaluations be carried out by a team (multiple evaluators) trained in using evidence-based observations aligned with professional teaching standards.

Teacher Evaluation Component

Under Q Comp every teacher must be evaluated multiple times, every year. The performance-pay system must involve a comprehensive standards-based professional review system for teachers that utilizes input from a variety of sources. The review system must be based on scientifically based education research. Peer reviewers, such as master and mentor teachers, along with principals, will evaluate each teacher's performance at several points in time during the school year. The evaluations must be one consideration for teacher bonuses.[20]

A locally selected evaluation team develops a common set of skills to be measured and measures them with a common rating scale. Adopted from the TAP evaluation system, this system also relies on principles from the Interstate New Teacher Assessment and Support Consortium, the National Board for Professional Teacher Standards, Massachusetts' Principles for Effective Teaching, California's Standards for the Teaching Profession, Connecticut's Beginning Educator Support Program, the New Teacher Center's Developmental Continuum of Teacher Abilities, and the Danielson Framework for Teachers.

In order to assure fairness, all evaluators are required to use the same evaluation criteria. There are many ways teams can be structured at each of the various school levels, including: interdisciplinary, inter-grade level, departmental, grade level, etc. Q Comp guidelines establish that teams should include specialists in the areas of special education, elective courses (industrial technology, home economics, foreign language, band, choir, music, etc.), school nurses, administrators and counselors.

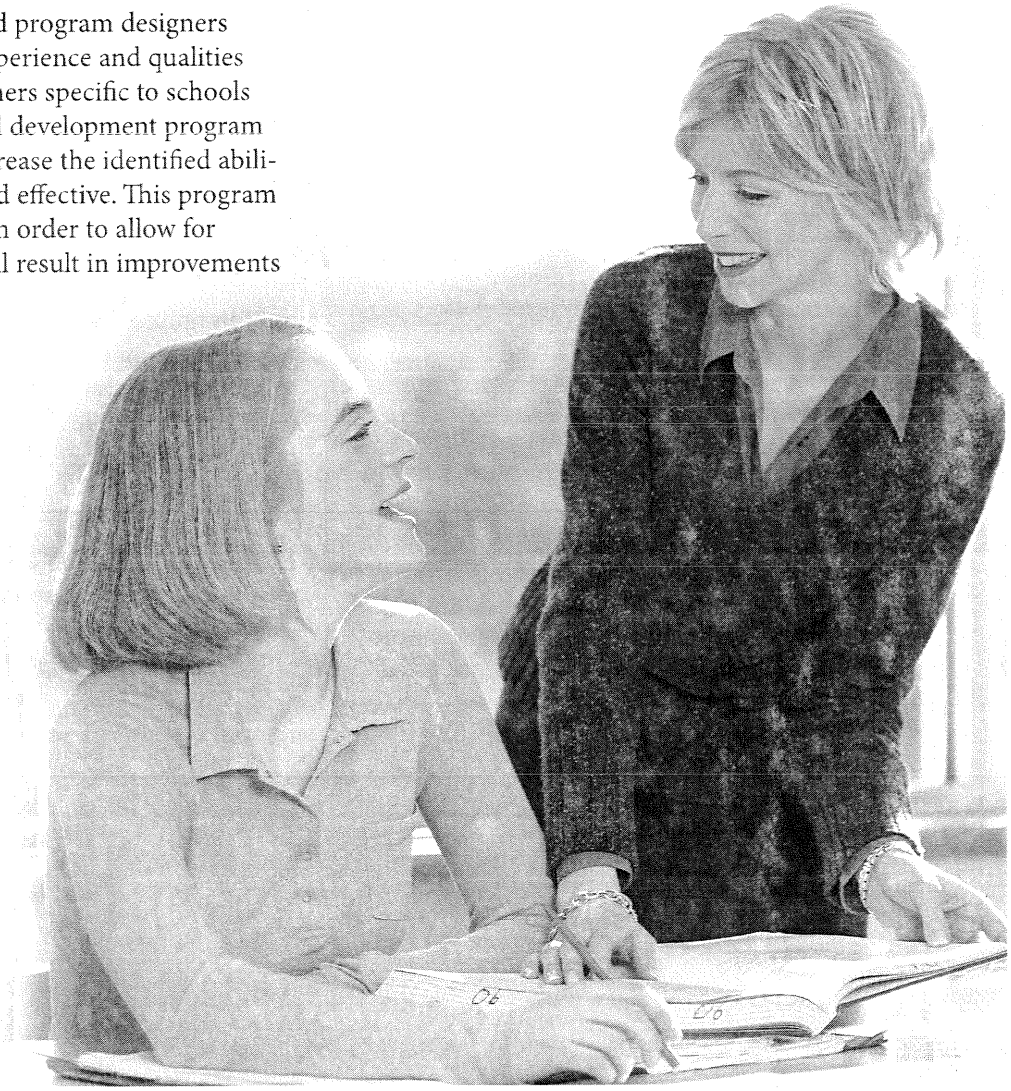
For more information on the Minnesota Q Comp programs visit the **Minnesota Department of Education Q Comp** site (**http://children.state.mn.us/mde/Teacher_Support/QComp/Program_Components/index.html**).

CONCLUSION

There are many factors to consider in planning, designing and implementing a teacher evaluation system. During the design process policymakers need to address the issue of data systems and the integration of the data systems used to track students, teachers and classes with the existing human resources system so the evaluation and compensation of teachers is technically possible.

A collaborative planning process has increased teacher buy-in in many programs, and the need to base evaluations on multiple observations by multiple trained evaluators is also reported to increase teachers' beliefs in the fairness of the evaluation process. The definition of a good teacher is dependent on context. A highly effective teacher at Vaughn might not be as effective at another school. For this reason it is important for policymakers and program designers to define what knowledge, skills, experience and qualities are needed to cultivate quality teachers specific to schools and regions. A targeted professional development program needs to be in place and should increase the identified abilities a teacher needs to be considered effective. This program needs to be funded and accessible in order to allow for teachers to develop abilities that will result in improvements in student learning.

The process of planning, designing, implementing, integrating, reviewing and revamping a teacher evaluation system is an arduous task, fraught with challenges. Time, dedication, a shared vision and willingness to collaborate with the goal of improved student achievement are necessary to craft an effective evaluation system. The process of implementing teacher evaluation systems may seem daunting; however, research has linked teacher evaluation scores from these systems to increases in student achievement across multiple programs. Components of successful programs, as well as lessons learned from these programs, inform the recommendations to policymakers outlined below. Redesigned teacher evaluation systems – if implemented with sincere consideration and caution – can and do fulfill the ultimate purpose of every teacher, school, administrator and education policymaker – *to improve student achievement.*



Policy Recommendations

1. Pre-Implementation

- Consider what the data system needs are to track students, teachers and classes, and how those data systems will need to be integrated into the existing human resources system
- For consistency and comparability, make sure student assessments are aligned along grade-levels
- Be prepared to determine a set of teacher standards that describes in considerable detail what teachers need to know and be able to do
- Adapting existing standards or definitions of good teaching can save time.

2. Designing the Evaluation System

- Involve professionals, stakeholders and practitioners in the design of the teacher evaluation system
- Determine the multiple forms of data that will need to be collected and have a plan of action with dedicated staff time to collect these data
- Utilize evaluation criteria from a variety of sources
- Decide how to weight evaluation standards and communicate standard weights to all stakeholders
- Define multiple teaching levels to differentiate between new and veteran teachers
- Use student achievement as a factor in evaluating teachers, but not as the only factor.

3. Evaluation and Evaluator Criteria

- Develop a related set of scoring rubrics that provide guidance to evaluators on how to score the data to various performance levels
- Train evaluators to be responsible for observing teachers in the classroom
- Have a system for averaging the multiple scores of the different evaluators to determine a teacher's instructional performance
- Implement multiple evaluations per year performed by more than one trained evaluator.

4. Incentives

- Decide how performance evaluation results will be attached to incentives
- The more substantial the incentive, the more effective the model is in motivating knowledge and skill acquisition by teachers
- Recognize that set bonus are a higher percentage of a starting teacher's salary than of an experienced teacher's salary.

5. Teacher Support and Growth

- Provide evaluators' feedback to teachers on evaluation outcomes
- Link targeted professional development to knowledge and skills needed to receive incentives
- Incorporate an appeals process for teachers
- Have an action plan for assisting teachers that receive low scores
- Provide teachers with time during the school day to collaborate in professional development teams.

6. Program Review and Refinement

- Track evaluation scores to ensure inter-rater reliability
- Have a review process in place to gauge teacher and administrator perceptions of the evaluation system and revise the system as necessary
- Address concerns about the evaluation system through a committee to guide ongoing work on the evaluation system.

REFERENCES

Azordegan, Jennifer; Byrnett, Patrick; Campbell, Kelsey; Greenman, Josh; Coulter, Tricia. *Diversifying Teacher Compensation*. Denver: Education Commission of the States, December, 2005. http://www.ecs.org/clearinghouse/65/83/6583.pdf (Accessed 12/2006)

Baber, Angela. Diversifying Teacher Compensation Database, Denver: Education Commission of the States, Policy Database, 2006. http://www.ecs.org/html/T_Comp.htm (Accessed November 2006)

Cincinnati Public Schools Web site. Teacher Evaluation Web site. http://www.cps-k12.org/employment/tchreval/tchreval.htm (Accessed December 1, 2006)

Danielson, Charlotte. "Teacher Evaluation Instrument, Rubrics from: Charlotte Danielson" *Enhancing Professional Practice: A Framework for Teaching*. 1996. http://www.cesa11.k12.wi.us/Content/ProfessionalDevelopment/Initiatives/PI34/Pages/Danielson%20Rubric.pdf (Accessed January 7, 2007)

Denver ProComp Web site. http://www.denverprocomp.org (Accessed November 2, 2006)

Education Commission of the States. Pay-for-Performance: Key Questions and Lessons from Five Current Models, Issue Paper. Denver: Education Commission of the States. June 2001. http://www.ecs.org.clearinghouse/28/30/2830.htm (Accessed November 5, 2006)

Gallagher, H. Alix. "Vaughn Elementary's Innovative Teacher Evaluation System: Are Teacher Evaluation Scores Related to Growth in Student Achievement?" *Peabody Journal of Education*, 79(4), 79-107. Lawrance Erbaum Associates, Inc. 2004.

Goldrick, Liam. *Improving Teacher Evaluation to Improve Teaching Quality*, Issue Brief. NGA Center for Best Practices, Education Policy Studies Division, December 9, 2002.

Gordon, Robert; Kane, Thomas; and Staiger, Douglas O. The Hamilton Project. *Identifying Effective Teachers Using Performance on the Job*, Discussion Paper 2006-01, The Brookings Institute. April, 2006. http://www.brook.edu/views/papers/200604hamilton_1.pdf (Accessed November 2006)

Halverson, Richard; Kelley, Carolyn; Kimball, Steven. "Implementing Teacher Evaluation Systems: How Principals Make Sense of Complex Artifacts to Shape Local Instructional Practice." *Educational Administration, Policy and Reform: Research and Measurement Research and Theory n Educational Administration*, Volume 3. Greenwich, CT: Information Age Press. 2004.

Hanushek, Eric A.; Kain, John F.; O'Brian, Daniel M; and Rivkin, Steven G. *The Market for Teacher Quality*, National Bureau of Economic Research. February, 2005.

Keller, Bess. "Cincinnati Teachers Rebuff Performance Pay." *Education Week* (May 29, 2002): 5. http://www.edweek.org/ew/articles/2002/05/29/39cincy.h231.html (subscription required).

Kellor, Ellen; Milanowski, Tony; Odden, Allan; Gallagher, H. Alix. *How Vaughn Next Century Learning Center Developed a Knowledge- and Skill- Pay Program*. Madison: University of Wisconsin, Wisconsin Center for Education Research, Consortium for Policy Research in Education. May, 2001. http://www.wcer.wisc.edu/cpre/papers/Vaughn%20KSBP%208-01.pdf (Accessed January 2, 2007)

Los Angeles Unified School District. Vaughn Next Century Learning Center Web site. http://vaughncharter.com/s2/index.php?option=com_content&task=view&id=23&Itemid=41 (Accessed December 12, 2006)

McCaffrey, D.F.; Lockwood, J.R.; Koretz, D.M.; and Hamilton, L.S. *Evaluating Value-Added Models for Teacher Accountability*. Santa Monica, CA: RAND Corporation, 2003.

Milanowski, Anthony. *The Criterion-Related Validity of the Performance Assessment System in Cincinnati*, Working Paper Series TC-03-05. Madison: University of Wisconsin, Wisconsin Center for Education Research, Consortium for Policy Reseach in Education. March, 2003. **http://www.wcer.wisc.edu/cpre/papers/CinciAEFA03_TE.pdf** (Accessed December 2, 2006)

Milanowski, Anthony. "The Relationship Between Teacher Performance Evaluation Scores and Student Achievement: Evidence from Cincinnati" *Peabody Journal of Education*, 79(4), 33-53. Lawrence Eribaum Associates. 2004.

Milanowski, A.T. *The Varieties of Knowledge and Skill-Based Pay Design: A Comparison of Seven New Pay Systems for K-12 Teachers*, Research Report Series RR-050. Pennsylvania: University of Pennsylvania, Graduate School of Education, Consortium for Policy Research in Education. **http://epaa.asu.edu/epaa/v11n4/**. (Accessed December 11, 2006).

Milanowski, A.T. and Kimball, S.M. *The Framework-Based Teacher Performance Assessment Systems in Cincinnati and Washoe* (Working Paper TC-03-07). Paper presented at the Annual Meeting of the American Educational Association held in Chicago, IL. Madison: University of Wisconsin, Wisconsin Center for Education Research,Consortium for Policy Research in Education, April 2003.

Milanowski, Anthony; Kimball, Steven; and White, Brad. *The Relationship Between Standards-Base Teacher Evaluation Scores and Student Achievement: Replication and Extension at Three Sites*, Working Paper Series, TC-04-01. Madison: University of Wisconsin, Wisconsin Center for Education Research, Consortium for Policy Research in Education. March, 2004. **http://www.wcer.wisc.edu/cpre/papers/3site_long_TE_SA_AERA04TE.pdf** (Accessed January 7, 2007)

Milanowski, A.T.; Kimball, S.M.; and White, B. *The Relationship Between Standards-Based Teacher Evaluation Scores and Student Achievement: Replication and Extensions at Three Sites*, Paper presented at the American Educational Research Association held in San Diego, CA. Madison: University of Wisconsin, Wisconsin Center for Education Research, Consortium for Policy Research in Education, April 2004.

Milken Foundation. Teacher Advancement Program Foundation Web site. **http://www.tapschools.org/tap/tap.taf?page=tapstates**. (Accessed December 16, 2006)

Minnesota Department of Education. Minnesota Department of Education Quality Compensation Web site. **http://children.state.mn.us/mde/Teacher_Support/QComp/Program_Components/index.html** (Accessed November 2, 2006)

Odden, Allen. "Lessons Learned About Standards-Based Teacher Evaluation Systems" *Peabody Journal of Education*, 79(4), 126-137. Lawrence Erbaum Associates, Inc. 2004

Odden, A., and Kellor, E. *How Cincinnati Developed a Knowledge and Skills-Based Salary Schedule: Consortium for Policy Research in Education*. Madison: University of Wisconsin-Madison, Consortium for Policy Research in Education, March 2000. (Accessed 01/2007)

Odden, A.R.; Kelley, C.; Heneman, H.; and Milanowski, A. *Enhancing Teacher Quality through Knowledge- and Skills-Based Pay*. Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education, 2001. **http://www.cpre.org/Publications/rb34.pdf** (Accessed December, 2005)

Odden, A.R. and Kelley, C. *Paying Teachers for What They Know and Do: New and Smarter Compensation Strategies to Improve Schools*. Thousand Oaks, CA: Corwin Press, 1997.

ENDNOTES

[1] Allan Odden and Carolyn Kelley, *Paying Teachers for What They Know and Do: New and Smarter Compensation Strategies to Improve Schools.* (Thousand Oaks, CA: Corwin Press, 2002).

[2] H.P. Hatry, J.M. Greiner, and B.G. Ashford, *Issues and Case Studies in Teacher Incentive Plans.* (Washington, DC: The Urban Institute, 1994); A.M. Mohrman, S.A. Mohrman, and Allan R. Odden, "Aligning Teacher Compensation with Systemic Reform: Skill-Based Pay and Group-Based Performance Awards." *Educational Evaluation and Policy Analysis*, 18 (1996): 51-71; R.J. Murnane and D.K. Cohen, "Merit Pay and the Evaluation Problem." *Harvard Educational Review*, 56(1) (1986): 1-17.

[3] For a review of student performance assessment see another brief in this series, Student Performance Assessment in Diversified Compensation Systems.

[4] Anthony Milanowski, *The Varieties of Knowledge and Skills Based Pay Design: A Comparison of Seven New Pay Systems for K-12 Teachers*, CPRE Research Report Series RR-050 (Madison: University of Wisconsin, Wisconsin Center for Education Research) October, 2002.

[5] Anthony Milanowski, *The Varieties of Knowledge and Skills Based Pay Design: A Comparison of Seven New Pay Systems for K-12 Teachers*, CPRE Research Report Series RR-050 (Madison: University of Wisconsin, Wisconsin Center for Education Research) October, 2002.

[6] Robert Gordon, Thomas Kane and Douglas O. Staiger, *The Hamilton Project. Identifying Effective Teachers Using Performance on the Job*, Discussion Paper 2006-01, (The Brookings Institute. April, 2006. **http://www.brook.edu/views/papers/200604hamilton_1.pdf**)

[7] Cincinnati Public Schools Web site. Teacher Evaluation Web site. **http://www.cps-k12.org/employment/tchreval/tchreval.htm** (Accessed December 1, 2006)

[8] Anthony Milanowski. *The Relationship between Teacher Performance Evaluation Scores and Student Achievement: Evidence from Cincinnati.* (Peabody Journal of Education, 79(4), 33-53 .2004).

[9] Allen Odden, Eileen Kellor, *How Cincinnati Developed a Knowledge and Skills Based Salary Schedule.* (CPRE. March, 2002).

[10] *Vaughn Elementary's Innovative Teacher Evaluation System: Are Teacher Evaluation Scores Related to Growth in Student Achievement?*,

[11] Los Angeles Unified School District. Vaughn Next Century Learning Center Web site. http://vaughncharter.com/s2/index.php?option=com_content&task=view&id=23&Itemid=41

[12] Milken Foundation. Teacher Advancement Program Foundation Web site. **http://www.tapschools.org/tap/tap.taf?page=tapstates.**

[13] James Rowley and Patricia M. Hart. *High-Performance Mentoring*, (Thousand Oaks, CA: Corwin Press, 1999).

[14] Jennifer Azordegan, Patrick Byrnett, Kelsey Campbell, Josh Greenman and Tricia Coulter, *Diversifying Teacher Compensation* (Denver: Education Commission of the States. December, 2005)

[15] Minnesota Department of Education. Minnesota Department of Education Quality Compensation Web site. **http://children.state.mn.us/mde/Teacher_Support/QComp/Program_Components/index.html**

[16] Denver ProComp We site. **http://www.denverprocomp.org**

[17] Denver ProComp We site. **http://www.denverprocomp.org**

[18] Minnesota Department of Education. Minnesota Department of Education Quality Compensation Web site. **http://children.state.mn.us/mde/Teacher_Support/QComp/Program_Components/index.html**

[19] Minnesota Department of Education Web site. *Minnesota Department of Education; Instructional Observation and Standards-Based Assessment* **http://children.state.mn.us/mdeprod/idcplg?IdcService=GET_FILE&dDocName=008189&RevisionSelectionMethod=latest&Rendition=primary**

[20] Minnesota Department of Education. Minnesota Department of Education Quality Compensation Web site. **http://children.state.mn.us/mde/Teacher_Support/QComp/Program_Components/index.html**

National Conference of State Legislatures

LEGISBRIEF

BRIEFING PAPERS ON THE IMPORTANT ISSUES OF THE DAY

OCTOBER 2009

VOL. 17, NO. 38

Connecting Student-Teacher Data

By Michelle Exstrom

Linking student achievement and teaching quality has many uses.

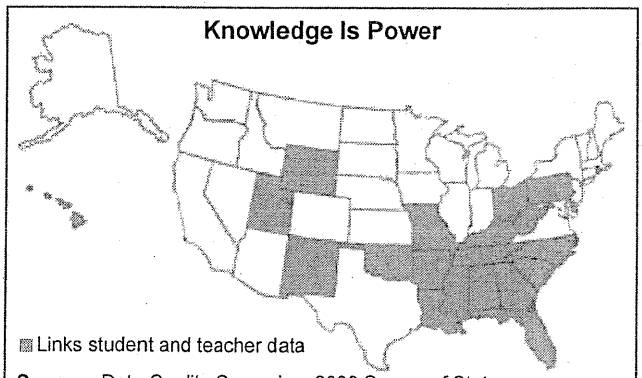
To make informed decisions and craft state laws to improve student achievement and teaching quality, policymakers need access to relevant data connecting the two. State lawmakers could use this information to measure whether teachers are effectively improving student achievement and by how much, identify tools to improve instruction, and measure the success of the institutions that prepare teachers.

During the past decade, states have made notable strides to build statewide longitudinal databases that track students across districts through their academic careers. This can be attributed, in part, to the federal No Child Left Behind requirements for collecting, aggregating and disaggregating such data.

A teacher-student match is one component of a statewide data system.

Progress in matching student data to teachers, however, has been slower. Fewer than half the states have such a comprehensive database, according to the Data Quality Campaign. The campaign is a national collaborative to encourage and support state policymakers as they improve the availability and use of high-quality education data. Among the 10 essential elements of an effective longitudinal data system identified by the campaign is creating a statewide teacher identifier with a teacher-student match.

Knowledge Is Power



Links student and teacher data

Source: Data Quality Campaign, 2008 Survey of States.

Data security and student-teacher privacy are concerns.

Those in favor of linking student and teacher data argue that state policymakers will gain valuable knowledge by creating sophisticated, comprehensive systems to track where teachers are trained, what subjects and grades they teach, who their students are, and how their students perform. Ultimately, researchers can use the data to provide insight to policymakers, identify areas for improvement, and evaluate the effectiveness of state and district policies.

Others are concerned about data security and the privacy of both students and teachers, and also note the significant financial commitment of developing these databases. Educators are wary that the focus could become punitive rather than on needed system reforms and support for educators.

National Conference of State Legislatures

Executive Director
William T. Pound

Denver
7700 East First Place
Denver, Colorado 80230
Phone (303) 364-7700
www.ncsl.org

Washington, D.C.
444 North Capitol Street, NW, Suite 515
Washington, D.C. 20001
Phone (202) 624-5400

State Action With increased pressure from the Legislature for accountability and improved student achievement, Florida became one of the first states to begin to track student data using a statewide database in the 1980s. Since 2002, Florida houses data from many sources—including student classroom assignments, education facilities, curriculum and instructional staff for the P-20 education system—in the Florida Education Data Warehouse. The single repository allows researchers to study student and teacher longitudinal data from the 1995-1996 school year to the present. Grants in 2006 and 2009 ($1.5 million and $2.4 million, respectively) from the Institute of Education Sciences at the U.S. Department of Education allowed Florida to continue data system improvements and build a web interface for better reporting capabilities.

Florida has used a statewide database since the 1980s.

In 2005, Louisiana integrated several existing data systems into the Louisiana Education Accountability System, which links students and teachers to classes It allows the state to monitor class sizes, monitor vocational and special education populations for funding purposes, oversee teacher credentials to ensure that the state is meeting NCLB requirements, and track the percentage of students who take core classes. The data can help determine the effect of a particular teacher by tracking the progress of the teacher's students. The data also are being used to track effectiveness of state teacher preparation programs. A $4 million grant from the Institute of Education Sciences in 2009 will help Louisiana further centralize student and teacher data that can more easily generate reports on demand and track homeless populations and critical student performance measures.

Tennessee's data system, the Tennessee Value-Added Assessment System, was first created by Dr. William Sanders as a result of 1992 legislation. Researchers can obtain information from the system to show the value a particular teacher adds to student achievement. Dr. Sanders' use of the data system resulted in ground-breaking discoveries about the importance of effective, high-quality teaching. The system can predict how well a student should perform and provides information to teachers and principals about the teacher's strengths and weaknesses. It is not used for teacher evaluation, however. Teachers can use the information to better tailor instruction, and principals can use it to better match students with teachers. A $3.2 million grant in 2006 from the Institute of Education Sciences allowed Tennessee to continue system improvements.

Federal Action The American Recovery and Reinvestment Act of 2009 creates several opportunities for states to receive funding to improve state data systems. Funds are provided through the State Fiscal Stabilization Fund; a $245 million grant program for state data systems distributed by the Institute of Education Sciences; and a discretionary fund distributed by the secretary of education in a national competition, Race to the Top. The Stabilization Fund and Race to the Top require states to develop and use statewide longitudinal data systems to improve teacher effectiveness and the equitable distribution of effective teachers.

Federal funds can help states implement statewide longitudinal data systems.

In December 2009 and June 2010, states can submit applications to compete for the $4.35 billion Race to the Top Fund. Proposed priorities and guidance for the program indicate that states will be disqualified for not linking data on student achievement or student growth to teachers and principals for teacher and principal evaluation.

Contacts for More Information

Michelle Exstrom
NCSL—Denver
(303) 364-7700, ext. 1564
michelle.exstrom@ncsl.org

The Data Quality Campaign
www.dataqualitycampaign.org

Education Provisions of the
American Recovery and Reinvestment Act
www.ed.gov/policy/gen/leg/recovery/index.html

PLEASE NOTE: Legislative Information ***cannot*** perform research, provide legal advice, or interpret Maine law. For legal assistance, please contact a qualified attorney.

Resolve, To Encourage Alternative Compensation Models for Teachers and School Administrators

Sec. 1 Department of Education; evaluation of alternative compensation models for educators. Resolved: That the Department of Education shall review alternative compensation models established in other states and governmental or educational entities that allow for performance-based compensation, including bonuses for teachers and school administrators and the bases upon which the alternative compensation is determined; and be it further

Sec. 2 Department of Education; application for federal funds. Resolved: That the Department of Education shall review the requirements of the federal Teacher Incentive Fund program and any other federal grant program under which funds may be used for establishing alternative compensation models for educators. The department shall prepare and submit an application for federal grant funds from the federal Teacher Incentive Fund and any other applicable federal program to develop a state-based alternative compensation grant program for school administrative units; and be it further

Sec. 3 Department of Education; alternative compensation grant program. Resolved: That the Department of Education shall establish an application process whereby school administrative units may apply to participate in the alternative compensation grant program under section 2, referred to in this section as "the grant program." Interested school administrative units must agree to abide by the requirements of the federal grant programs in order to be considered for the grant program. The department shall develop requirements for use of grant program funds, including reasonable timelines for the development and implementation of alternative compensation models and for school administrative units to report progress. To the extent that federal funding requirements allow, the grant program funds may include funding for the department to administer the grant program, to provide technical assistance to school administrative units and to pay for an independent evaluation of the alternative compensation models that are developed. School administrative units must be encouraged to experiment with any number of alternative compensation models. Any alternative compensation plans developed by a school administrative unit must be approved by the participating local bargaining units consistent with the Maine Revised Statutes, Title 26, chapter 9-A and related rules pertaining to collective bargaining for teachers employed by school administrative units; and be it further

Sec. 4 Grant program evaluation. Resolved: That the Department of Education shall submit annual reports to the joint standing committee of the Legislature having jurisdiction over education and cultural affairs by January 15, 2011 and January 15, 2012 describing the progress of the school administrative units participating in the alternative compensation grant program under section 2 and report the results of any independent analysis conducted on the effects of alternative compensation systems, including but not limited to student outcomes, teacher recruitment and retention. The department shall seek outside funding and technical support for use in the development, implementation and evaluation of any alternative compensation models developed through the alternative compensation grant program.