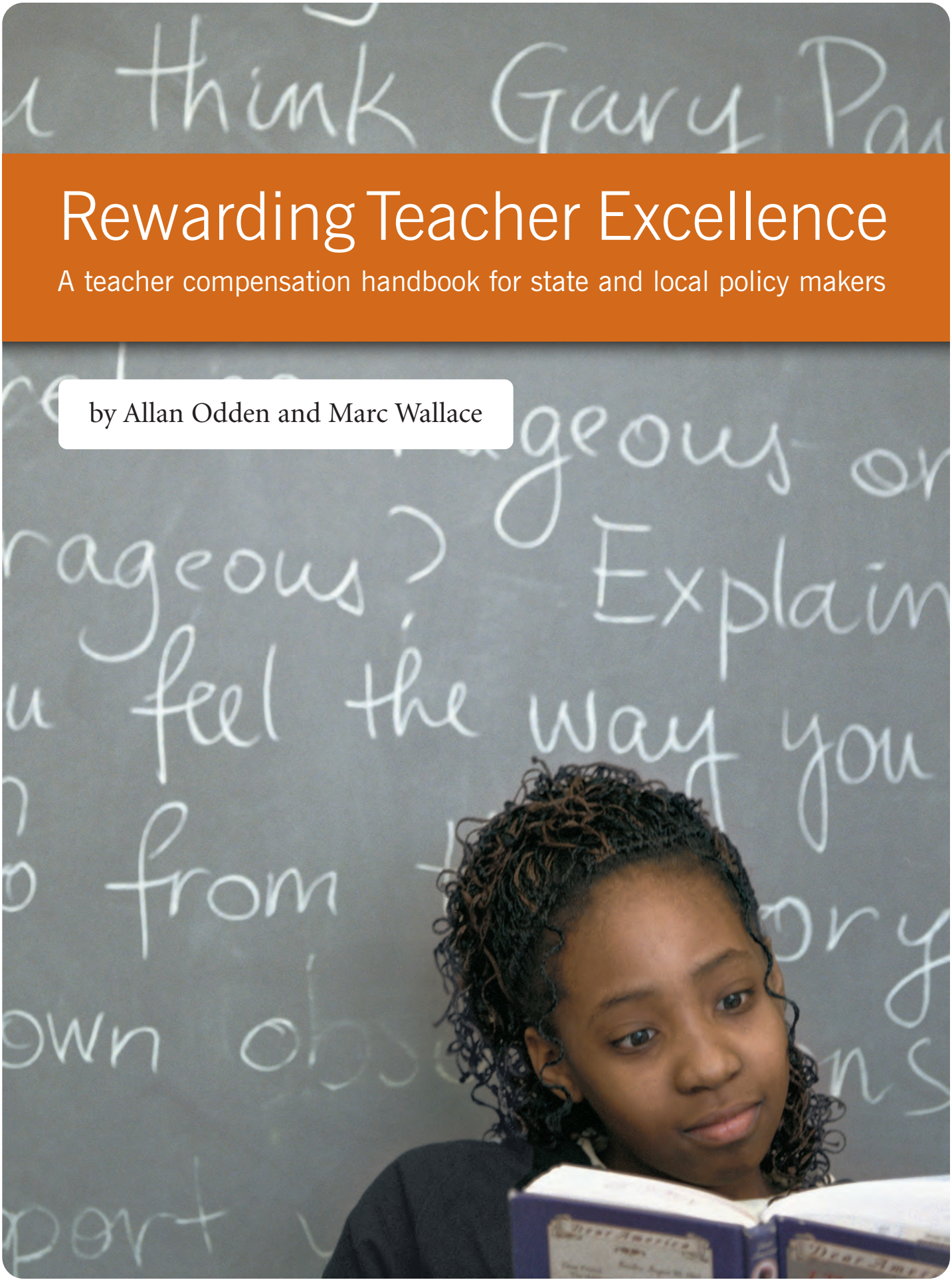
Rewarding Teacher Excellence

A teacher compensation handbook for state and local policy makers

by Allan Odden and Marc Wallace



The quest to change

the way teachers are paid is not something new to this first decade of the 21st century. During the last half of the 20th century teacher pay structure changes were tried at least once a decade. Nearly all of the 1980s Nation at Risk reform recommendations included proposals for changing teacher pay structures to some sort of merit- or performance- basis, but most failed and only a handful survived more than a few years. Today, the single salary schedule that pays teachers on the basis of their years of experience and numbers of education credits and degrees is the schedule used by the vast bulk of states and districts across the country. However, in 2007, there are a few reasons for cautious optimism about teacher pay improvements: (1) the proliferation of new school-level teacher pay models, (2) the emergence of several new state and district-level programs, and (3) the creation of a new $100 million federal Teacher Incentive Fund.

REWARDING TEACHER EXCELLENCE
A teacher compensation handbook for state and local policymakers
Allan Odden and Marc Wallace

The developement of this handbook was supported in part by a generous grant from the Joyce Foundation.

Madison, Wisconsin
Consortium for Policy Research in Education
Wisconsin Center for Education Research
University of Wisconsin

February 2007

Contents

Allan Odden PROFESSOR AND CO-DIRECTOR
Consortium for Policy Research in Education (CPRE)

Allan Odden is Professor of Educational Leadership and Policy Analysis at the University of Wisconsin-Madison. He also is Co-Director of the Consortium for Policy Research in Education (CPRE), a national center studying how to improve state and local education policy. He formerly was professor of education policy and administration at the University of Southern California (1984-1993) and Director of Policy Analysis for California Education (PACE), and from 1975-1984 held various positions at the Education Commission of the States. He was president of the American Educational Finance Association in 1979-80 and received AEFA's distinguished Service Award in 1998. His research and policy emphases include school finance redesign and adequacy, effective resource allocation in schools, the costs of instructional improvement, and teacher compensation. Dr. Odden has written widely, publishing over 200 journal articles, book chapters, and research reports, and 30 books and monographs. His newest book, co-authored with Marc Wallace, is entitled How to Create World Class Teacher Compensation (St. Paul: Freeload Press).

Marc J. Wallace, Jr. FOUNDING PARTNER
Teacher Excellence through Compensation

Marc J. Wallace, Jr. is a founding partner of Teacher Excellence through Compensation (TEC). He is based in Lake Bluff, IL where he serves clients by bringing to bear his extensive consulting and research experience in compensation, workforce effectiveness, human resource strategy and labor relations. Prior to founding TEC he served as President of the Center for Workforce Effectiveness (CWE) in Northbrook, IL from 1992 to 2002 serving clients in both the private and public sector.

Dr. Wallace (with Allan Odden) has pioneered the successful adaptation of proven compensation strategies and techniques in the field of education. His work has resulted in the successful introduction of teacher pay for performance in a variety of schools, school districts and states. Also, Dr. Wallace has co-authored over 60 articles and ten books on management and human resources.

Dr. Wallace holds a Ph.D. and a Master's degree in Industrial Relations from the University of Minnesota. He holds a Bachelor's degree from Cornell University.

Introduction

As state and local policymakers hear more about ways to pay teachers differently, they could use answers to several key teacher compensation questions. This handbook has been developed to ask and answer many of those questions. Each section is organized by key questions that policymakers should ask themselves as they consider new ways to pay teachers, and attempts to summarize the relevant teacher compensation research and highlight existing programs.

Why has the single salary schedule remained for so long?

There are several reasons for the resilience of the single salary structure. First, it is easy to administer. Second, it is objective and not subject to administrative whim. Third, at least when initially adopted in the 1920s and few teachers had a full college education, it encouraged all teachers to obtain at least a bachelors degree, and over time even a Master's degree. Fourth, it was created to eliminate the inequities of paying women less than men, minorities less than non-minorities and elementary teachers less than secondary teachers. These features were prime goals for the single salary schedule when it was created in the early decades of the 20th century (Odden & Kelley, 2002) and the schedule has implemented these goals for nearly 100 years.

Is the time right for teacher compensation reform?

While there is some agreement that the single salary schedule may not be the most effective way to pay teachers, there is less agreement on what solutions or alternatives should look like. However, that may be starting to change, since there is a confluence of factors that have elevated teacher compensation to the top of policymakers' agendas in 2007.

First, proposals to change the way teachers are paid are more popular than ever in both political and policymaker conversations. Indeed, 20 governors mentioned teacher compensation reform in their 2005 state of the state speeches, and nine proposed specific pay for performance plans.

Second, recent national commissions on teaching have highlighted teacher compensation reform efforts. The 2004 Teaching Commission report included teacher compensation reform as one of its three primary strategies, and the National Commission on Teaching for America's Future has also recently been promoting pay reform as a major strategy.

Third, there is more school-level experimentation with new compensation models than ever before. Specifically, the Teacher Advancement Program (funded by the Milken Family Foundation) has been adopted by more than 100 schools in 10 states.

Fourth, more states and districts are experimenting with new pay programs as well. State-level experimentation is relatively new, but programs in MN, AZ, and IA will provide interesting lessons for other states that want to move forward. District efforts are prominent in Denver (which passed a $25 million levy to support its new program), Mobile, Toledo, Columbus, Charlotte-Mecklenburg, Minneapolis, Dallas and are being closely monitored by other districts considering similar programs.

Finally, a new Teacher Incentive Fund program being run out of the U.S. Department of Education will allow interested states and districts to compete for $100 million in funding to reward teachers who produce improvements in student performance and on other measures.

In addition to these recent developments, new ideas about paying people differently have been gaining momentum over the last 25 years. Specifically, two ideas—paying for knowledge and skills rather than experience and including bonuses for improved organizational performance—were tried in the private sector in the 1980s and discussed by our nation's teacher union leaders in the 1990s. This experimentation set the stage for much of the best work in the field today.

Why change the teacher salary schedule?

Pay structure changes should be considered when they reinforce attaining the strategic goals of the organization. When the current, single salary structure for teachers was adopted in the first half of the 20th century, it was done so to provide equity in pay for teachers and encourage many to obtain a bachelors degree—reasonable strategic goals for those times. But those goals have been attained and the current structure no longer reinforces the strategic goals of the education system.

Since changes in teacher pay structures generally follow changes in pay structures in the private sector (Odden & Kelley, 2002), it is useful to understand recent pay structure changes in that arena. When the private sector began paying workers differently in the last two decades of the 20th century, it did so for strategic reasons. Buffeted by pressures to dramatically increase performance, cut costs and move new services and products more quickly into the market place, many private sector companies shifted to paying workers for knowledge and skills, rather than just years of experience, augmented with group-based performance bonuses. The knowledge and skill pay element was designed to encourage workers to acquire and use new expertise needed to work more effectively in restructured work organizations. Performance bonuses, most often for groups of workers rather than individuals, were adopted to keep everyone focused on core organizational goals and to have employees, including front line workers, share in the organization's financial success when performance targets were met and profits rose (Lawler, 2000).

How do teacher pay changes link to core educational goals?

Today education is being buffetted by similar performance pressures. State initiated and designed standards-based reforms seek to dramatically increase the performance of the education system. Indeed, when the data are closely scrutinized, the goal is to double and triple education performance—student achievement—over the next ten to twenty years. The federal No Child Left Behind (NCLB) Act gives further impetus to these goals, requiring both more specific time frames for attaining the goals, and meeting the goals for all students, including students from poverty backgrounds, from families whose primary language is not English, and for students with disabilities.

Moreover, the country has reached a strong consensus that the classroom teacher is the most critical and important educational variable linked to accomplishing these lofty student achievement objectives (e.g., Wright, Horn & Sanders, 1997; Azordegan et al., 2005). Thus to accomplish the goals of dramatically improving the achievement of all students, most agree that the country must:

- Put a high quality teacher in every classroom in the nation
- Insure that each teacher covers the state and district's curriculum standards
- Enhance the actual instructional practices used so that students learn the content in ways that let them think, problem solve and communicate about the content.

What are the new goals and objectives for teacher compensation?

These teacher quality goals nicely lead to clear and specific teacher labor market goals, which can be assimilated into a teacher compensation strategy (see Chapter 2 of Odden & Wallace, 2007), i.e., a set of goals and redesign principles for a new teacher compensation structure.

Such a teacher compensation strategy would have the following goals:

- To recruit and retain high quality teachers overall and specifically for schools in:
  - Large, urban districts
  - Geographically isolated rural districts, and in
  - Content areas experiencing shortages—math, science and technology.
- To enhance the instructional expertise, skills and knowledge of teachers so they know the content they are teaching and can educate students to use the content to solve problems.
- To have teachers focus like a laser on improving student achievement in the core subjects—mathematics, science, reading/English/writing and history—and to expend their energies and talents first on accomplishing these core achievement goals.

These goals in a new teacher compensation strategy can be incorporated into revised teacher salary structures. Appropriate objectives would be as follows:

OBJECTIVE 1. Identify an overall average teacher salary level that would allow the state and each of its districts to compete for needed talent in the broader labor market.

OBJECTIVE 2. Identify the higher salary levels that would be needed to enable large, urban and geographically isolated rural districts to compete for teacher talent; identify similar salary targets for high poverty, and low performing schools; as well as for subject areas where there are teacher shortages.

OBJECTIVE 3. Set clear, ambitious but attainable achievement goals, specifically, targets for improved student performance, on both an individual classroom and schoolwide basis. Set aligned targets for the major student sub-groups: those from lower income backgrounds, those struggling to learn English and those with disabilities but who still should be above to learn the regular, core curriculum. Incorporate these improvement targets into an element of teacher compensation.

OBJECTIVE 4. Identify an instructional vision that can serve as a focus for teacher professional development and a target for enhancing a teacher's instructional practice over time. Create a performance assessment system that measures each teacher's instructional practice to various levels of performance vis a vis the new instructional vision, the results of which can be used to operate a knowledge- and skills-based pay structure.

Setting these objectives is just the first step. The next step is to incorporate these objectives into the appropriate elements of redesigned teacher salary structures so that the new salary structures indeed implement these important objectives. This design challenge requires an understanding of the various elements of teacher compensation structures and how they should be changed to incorporate the objective so the new salary structure not only implements these objectives but also and simultaneously reinforces the education system's goal of doubling and tripling student academic achievement.

As state and local policymakers face the need for and challenges involved in redesigning teacher salary structures, they need to determine which "problems" of the education system they are trying to ameliorate with teacher salary structure changes, and set priorities on the sequence of pay changes they will design and implement.

The rest of this handbook is organized into five questions (related directly to these four objectives) and a last section that focuses on how to design, implement, evaluate, and fund redesigned teacher compensation systems. Each section will be organized around key questions but space is limited. This handbook is based on a larger book (Odden & Wallace, 2007) that you may want to look at for more information on certain topics.

Teacher Compensation Terms

Like most organizations, there are several elements of total compensation for teachers. In the redesign process, it is important to understand the various elements of total compensation for teachers and then to identify which element is the best for implementing each of the above teacher compensation objectives.

To begin, there are five critical elements of total compensation for teachers:

- Base pay
- Variable pay
- Benefits
- Career opportunity
- Working conditions

BASE PAY is the largest element of pay. It is the monthly check. For most workers, including teachers, base pay is the most important pay element and entails the largest amount of compensation dollars. This is not to say that the other elements of compensation are unimportant; it is only to say that base pay is critical to a teacher's decision to work in a school/district and to remain in that instructional position over time.

Base pay has three important components:

- Beginning or starting pay
- Pay progression over time, mainly annual pay increases
- Top pay

Beginning or **starting pay** would be the salary offered to the brand new teacher. Research shows that beginning pay is core and critical for recruiting individuals initially into the teaching profession (Goldhaber & Player, 2005). **Base pay progression** is how a teacher earns or receives a pay increase year after year. Currently, base pay progression is determined largely by years of experience, and education degrees and credits; research shows that neither of these variables is strongly related to student achievement gains. So the pay elements that allow teachers today to earn more pay over time do not reinforce the critical education goals of improving teachers' instructional practice or boosting student achievement.

Most redesigned teacher salary structures today include *different variables that serve as the basis for pay progression*, such as a direct measure of knowledge, skills and instructional expertise. The objective is to use pay elements that are linked to student achievement gains in the mechanisms that determine teacher pay increases over time so that teachers have an incentive (a base pay increase) to change their instructional practice in ways that contribute more to the bottom line of the education system—more student learning.

Top pay is the highest salary a teacher can earn, given the other elements of the salary schedule.

VARIABLE PAY is an additional aspect of pay. Variable pay is almost always provided as a bonus to an employee or a teacher. Variable pay is provided when the individual teacher or all teachers on a faculty produce some predetermined increase in student achievement. It is conditional: if the improved achievement is produced, the variable pay bonus is provided; if not, the bonus is not provided. It is called variable pay because it varies each year depending on whether the performance improvement target is met. Most compensation experts strongly urge organizations, including school systems, to provide variable pay as a bonus—rather than as an increase in base pay. If a performance bonus is added to base pay, then the teacher is rewarded for producing that one-time, annual performance improvement every year thereafter. If provided as an annual bonus, that pay element, then, must be re-earned each year.

BENEFITS, including health and retirement benefits, are also very important, as are **CAREER OPPORTUNITIES** and **WORKING CONDITIONS**. And structuring each of these elements so that they help to accomplish the education system's strategic goals takes resources. But analyzing teacher benefit structures and working conditions is beyond the scope of this handbook. It should be noted that on average, teacher benefits are better than those for the average worker but similar to what large corporations provide; most teachers have adequate health insurance and pension benefits (Allegretto, Corcoran & Mishel, 2004)

WORKING CONDITIONS are also important, especially for teachers. Research shows that most teachers require *both* a salary premium *and* good working conditions—strong principal leadership, decent class sizes and appropriate instructional materials—in order to be attracted to and remain in urban schools and classrooms. Chattanooga and Charlotte-Mecklenburg understood this reality and addressed both issues. We do not want to under estimate the importance of good working conditions, career opportunities and the benefit packages for teachers. However, the handbook focuses on redesigning cash compensation for teachers—base pay, or the monthly check, and variable pay, or bonuses for improving student academic performance.

How can states and districts determine adequate beginning and average teacher pay levels?

An important step in restructuring teacher salary schedules is to identify the salary benchmarks that would allow the state and each group of local districts to compete for quality teacher talent in the broad labor market. This first requires states and districts to identify salary benchmarks for both beginning and average salaries. Paying adequate overall salaries is critical to each district's being able to recruit and retain a quality teacher in every classroom, despite the school's location, education challenges it presents, or subjects that must be taught.

Beginning salaries should be set at a level that allows the education community to compete for beginning teaching talent from among new college "graduates. Beginning salary is one of the most important factors in any individual's decision to choose a profession; of course, other factors, such as interests and skills, also matter. But a low beginning salary can deter many individuals from choosing to enter a profession, especially teaching.

Average salaries are critical for retention; average teacher salaries should be set at some competitive level to enable the education system to retain teachers that have entered teaching and at some point decide whether or not they want to remain in the profession.

What labor markets should be analyzed to identify competitive pay levels?

Two types of labor markets can be used to determine appropriate beginning and average salaries. The first is the education labor market and the second is the broader, total labor market.

COMPETITIVE IN THE EDUCATION LABOR MARKET. Many states identify surrounding states as their education competitors for teachers, and set a benchmark with respect to those states. Another approach is to identify a region of states, such as the Southern Regional Education Board states. A third approach is to identify the national average teacher salary as the education benchmark.

Most local districts also can identify other districts with which they compete for teacher talent; usually but not always, it is surrounding districts. Steamboat Springs, a resort community, identifies other resort communities in Colorado (e.g., Vail, Breckenridge, Telluride) as its primary competitors.

COMPETITIVE IN THE BROADER LABOR MARKET. A second labor market, which is taking more and more teaching talent away from education, is the broader labor market that includes all government as well as private sector jobs. As the U.S. economy has shifted into a knowledge and service economy, many of the jobs in this new economy require knowledge, skills and competencies quite similar to those for teachers. So education competes for teacher talent in the broader labor market as well. The competition is even stiffer for math, science and technology teachers, an issue discussed in the next section.

The U.S. Bureau of Labor Statistics has created a data base, called O*NET, that allows states and regions within states to identify the salary levels of private sector jobs that compete with public education and to which teacher salaries can be benchmarked. In Arkansas, we conducted a factor analysis of the skills and responsibilities of teaching and other job categories and found the following to be similar to public school teaching: 1) college teaching; 2) adult and vocational education teachers; 3) jobs similar to teaching such as training specialists, extension agents, and librarians; 4) social service workers such social workers, counselors and psychologists; and, 5) health care occupations such as registered nurses, physical therapists, audiologist, nutritionists and pharmacists. As noted below, the occupations that compete for math and science teachers include accountants, engineers, computer and information technology workers, professional level sales, and science (e.g., biologists, biochemists, chemists, conservation scientists, foresters, atmospheric and space scientists, etc.).

If the O*NET data set does not provide sufficient data for beginning salaries, the National Association of Colleges and Employers (NACE) tracks entry salary offers in a variety of occupations within each of the fifty states, for graduates of select universities. In a 2003 school finance adequacy study, Arkansas reviewed starting salaries for the following occupations and compared them to the typical teacher—agriculture and natural resources, health and related occupations, and humanities and social science, and the following occupations for math and science teachers—business, computer technology, engineering, and science.

What does "market competitive" mean?

States and districts need to define what they mean by "market competitive." One approach is to be "at the market," i.e., in the middle (average, mean or median) of the market; another approach is to be "leading the market," i.e., someplace in the top half, quarter or even the top jurisdiction in the market. Of course, the latter is more expensive, but if a state or district now leads a market, it might want to remain a market leader.

What are typical state or local salary benchmarks?

Each state and district, or region within a state, should look at teacher salaries with respect to these other occupations, and set benchmarks for average and beginning salaries.

BEGINNING SALARY BENCHMARKS. The most difficult decision for most states and districts will be to determine where in the market it wants to compete. In years past, a market benchmark has been to have beginning salaries set at least at the average of college graduates with liberal arts degrees; in 2005 both were about $31,000. Given the changing nature of the economy and the fact that the vast bulk of jobs have service and knowledge work embodied in them to some degree, a more appropriate target today might be the average salary for all college graduates, which is closer to $40,000. In early 2006, Arizona Governor Janet Napolitano and Maine Governor John Baldacci, both states with low teacher salaries, proposed raising beginning salaries to $30,000.

AVERAGE SALARY BENCHMARKS. During its 2003 school finance adequacy study, Arkansas set its teacher salary benchmark at the average of the Southern Regional Education Board (SREB) states. In 2006, Iowa Governor Tom Vilsak

and Virginia Governor Timothy Kaine proposed to bring their teacher's salaries up to the national average, and Kentucky Governor Ernie Fletcher proposed a hike to the average of its surrounding states.

Should teacher salaries be "adjusted" because of their "shorter" work day and year?

A final issue in making salary comparisons is whether teacher salaries should be "adjusted" to account for the fact that the typical teacher "works" only 9 or 10 months of the year, or even just 5-6 hours a day. This is a hotly debated issue within the education policy community, with many arguing for an adjustment and others arguing just as vociferously for no adjustment. But as Allegretto et al. (2004) learned from the U.S. Bureau of Labor Statistics (BLS), such an adjustment is not warranted because it is difficult to determine how many hours or even weeks that teachers work. Teachers prepare lessons and correct papers outside of the regular school day and often engage in training or curriculum development over the summer months. In comparing salaries among professions, the BLS makes no adjustments when "work" hours are difficult to determine, such as the number of hours airplane pilots work, or college professors work, and suggest that salary comparisons for such jobs, including teachers, be made on an annual salary basis. We agree with that position.

What special circumstances require wage premiums above and beyond the general teacher salary schedule and how large should those premiums be?

Paying teachers with the same characteristics the same salary has been a salary "equity" goal in education for almost a century; this type of pay equity has also been a strong value in the private sector. But when special circumstances arise that lead to substantial teacher shortages, then pay equity for all must be put aside.

Should there be wage premiums for high poverty/low performing schools?

Yes. It is growing common knowledge that many schools with large percentages of students from lower income or poverty backgrounds and/or have students performing at low levels on tests of academic achievement have large numbers of un- or under-qualified teachers (Lankford, Loeb & Wyckoff, 2002). This is one and probably the most important reason why students in those classrooms achieve at unacceptably low levels. It is hard for students to learn if their teacher does not know the subject very well, cannot manage the classroom, and lacks important pedagogical expertise. Thus getting knowledgeable and expert teachers in those classrooms is critical.

Paying a wage premium for teachers in these schools will be required in order for each classroom to be staffed by a knowledgeable and quality teacher. Since the largest percentage of students not performing at an acceptable level attend schools with high concentrations of poverty, creating wage premiums for teachers in these schools should have a high priority on each state's teacher compensation change agenda.

Are wage premiums for subject area shortages needed?

Again, the answer is yes. Many districts across the country, again particularly large urban districts, have difficulty recruiting teachers in subjects such as mathematics, science, technology and sometimes special education and teachers for ELL students (students who come from home where English is not the primary language and must learn English as well as content). When shortages emerge, one part of the solution is a wage premium.

To tackle this issue, states and districts need to define how to specify a subject as a "shortage" area. Many states and districts have engaged in this task, and the handbook does not provide guidelines for that task.

How large should these wage premiums be?

Of course, the key pay issue is to identify how large the wage premium should be. There is some but not a lot of research on this topic.

Some researchers have suggested that a 40 percent pay premium might be needed for teachers in high poverty, urban schools (Rivkin, Hanushek & Kain, 2004; Imazeki, 2005). Other research suggests that the premium might be in the 15-20 percent range, as compared to the state average. The findings from both imply that it will take more than an $1000-$2000 to recruit and retain teachers in such schools. Wage premiums of at least $5000-6000 are likely needed, over whatever the state average salary benchmarks would be.

Fairfax County (VA) provided a wage premium of 12 percent for teachers in their lowest performing schools and Miami provides a 20 percent premium, each also requiring those teachers to work approximately an hour longer each day.

We fully understand that a higher pay level alone will not solve the problem of recruiting and retaining high quality teachers in high poverty schools. Strong principal leaders, decent class sizes and ample instructional materials and supplies are also needed. But a substantial wage premium is also necessary.

The premiums currently offered across the country for subject area shortages are quite small, ranging from about $1000 to $2000 per year, amounts probably not large enough to make a real difference. Indeed, Clotfelter and Ladd (2005) investigated the impacts of a $1600 mathematics and science premi-

um offered by North Carolina and found the program had quite modest effects.

Milanowski (2003) conducted focus groups in a mid-western research university seeking to determine what undergraduates in engineering, mathematics and science would identify as premiums that would have them consider altering their career paths to enter teaching. Although many students stated that no premium would alter their desire to become an engineer, mathematician or scientist, a significant portion said that if the difference between what they could earn in a technical field and education were reduced by half, they would seriously consider teaching. That difference equated to about $5000 at the time of the study.

Though more analysis needs to be conducted, states and districts would be wise to think in terms of wage premiums for teachers in subject area shortages, particularly math, science and technology, of at least $5000 per teacher per year, and then track the impacts on both recruitment and retention of teachers in those areas to see if the premium is sufficiently high. In 2006, the Governor of Mississippi and in 2005 the Governor in Virginia proposed substantial wage premiums for teacher in high need subjects such as mathematics, science, foreign language and special education.

Should these wage premiums be bonuses or permanent wage increases?

In most cases, a wage premium that is a one time annual bonus is not sufficient. Teachers must not only be initially incented to teach in a high poverty school or a subject that is experiencing a shortage but also must decide to stay teaching in that school or subject area. One time bonus payments are not effective strategies for reducing (if not eliminating) these types of teacher shortages.

The reality is that to recruit and retain teachers in high poverty schools as well as "hot" subjects, the premium needs to be "permanent," or at least provided as long as there are shortages of qualified teachers. But even if at some time period there are an adequate number of teachers, the wage premium will probably need to be continued in order for the education system to continue to recruit and retain teachers as natural turnover occurs. At some point, though, the labor market could change and the premium may no

longer be needed, but for education, that will probably be several years, at least for teachers in many high poverty schools or teachers in mathematics, science and related technologies.

Should the wage premium be coupled with an effectiveness criterion as well?

Again the answer is yes, in order to have an effective program. The goal is not just to have "anyone" teach in a high poverty classroom or in a mathematics or science class; the goal is to have someone who knows math or science and is an effective teacher covering those classes. Thus, schools, districts and states should provide wage premiums only to "effective" or "trained" teachers in high poverty schools or subjects where there are shortages. Each teacher qualifying for the wage premium should have to meet an "effectiveness" criterion as well. This is a particularly salient issue for teachers already in education. Requiring a major or a minor or a certain number of academic credits in math or science, or the subject area shortage, could be another criterion for qualifying for the incentive. And for individuals already in education, creating some measure of instructional effectiveness could be a third criterion. If the state or district had a performance assessment of teachers, an issue discussed below, the education system could require a certain score on the performance assessment to qualify for the subject area wage premium; a score of 2 or 3 on a 1-4 (low to high) scale would be a reasonable standard.

Chattanooga used Tennessee's valued-added system for assessing student learning gains to identify those teachers who would qualify for a wage premium for teaching in its high poverty schools. That wage premium was coupled with stronger principals and other school improvement efforts, which combined turned around the student learning in several of its high poverty schools in a short period of time.

Are wage premiums also needed for urban and rural districts?

Again the answer is yes. In addition to schools and subjects that experience teacher shortages and thus warrant a wage premium, higher salary levels are also required generally for teachers in large urban districts and sometimes also in geographically sparse districts. Mississippi's governor proposed a wage premium for teachers working in isolated districts in his 2006 education recommendations. It is hard for large, urban districts to compete for teacher talent in the metropolitan labor market if all the surrounding dis-

tricts pay salaries at a higher level. Most suburban districts offer less challenging education environments; when coupled with higher salaries as well, urban districts are doubly disadvantaged in that metropolitan labor market.

Emerging research is showing that big urban districts—e.g., Chicago, Los Angeles, Milwaukee, New York City, etc.—probably need to pay average teacher salaries that are from 15 to 40 percent above the state average. Although it will be difficult financially and politically to meet the upper levels of that kind of wage premium, hitting the lower levels of the range is very important, or the large urban districts, which enroll the bulk of low performing students across the country, will be forced to staff their schools with under qualified and thus less effective teachers.

How should bonus programs based on improvements in student learning be structured?

Since improving student academic achievement, overall and for various sub-groups of students, is such a preeminent education goal, the time has undoubtedly come for creating a variable pay element for teachers, i.e., a pay element that would provide bonus awards to teachers for improving student achievement. Although this type of pay element can be controversial, and although it is true that factors outside the education system impact student achievement, creating a pay element based on student learning gains is seen by many policymakers—and many within education as well—as critical to any viable and relevant effort to redesign how teachers are paid.

Since the education system has been experimenting with such bonus programs for over a decade, the country has developed extensive knowledge and expertise on how to design and implement effective bonus programs (e.g., Odden, Kellor, Heneman & Milanowski, 1999; Odden & Kelley, 2002). Such programs enhance teacher motivation to improve student performance and, when provided to everyone in a school, do not cause competition among teachers (Kelley, Heneman & Milanowski, 2000).

Kentucky was the first state with such a program; it was based on improvements in the performance of cohorts of students over time (this year's grade 4, 8 or 10th students versus last year's grade 4, 8 or 10th grade students). North Carolina also has operated a statewide program for years, based on a value-added model that provided schools with bonuses if they produced greater than historic annual increases in student performance. Charlotte-Mecklenburg began a bonus program on its own before the state acted, and then modified its program to conform to the state program; both continue to operate today. Dallas has operated a performance bonus program for over a decade.

There are eight questions that must be answered in creating bonus programs that reward teachers for improving student performance:

1. What performance elements will the program include?
2. How will each performance element be measured?
3. How are annual improvement targets set?
4. How can a level playing field be created to provide every school a fair opportunity to produce targeted learning gains,
5. Should the awards be one time bonuses or added to base pay?
6. How large should the bonuses be and should there be multiple bonuses?
7. Should the awards be provided on an individual teacher or whole school basis?
8. What are other important eligibility rules?

1. What performance elements will the program include?

Selecting the performance elements to include in a variable pay, bonus plan, is critical and necessary. The elements that are included become the most important performance factors for the education system and school. The idea is to have teachers focus their professional expertise and energies first on the elements in the performance measure. As such, selecting the performance factors is a value judgment.

It could take up to a year to talk through all the issues related to selecting the key performance factors. At the end of the process, the goal would be to have the large bulk of teachers, administrators and policymakers agree that the factors selected are the most important education goals.

Most such programs across the country ultimately identify student academic achievement in the core academic subjects—mathematics, science, reading and writing, and sometimes history—as the core performance factors. Usually these factors comprise 50-100 percent of the set of performance measures, averaging 80 percent. Often, non-academic factors are also included, such as the four-year graduation rate, the percentage of high school students taking advanced courses, a decline in the percentage of elementary students scoring at or below the basic level, or the percentage of 8th grade students taking algebra. The assumption is that these other performance factors are means to accomplishing the achievement goals.

Today, most programs have some measure of Adequate Yearly Progress (AYP) as this is an important NCLB accountability factor and it also is a measure of closing the achievement gap, or insuring that learning gains are produced by students from poverty backgrounds, ELL students, and students with disabilities. Since AYP is such a stringent performance improvement goal, however, it often comprises only one of many performance factors in a variable pay plan.

2. How will each performance element be measured?

Although there is great debate on how best to measure student academic achievement for a bonus program, most states and districts use some combination of state and local tests that are currently being administered to measure student academic achievement. To be sure, states and districts are strongly advised to use the "best" tests possible, including tests with performance items. But because the tests states and districts currently use ipso facto are the key measures of student performance, they become the default measures for nearly all variable pay programs.

The most critical criterion for any performance measure is that it is valid and reliable. Most standardized achievement tests, created either by private companies or state departments of education, meet these psychometric requirements. Few district or teacher developed tests meet validity and reliability criteria because developing valid and reliable tests is a complex and expensive undertaking, and few local districts devote either the resources or time to develop tests that meet these psychometric standards.

Thus, most state and local variable pay plans will be based heavily on state and local standardized test scores that are currently in use, or are being put into use as a result of the requirements of NCLB to test students in grades 3-8 and 10 in mathematics, reading and science. We generally concur with the decision to use extant state and local tests of student academic performance, while simultaneously encouraging both entities to use better tests over time.

Any non academic achievement performance factor must also be measured in a reliable and valid manner in order for it to be used fairly in a variable pay program. So if a graduation rate is used, it must be defined carefully and collected the same way from every school and district.

Some performance factors, considered important by large numbers of people, are often left out of variable pay programs because they cannot be measured in valid and reliable ways. One district that tried to include a measure of student virtues and responsibilities finally gave that factor up after teachers complained that the factor was not measured fairly across schools. Those who want to enhance student "character" rarely have a good measure of that particular student characteristic. Hard as it might be, it is better to exclude factors that cannot be validly and reliably measured than to include them in variable pay programs.

3. How are annual improvement targets set?

After selecting and measuring the key performance factors, the next step is to set improvement targets. This is a critical step. The general principle is that the improvement targets should be a "stretch" but attainable. The goal is to produce larger annual gains than have been produced in the past, so that student performance begins to rise year after year. But the improvement targets must be attainable or they will not motivate teachers and administrators to change their behaviors in ways to boost system performance.

Another principle is that the improvement targets must be larger than the "measurement error" for the performance measures used. Those who have been engaged in setting AYP targets for NCLB have been embroiled in this issue for the past several years. Those who argue for a "value added" measure of improvement have also been cognizant of the measurement error issue (see for example, McCaffrey et al. 2003), as have variable pay designers that use simpler gain scores or gain to a standard score (Milanowski, 1999). Dealing with measurement error is much more complex when the bonus is targeted to individual teachers. Such programs usually need two if not three years of data to create stable improvement estimates.

Another issue in setting performance improvement targets is the methodology to use. Some argue for simple gain scores: this year's score minus last year's score. As long as gains are set above historical gain movements, this can be an easily understood approach. We recommend states consider the KISS principle in making these decisions: "Keep It Simple Stupid."

Another approach, embedded in the 1990s Kentucky program as well as NCLB, is gain to a standard. The notion here is that over some time period,

the goal is to have all, or more realistically, a large portion (85, 90 or 95%) of students achieving at or above some standard, such as proficiency. Setting the improvement target as increasing the percentage of students at or above the performance standard requires schools currently farther below the target to produce larger annual improvements than schools closer to the target. If the time frame for such improvements is too ambitious, as many claim with NCLB AYP targets, the local response can be frustration and anger rather than resolve to produce the attainable but more than historic average annual improvements.

A third approach is to use a value-added method. In this approach, a current year test score is regressed on the previous year's score, sometimes with demographic data as control variables. The result allows the state or district to calculate an historic "expected gain" score. Schools that produce gains above this "expected" level then meet the performance improvement target. North Carolina has operated a program like this for over 5 years. It does not control for demographic factors under the argument that to do so would inherently require lower levels of performance for students characterized by the controls, such as low income, minority or ELL students. As schools and districts begin to meet these more aggressive improvement targets, the initial regression analysis should be redone, to set an even higher "expected" gain target. Such recalibration should probably be done about every 5 years or so.

Another issue is on what group of students the improvement target is calculated. Kentucky used a cohort approach, e.g., for elementary schools, this year's fourth graders versus last year's fourth graders. By setting the target at a proficiency level, AYP counts mainly the students that achieve over the proficiency bar each year. The value added approach uses data from all students, at least all students that are tested, and aggregates gains over them.

None of these approaches is the single best approach, and different policy analysts argue strongly for the method they prefer. We are somewhat neutral about the particular method used, but as is indicated in the section on pay-out levels, we favor the simpler gain or gain-to-a-standard rather than the value-added approach.

The most important aspect of setting improvement targets is to set them above measurement error, to set them above historical improvement trends, and to make them "stretch" but attainable goals.

It should also be clear that using gains scores is the smartest way to create bonus programs; the goal is to incent improved performance. A program that counted only students already at an achievement level would simply reward the best schools, not those that are improving over time.

4. How can a level playing field be created to provide every school a fair opportunity to produce targeted learning gains?

In both measuring performance and setting improvement targets, adjustments must be made to make them "fair" for all school districts, i.e., to level the playing field so all schools, or all teachers if the award is targeted to individual teachers, have an equal chance of meeting the target.

There are several such issues. The first is setting a minimum percentage of students in the classroom and school that must take the test; most set the percentage in the 95% plus range. A second is whether scores for ELL and students with disabilities should be included. The practice today is to include the scores of ELL students; the major impact of doing so to lower the base year score. The practice also is to include the scores of students with disabilities; we would suggest including only those scores that are test scores, under the assumption that performance measures for students with multiple disabilities are more behavioral in nature and cannot be equated to academic achievement test scores. A third issue is how to include the scores of students who do not attend school for the full school year; no single method of addressing this issue has emerged across the country. So the exact adjustment is determined by the jurisdiction creating the program. The general solution is to include the scores of students that have attended school for some minimum number of days, like 90 or more.

There may be other adjustments that are particular to a specific state or district. For example, Cincinnati included a magnet school "drop back" rate, to neutralize the alleged practice of magnet schools moving students back to their neighborhood school if they were not performing well in the magnet school.

The key here is to recognize all the relevant factors for "creating a level playing field," address them during the design process, and make appropriate adjustments. This communicates to everyone involved that the issues are legitimate, and that the program has made accommodations to insure fairness across all schools.

Finally, there is a school finance side to this issue of leveling the playing field. If all students, including students from poverty and ELL backgrounds and students with disabilities are expected to learn to the same achievement standards over time, then schools enrolling them need extra resources to help them make the larger gains the system hopes for over time. Thus, states (and the federal government) should ensure that extra funds are provided for students from poverty backgrounds, who are ELL and who have disabilities.

Although most states provide extra resources for such students, as does the federal government, the fact that many of these programs are not fully funded can lead to teacher and school claims that the accountability systems or improvement targets are more ambitious than their resources allow. This has been a particular issue with the strident AYP requirements of NCLB and the lack of full funding of the program. The point of the above paragraph, however, is not to debate the specific extra funding level needed, but to say that an additional aspect of leveling the playing field is to provide teachers and schools with students who need extra help to learn to standards with additional resources so they can provide sufficient amounts of that extra help.

5. Should the awards be one time bonuses or added to base pay?

This question concerns the issue of whether the award for improving performance should be provided as an annual bonus, outside of base pay, or as an amount that is added to base pay. Compensation experts strongly urge all organizations to provide performance awards as one time bonuses. This makes the payment contingent on improving performance each year. If the performance improvement target is met, the bonus is paid; if it is not met, the bonus is not paid. And each year new improvement targets are set. When annual performance awards are added to base pay, the individual is rewarded for the rest of his or her career for that one time, annual performance accomplishment. Organizations that add annual performance awards to base pay find that, over time, their highest paid workers are their oldest workers, not their most productive workers.

6. How large should the bonuses be and should there be multiple bonuses?

A general principle is that the average bonus awards should be at least between 4 and 8 percent of base pay which, at an average teacher salary of $50,000, is from $2000 to $4000 per teacher. Smaller bonuses might not be large enough to garner teacher and administrator attention; we would not be opposed to larger bonuses but note that they are more expensive. Some bonus programs around the country have provided annual bonuses in the range of $15,000 to $25,000; our professional perspective is that those levels are too high.

Another issue is whether there should be a single or multiple levels of bonuses. As is clear below, we suggest multiple levels. Cincinnati and Charlotte-Mecklenburg had two levels of awards, the former with an "almost-made-it" award at half the level of the target award, and the latter with a target and above-target award. We would suggest at least three levels of awards, and the balanced scorecard discussed below has four levels.

CREATING A BALANCED SCORECARD. One of the most straightforward ways to design a bonus pay-out structure is to use a balanced scorecard. The balanced scorecard makes the variable pay structure transparent and easier to explain. An example of such a scorecard is provided in Figure 1 on page 35.

The balanced scorecard succinctly shows everyone in the school:

- The performance objectives in the award program
- The weights given to each objective, showing which ones are weighted the highest
- The improvement targets
- The way the school will earn its percentage of the target award.

The balanced scorecard is a powerful but simple way both to communicate the most valued objectives of the school, and the major features of how the incentive program will work. We recommend this approach, knowing that there are other approaches as well.

7. Should the awards be provided on an individual teacher or whole school basis?

It is important to specify up-front who is eligible for the performance award. In education today, the big debate is whether to design a program for individual teachers, or for groups of teachers, such as all faculty in a school. The balanced scorecard (Fig. 1) was designed for all teachers in a school; it could also be used for individual teachers but measurement error would be more problematic unless multiple years of data were used, such as a rolling three-year average. Though bonus programs based on improvements in student achievement are somewhat controversial in education, programs that provide awards for all teachers (as well as administrators and classified staff) in a school are less controversial than programs that reward individual teachers.

Nevertheless, we would encourage states and districts to consider designing bonus programs for individual teachers as well as groups of teachers. Our reason is political as much as substantive. One aspect of the genius of the new Denver teacher compensation plan is that it includes an element that rewards teachers for the achievement gains of the students in their own, individual classrooms. This element has given the program tremendous political credibility. In fact, many in the media characterize the Denver plan as a compensation plan that finally links pay to the achievement gains of each teachers' individual students. In fact, that element of the program comprises a very small portion of the new dollars in the program (see Chapter 4 of Odden and Wallace, 2007), only about 35 percent of all teachers are eligible for this element (only teachers for whom student gain scores can be calculated over time), and only a small portion of those teachers are expected to earn that compensation bonus. But this program element has really "carried" the program among policymakers and the public. So whatever objections teachers might have to such a compensation element, we urge them to consider the tradeoffs between having such a program element (that will impact only a small number of teachers) and the political support it provides, and their opposition to it. Simply opposing such a compensation element, which seems so "natural" to most individuals outside of education, has the effect of making teachers seem opposed to attaining the prime goal of schools—student achievement.

FIGURE 1

A Balanced Scorecard

1	2	3	4	5	6	7	8	9	10
Objective	Measure	Weight	Prior Year Actual	Below Threshold	Threshold	50% of Target	100% of Target	125% of Target	Percent Accomplishment
				0%	25%	50%	100%	125%	
Achievement	Mathematics	20%	20%		60	62.5	65	66.25	10%
Achievement	Reading/Writing	20%	20%		70	72	74	75	20%
Achievement	Science	20%	20%		50	53	56	57.5	10%
Achievement Gap	Meets AYP	30%	30%	No			Yes		0
Retention Rate	Third Grade Retention	10%	10%		20	18	16	15	5%
Totals		100%	100%						45%

The scorecard shows that a small number of performance objectives and measures are included in the program. The general rule is to have 3-6 objectives. This scorecard has five different performance measures, three focusing on student achievement in core subjects—mathematics, reading/writing and science, one on making AYP under the federal NCLB requirements, and one for reducing grade level retention in third grade, a key grade for having all students be proficient readers. The scorecard shows how AYP can be included in but not be the sole determinant of a state or locally designed school-based incentive program. AYP also is a proxy for closing the achievement gap because meeting AYP requires improvement for all key subgroups of students—ELL students, those from poverty backgrounds, each minority group and females as well as males.

Column three shows that the balanced scorecard does not value all performance goals at the same level. The balanced scorecard allows the organization to assign a specific weight to each performance measure; the weights need to sum to a total of 100%. In this example, the achievement scores in the three core subjects are each weighted equally at 20 percent each. AYP, the proxy for closing the achievement gap, is valued higher and given a 30 percent weight. And the third grade retention rate is valued at 10 percent. These weights are value and policy judgments; different states and districts could decide on weights that are different from these figures.

The scorecard identifies a target performance improvement goal for each measure in column 8, but then, as will be shown, provides partial rewards for partial progress in a school's hitting the target, columns 5-7, as well as "extra credit" for beating the target—column 10. Indeed, column 6 indicates a minimum threshold that will qualify the school for at least some reward for each particular objective. Note that the way this scorecard is designed, maintaining one's level of performance is defined as the threshold amount. That might or might not be appropriate in all schools. But as this scorecard shows, meeting the threshold would earn the school 25% of the award, producing 50% of the target improvement would earn the school 50% of the target, and producing the targeted improvement would earn the full award. Column 11 indicates that beating the target by 25% or more will produce an incentive amount that is 125% of the target award.

To calculate the amount of the award, the percent of target improvement attained (0, 25, 50, 100 or 125%) is determined by the actual score for each objective; this is the column heading for each shaded cell in the example. This percent, then, is multiplied by the weight for that objective, and the result is included in column 10, Percent Accomplishment. So for example, this school scored at 50% of target in mathematics which, given the 20% weight for this objective, produced a Percent Accomplishment for this row of 10%. Likewise, the school scored at 100% of target in reading/writing, producing a 20% accomplishment; at 50% of target in science, producing a 10% accomplishment; did not make AYP so earned 0% for this row; and scored at 50% of target for third grade retention so earned 5% of accomplishment (50% times 10%). The total earned accomplishment is the sum of each row of column 10, which in this case is 45 percent. In other words, for this scorecard, the school earned 45 percent of the target award.

The target award could be set at a percentage of salary or a fixed amount. The example shows both. For the approach that sets the award at a fixed dollar amount of $1250 per certified staff member in the school, the award earned is $1125, which is 45% of $2500. Given this way of setting the award, all certified staff in the school would be provided the same award. This award could also be given to all administrative staff as well, including the principal. In this way, all professionals in the school are focused on the same set of core objectives and rewarded in the same way if the objectives are attained.

Alternately, if the target award were a percentage of base salary, say 6 % of salary, then the award earned would be 2.7 % of salary, which is 45% times 6%. If a teacher's salary were $45,000, then the award amount would be $1215, which is 2.7 % of $45,000. A teacher with a salary of $30,000 would receive the same percentage award, 2.7%, but this would be just $810 for this teacher. The award for a principal with a salary of $80,000 would be $2160.

One approach: Target incentive: $2500
Percent accomplished: 45%
Award: 45% x $2,500 = $1125

A different approach: Target incentive: 6% of base salary
Percent accomplished: 45%
Award: 45% x 6% x base salary
Award at base salary of $45,000 = 2.7% of 45,000 = $1215
Award at base salary of $30,000 = 2.7% of 30,000 = $810

Having made this suggestion, we also know that if teachers strongly oppose such a individual teacher bonus program, as was the case in Colonial, Pennsylvania and seems to be the case in the recently adopted programs in Florida and Houston, imposing such a program can lead to strife and political turmoil, hardly the context for having teachers work hard to boost student learning. Indeed, when such a program was "forced" on teachers in Colonial, the following year was characterized by anger and unsettledness. Furthermore, many teachers who received the bonus checks at the end of the year actually turned them back to the district in protest. The district dropped the program after one year of operation. We can hope similar turmoil does not emerge in Florida and Houston.

8. What are other important eligibility rules?

If a school-wide bonus program is adopted, such as the one reflected by the above balanced scorecard, it might be wise to make everyone in the school eligible for the bonus—teachers, administrators and classified staff. Indeed, Cincinnati ran separate school-based bonus programs for principals and teachers, but then folded them into one program after principals requested a single program. Although Cincinnati provided the same dollar bonus to principals and teachers alike, it also would be appropriate to provide a higher bonus for principals, who are paid at a higher average level, than teachers. The proposal by Alaska's governor this year recommended that teachers and administrators receive bonuses at the same level, and that classified staff also be included, but at half the level of the licensed staff.

It also is a good idea to include classified staff in the school-based bonus program, probably at a lower level than for teachers. In one Kentucky school, where teachers decided who was eligible for the bonus and, in one school, decided to exclude classified staff, the kitchen staff, which for years had always baked donuts for faculty meetings, ceased doing so. Although including classified staff increases the cost of the program, it might be better to include everyone in the school—at least to some degree—than exclude groups who contribute in their own special ways to boosting student learning.

Bonus programs, especially those that are school wide, need clear rules on other eligibility issues. Rules need to cover teachers who go on long term sick leave and on maternity leave. Rules also need to cover individuals who spend

part time at the school, such as for example, "traveling" art and music teachers, or psychologist and social workers, all of whom might spend just 2-3 days a week at the school. Usually, such individuals receive a bonus in proportion to the amount of time they spend weekly or monthly in the school.

There also should be rules on the level of the awards; one approach is to set the award level as a percentage of each eligible individual's salary, so administrators would earn a slightly higher award and classified staff would earn a lower award level as compared to teachers. Most programs have fixed amounts that vary by administrators, teachers and classified staff. Fixing the award numbers beforehand is a wise strategy.

Addressing all these issues before the award program is implemented is important. A state or district would be well served by a committee charged with raising all these nitty-gritty issues, and proposing regulations that would guide the program. Ignoring such issues can lead to disappointment, controversy, and legitimate questioning of the program's fairness. Although there are no "correct" answers to all of these issue, putting them on the table, debating alternative approaches, and deciding on how to proceed makes the rules of the program transparent and reduces, if not eliminates, surprises and discontentment.

How can base pay progression elements reinforce better instruction?

Of course, the major aspect of pay—for teachers or any worker—is starting pay and how pay increases are earned over time. As is well known, pay progression for most teachers today is based on years of experience and education units and degrees. Although these pay elements implemented strategic goals in the past, they do not do so today. Research is pretty clear that after three years, a teacher's years of experience is not linked to student learning gains. Research also shows that neither education units nor degrees are linked to student learning gains, except for credits and courses in mathematics and science for teachers who teach those subjects (Monk, 1994).

Redesigned teacher compensation structures often include different factors that are used to provide base pay increases. Such factors can include:

- A major, minor or master's degree in the subject taught, on the assumption that knowledge of one's subject is important to teaching it well.
- Engagement in professional development, on the assumption that such activities will lead to better instructional practice.
- Years of experience, on the assumption that a state or district might want some element of pay for loyalty for staying in the system
- Implementing a project designed to improve student performance, on the assumption that the effort to design and implement such a program over time should lead to better instructional practice.
- Knowledge and skills, or a score on a performance assessment of teachers.

Indeed, the new Denver pay plan includes all of these pay elements (see Odden & Wallace, 2007) and (www.denverprocomp.org).

Note that only the last option is not based on an assumption that the activity is effective; it actually measures a teacher's instructional performance to a set of teaching standards, which as we show below are validated over time as linked to student learning gains. Although there are many ways that knowledge and skills-based salary structures can be designed, the following two examples use a performance measurement of a teacher's actual instructional practice and can provide the long term vision for how such a program could work (see also Odden and Wallace, 2007, Chapter 4, and Milanowski, 2002)

FIGURE 2

Cincinnati's Proposed 2000-2001 Knowledge and Skills Salary Structure

Teacher Category	Performance Required	Salary Range	Conditions
Apprentice	Entry level with teacher license	$30,000	Teachers who fail to advance to Novice level within two years are terminated
Novice	Must be rated 2 or better in all knowledge and skill categories (on a scale of 1-4), and Pass PRAXIS III assessment	$32,000 – $35,750	Teachers who fail to advance to Career level within two years are terminated
Career	Must be rated 3 or better in all categories	$38,750 – $49,250	No maximum on number of years in category
Advanced	Must be rated 4 in two categories, including instruction	$52,500 – $55,000	No maximum on number of years in category
Accomplished	Must be rated 4 in all categories	$60,000 – $62,000	No maximum on number of years in category

Cincinnati was the first school district in the country to propose a completely new salary schedule, with the score on a performance-based evaluation system the major factor determining the pay category for teachers. The schedule, shown in Figure 2, had five performance categories, from Apprentice to Accomplished. The apprentice category was for new teachers, who were provided intense help by the district's new teacher induction program. To enter category two, Novice, a teacher needed to earn the second tier of licensure,

which was passing the PRAXIS III assessment in Ohio, as well as earn a Basic score on the district's teacher evaluation system, which is based on Charlotte Danielson's Framework for Teaching. Each performance and pay category had 2-5 step increases, but the highest step in each category was lower than the initial step in the next highest category. Teachers' salaries would have been capped at the highest step in each category, unless a periodic performance-evaluation showed that the teacher's instructional expertise had reached the standards of the next performance and pay category, and the teacher was promoted. The district also required teachers to move into the third category by their fifth year of teaching, or lose their job in the district. In addition to this schedule, teachers also could earn salary incentives for certain higher education degrees, but only in their licensure area, for National Board Certification, and for teacher leadership roles.

Subsequent research showed that the evaluation system on which pay increases were based was both reliable across assessors, and valid, i.e., teachers with higher evaluation scores as a group produced more student learning gains in their classrooms (Milanowski, 2004; Milanowski & Kimball, 2005; Milanowski, Kimball & Odden, 2005). Implementation and roll out glitches, however, led to teachers' turning down the schedule in a subsequent district-wide vote, though the district continues to operate the performance evaluation system. One of the prime reasons for dissatisfaction with the pay structure was that teachers were transitioned into the new schedule based on their performance on the new evaluation system not on their actual salaries. As a result, many faced potential salary declines of up to $10,000, a transition policy that engendered strong opposition. We recommend that salary levels be guaranteed in any transition to a new schedule, with the new rules conditioning only future salary increases.

A variation of an ambitious knowledge and skills-based teacher salary schedule is displayed in Figure 3. The three lane model in Figure 3 shows directly that there are still incentives for some degrees. The schedule has four performance categories: Entry, Emerging Career, Career and Master. By design, it has the "look" of a single salary schedule, as there are several rows and three columns. But the proposed schedule represents substantial change. For most states and districts, the proposed schedule would replace a 20+ step and 6+ lane traditional schedule. The fewer number of columns sends the signal that miscellaneous units will no longer be rewarded. The units first have to produce a Masters degree, and then a doctorate degree. Though the number of

rows was reduced, the key aspect of the schedule is that it includes four performance and pay categories that are determined by a teacher's performance on a newly designed performance-based evaluation system.

The schedule could work the following way:

- Pay increases would be large for movement across categories and much smaller for step movements within categories. In the example given, the step increases are just 1.5% while the performance category increases are 10%. The message clearly is that teacher instructional performance is the main way to earn salary increases.

- Teachers would be screened for Entry; this would usually be the preliminary license provided after a post secondary training program, or perhaps some type of alternative training program. During the time period in Entry, teachers would be involved in an intense and focused new teacher induction program.

- Teachers would go through a performance-assessment at the end of year 3, depending on the state or district. Teachers would need to meet the performance standards for the Emerging Career level to move into that category. If that performance level were not met by the end of year 3, the teacher would lose their job in the district. So there is an up or out element, but based on performance.

  In many states that have a two tiered licensure system (Youngs, Odden & Porter, 2003), moving into Emerging Career could coincide with earning the professional licensure, which is usually done through a performance assessment of the individual's instructional practice. This level of performance is equal to an overall score of 2 on the TEC standards and rubrics discussed in the Chapter 5 of Odden and Wallace, 2007, and a 2 for the Connecticut and Indiana systems.

- After earning the Standard License and being in Emerging Career, teachers would continue with ongoing professional development and undergo a periodic performance assessment. Towards the end of year 3 in that category, teachers could request such an assessment and if their performance met the standards for the Career category, they could "jump" to Career, step 1. Their salary would be "capped" at Emerging Career Step 6 if their performance never reached the Career level.

FIGURE 3

A Knowledge and Skills Based Pay Plan

	Step Within Level	BA	MA	MA 60/Doct
Entry	1	$30,663	$31,890	$33,165
	2	$31,123	$32,368	$33,663
	3	$31,590	$32,853	$34,168
Emerging Career	1	$34,749	$36,139	$37,584
	2	$35,270	$36,681	$38,148
	3	$35,799	$37,231	$38,720
	4	$36,336	$37,789	$39,301
	5	$36,881	$38,356	$39,891
	6	$37,434	$38,932	$40,489
Career	1	$41,178	$42,825	$44,538
	2	$41,795	$43,467	$45,206
	3	$42,422	$44,119	$45,884
	4	$43,059	$44,781	$46,572
	5	$43,705	$45,453	$47,271
	6	$44,360	$46,135	$47,980
Master	1	$48,796	$50,748	$52,778
	2	$49,528	$51,509	$53,570
	3	$50,271	$52,282	$54,373
	4	$51,025	$53,066	$55,189
	5	$51,790	$53,862	$56,017
	6	$52,567	$54,670	$56,857

Percent Increase for Step	1.5%
Percent Increase for Skill Level	10.0%
MA, MA60/Doctorate	4.0%

The system could require that teachers meet the Career standard in order to stay in the system, a new tenure standard, if you will. Indeed, if the Professional License is granted after a teacher is working for 2-4 years and meets the standard for emerging career (the time period varies by state), it might make sense to postpone the tenure decision until a later time period.

- Once in Career, teachers would undergo a periodic performance assessment. Towards the end of year 3 in that category, teachers could request such an assessment and if their performance met the standards for the Master category, they could "jump" to Master, step 1. However, the stan-

dards for Master would be rigorous, and although not all teachers would be expected to perform at this level a large percentage should reach that level, and there would be no "quota" for this level.

- Thus, the schedule provides a "fast-track" to the top for the individual who enhances their instructional practice, and "caps" the salary of those teachers who do not.

Districts need a performance assessment system to operate such a salary structure. It is very difficult for districts to design a system from scratch, and there are some failed efforts across the country that went this route. The lesson we draw from these experiences is that adapting performance evaluation systems that are already operating, like the Connecticut system (see http://www.state.ct.us/sde/dtl/t-a/best/handbooks/portfolio_forms2.htm), the TEC system (see Odden and Wallace, 2007, Chapter 5) or the Danielson system, will likely produce an operating performance assessment system that could be used for pay categories, much more quickly than trying to design one from the bottom up (see next section of Handbook).

The salary structure in Figure 3 can be enhanced with additional incentives for:

- Subject areas shortages, such as mathematics and science. Districts should consider incentives in the range of $5000 for such content areas.
- Hard-to-staff, high-poverty, low-performing schools. Again, we would suggest incentives in the $5000+ range.
- Incentives for Certification by the National Board for Professional Teaching Standards. Incentives in the 10-20% range ($4000 to $8000 annually), rather than just a one time bonus, will get the attention of teachers to strive to enhance their practice to the high and rigorous standards set by the National Board

A new salary structure like that depicted in Figure 3 represents one of the most strategic ways to redesign how teachers are paid. It pays for the type of teacher expertise school systems need—powerful and effective instructional strategies, it signals that enhancing one's knowledge and skills is the way to higher pay levels, and it links the highest pay to the most effective teachers. Long term, a structure that resembles that in Figure 3 could be a long term goal for many states and local districts.

Since a schedule such as that depicted in Figure 3 is such a dramatic change from current teacher salary schedules, states and districts likely will need to transition into such a structure over time. A first step would be to create and precise a performance assessment system; Chapter 5 of Odden and Wallace (2007) identifies and describes a system that could be used. As mentioned above, systems developed and used by Connecticut and Indiana, as well as Charlotte Danielson's (1996) framework for teaching could also be used. And the next section addresses this issue in more depth.

In the first several years, the score on the evaluation system could be used to trigger a salary incentive on top of the state or district's single salary structure. This would entail grafting a new element onto the old structure. After the performance evaluation system was up and running well, more dollars could be put into the incentive element of the system; indeed, at some point, all new dollars could be put into the incentive based on the evaluation score. Then a transition to the structure depicted in Figure 3 could occur. In this way, some portion of pay would initially be contingent on the level of a teacher's instructional expertise, then a bit more, and finally, the level of teacher's performance on the performance assessment/evaluation would be the major determinant of a teacher's pay.

Have any states adopted this approach?

Yes, both Iowa and New Mexico. And both linked their pay system to their teacher licensure system. Iowa's program has four levels, each with a minimum salary, and New Mexico's program has three performance and pay categories, each with a minimum salary. Instead of specifying steps within each pay category, like the schedule in Figure 3, the state programs simply required minimum salaries in each of the three categories. But each state has large differences between the minimums in each category, sufficient to allow local districts to provide step increases within the category if they want to do so.

How should teacher performance be measured?

As indicated several times in the preceding section, operating a knowledge and skills-based teacher salary schedule requires a measurement of teachers' knowledge and skills. Both of the discussed schedules required a mechanism to place teachers in the different performance and pay categories. The challenge is to decide what performance measurement or assessment system to adopt.

Should teacher performance be measured by student performance?

The quick answer is probably not. Student performance is a result of what the teacher does, and the above performance bonus programs are based largely, if not solely, on student performance. We are strong supporters of having a key element of teacher compensation be based on student performance, but recommend that it be for an annual bonus, and not linked to base pay progression, or the annual raise.

Base pay is what districts and states spend to "buy" the instructional "assets" of teachers, much like any work organization needs to buy the human and physical assets needed to run the organization. Education systems need teachers, administrators, computers, school buildings, etc. to provide educational services. But states and districts need to buy teacher assets that matter, i.e., that are linked to student learning gains.

In the past, states and districts paid more for those assets when the teacher earned another year of experience or an education degree; but those indicators are not strongly linked to student learning gains. So states and districts now seek to identify other indicators of teachers that are directly linked to

student learning gains. A degree or major in math and science is one such indicator for math and science teachers, so paying for a focused Masters degree makes sense. As noted in the last section, many new indicators that districts have selected to use in new pay systems—engagement in professional development, implementing a student learning project, etc.—are based on the assumption that they do lead to learning gains. Some might question such assumptions. That is why we recommend selecting a measure that can and has been directly linked to student learning gains—a performance evaluation of a teacher's actual instructional practice.

What new pay factor differentiates teachers' effectiveness with students?

Our perspective is that the most powerful new approach for providing annual teacher salary increases are those based on a measure of a teacher's instructional practice, as produced by a performance evaluation or performance assessment system. Such pay for teacher knowledge and skills can enhance teacher quality (e.g., Odden, Kelley, Heneman & Milanowski, 2001) and, as the previous section showed, can be integrated into state efforts to create two-tier licensure systems for teachers (Youngs, Odden & Porter, 2003). Giving teachers a pay incentive for earning National Board Certification is an additional concrete example of paying teachers for knowledge, skills and instructional expertise. And both of these measures have been shown by research to identify teachers with greater impacts on student learning gains, a goal for a revised base pay structure.

How should you select a performance assessment system for teachers?

States have the resources and authority to develop a performance assessment system for teachers, but the task is too complicated for most districts. Thus, we do not recommend that districts create their own performance assessment or evaluation system, nor do we recommend that states encourage them to do so. It is too complex a task. Even for states, the most straight forward approach is to adapt a system that already has been developed.

There are four performance assessment systems that have been created. The BEST system was created and is used by Connecticut, as well as Indiana. Though this system was developed for use to confer the second tier of licensure, the scoring rubrics have four levels of performance, 1-4. A two is

required for full licensure. The system is robust enough to use for all teachers, with a score of 3 or 4 moving teachers into a higher performance and potentially higher pay category.

The second system was created by the National Board for Professional Teaching Standards, but that is a proprietary system and has standards only for accomplished teachers. It is best used as an adjunct to a salary schedule or perhaps as the top tier for a knowledge and skill-based salary structure.

Teaching Excellence through Compensation (TEC) has also developed a performance assessment system, and it is described in Odden and Wallace (2007), Chapter 5. Rather than including a separate set of standards and scoring rubrics for all subjects, the TEC system is a more generic version of the Connecticut system, largely because many districts argued that administering subject-specific standards would be too complex. The highest performance categories in these systems reflect the type of instruction that is effective in having most students learn to higher levels and be able to use the content for problem solving and other analytic purposes.

The fourth system is Charlotte Danielson's (1996) Framework for Teaching, a system popular among school districts. In our opinion, however, the Framework accommodates a variety of instructional visions and thus does not specifically promote a specific instructional vision that research shows is effective with all students.

How should a performance assessment for teachers be selected?

States have the competitive advantage. States have the resources to do this job well, whereas most districts do not. We strongly recommend that each state select, develop and then have each district use a statewide performance assessment system. The resources of the state can be used to adapt one of the above systems for state use. The system would provide common scores to be used to place teachers in a variety of locally designed salary structures, whereas placement with locally developed systems would be much more problematic when teachers moved to a different district. The system could be used to confer the second or professional license. And as importantly, the system could be used to emphasize a common, statewide view of powerful instruction. Indeed, since all the systems have a four level scoring rubric,

a state could easily have four licensure levels, each connected to a performance level in the system: initial or preliminary license, Standard license for level 2, Professional or Career for level 3 and Master for level 4, or some such labeling.

What is in the measurement system?

The Connecticut and TEC systems, as well as the portfolio aspect of National Board Certification, measure a teacher's instructional practice through a specified and fixed portfolio of instructional artifacts. The artifacts are usually part of teaching a 2 week curriculum unit and include:

- A description of the class, and its student characteristics.
- A description of the concept that is the goal of the curriculum unit, and what is known about how students learn that concept.
- All the lesson plans
- All the formal and informal assessments.
- At least two video clips of actual classroom instruction.
- Reflections on how the unit went, and ideas about improvement for the next time it is taught.

The precise nature of each of these elements of the portfolio is described by the different systems and each portfolio element, and the portfolio overall is assessed by trained assessors who use a common set of scoring rubrics. All elements of the portfolio also can be digitized, including the video clips of actual instruction, which facilitates both storage of the portfolio (which can be done electronically on the Web) and scoring, which can be done by trained assessors located anywhere in the state (or country or even the world). There also are specified ways to produce an overall score from the results of each of the sub pieces of the portfolio.

The result is a rich measurement of the instructional practice of teachers, using multiple forms of data, assessed by trained scorers to a set of specified teaching standards and scoring rubrics, the result of which can be sustained as valid and reliable.

If a state helps to develop and operate such a performance assessment/evaluation system, it can help assure common measurement across districts, which provides a stable teacher performance measurement system that can be used in quite different ways by districts in locally designed new teacher salary structures.

How can you eliminate hoops?

Because of the way districts have developed new performance evaluation systems in the past, teachers—and administrators—often see the new system as a set of "hoops" to jump through rather than something integral and important to their work. The Connecticut/Indiana and TEC systems can avoid this unhealthy claim if local and state leaders see the real connections among the evaluation system, professional development and the daily work of teachers.

In a standards-based environment, the prime job of teachers is to teach a series of standards-based curriculum units that systematically over the year cover the district's and state's curriculum content standards. An evaluation system, like the ones described above, with a curriculum-unit instructional portfolio as its foundation, simply requires that one such unit be formally presented as an example of a teacher's instructional practice. Further, if the prime approach to professional development at the state and local level is to help teachers create or hone standards-based curriculum units, which research has shown to be one of the most powerful forms of professional development (Cohen & Hill, 2002), then the evaluation approach, the professional development approach and the daily work of teachers will be completely and fully integrated, all around standards-based curriculum units. Such a clear integration would go a long ways towards diminishing the claim that the evaluation system was something different, onerous and just another hoop. It would make the evaluation system something intertwined with what the teacher did every day in both instruction and professional development.

How are teacher pay changes designed, implemented, evaluated and funded?

The process of designing and implementing teacher salary structure changes is very important. There are four major process steps, and each is important:
1. Design
2. Implementation
3. Evaluation and Change
4. Cost

1. Design

The design phase can be completed in one year. Longer time might be necessary, but more energy and resources should be put into implementation and evaluation than in the design phase.

It is best to include all major persons affected by the pay change in the design process: teachers from elementary, middle and high schools, and from all subjects taught; administrators from elementary, middle and high schools and the central office, and small numbers of others depending on local context.

We would suggest creating a Steering Committee to "guide" the design process and make policy decisions that ultimately arise. We recommend that this committee have 10-15 members. We suggest several subcommittees for each pay element that is being redesigned. For a bonus program based on improvements in student performance, we would suggest four sub-committees with the following missions:

- Identify the performance elements and how they are to be measured
- Identify the procedure (value-added, gain score, gain to a standard, etc.) to calculate performance change or improvement and to set improvement or change targets
- Design the pay-out structure, such as a balanced scorecard.
- Identify pay out levels and estimate costs.

We suggest that each of the sub-committees have 8-12 members, including at least 2 members of the Steering Committee. Each should be chaired by a Steering Committee member. We suggest that the sub-committees and Steering Committees meet monthly, on an every two-week schedule. Sub-committees meet week 1; short minutes would then be placed on a web site for wide communication. In week 3, the Steering Committee meets. Sub-committee chairs would bring back key decisions to the Steering Committee and ask for policy decisions when needed to continue to move forward. Steering Committee minutes and decisions would also be placed on a web site for wide communication. Then the sub-committees would meet in week 5, and so on until the design process is completed.

For creating changes to the base pay schedule, we would suggest creating sub-committees with the following charges:

- Identify an instructional vision and select a set of teaching standards and performance assessment scoring rubrics that reflect this vision (The TEC system, the Connecticut system, the Danielson system, etc.).
- Identify how the new teacher performance assessment system will augment or replace the current teacher evaluation system.
- Design the new KSBP structure, including the transition structures.
- Identify average pay levels, pay incentives for low performing schools, subjects with shortages, and for National Board Certification, and the costs of the structures.

We would recommend the same process for Steering and sub-committee meetings as described above for performance bonus programs. Each state and district will need to decide how much change can be designed in each semester of the school year; it could be that the performance bonus program is designed in the first semester, and changes in base pay are designed in the second semester, or some sort of sequencing of the design process.

2. Implementation

Most states and districts spend all their developmental resources on the design process and give insufficient attention to implementation. But implementation is even more important than design. Few programs fail because of design flaws; most programs that fail—in both the public and private sectors—do so because of implementation problems, most of which can be avoided.

There are several aspects of implementation that are important. First, there should be a pilot of each and every new pay element. The pilot should be operated in sufficient numbers of schools and for sufficient numbers of

teachers so that design glitches and uncertainties can be identified and fixed. The purpose of a pilot is not to determine whether the new pay element is valid and reliable; it is to make sure it operates smoothly and to get feedback from those impacted on problems, lack of clarity, or other aspects that need strengthening. A good pilot usually takes at least one year. Thus, in the best and fastest scenario, implementation cannot begin until year 3, after a year each of design and piloting.

Second, when implementation begins all teachers need to be informed of each new pay element that will be implemented—what it is, how it works, how and when they will be impacted, etc. This is best done through brochures, web sites, and briefings at every school—with teachers required to attend the briefings. All individuals involved in operating mechanisms of any new pay element (such as assessors in a performance evaluation system) need to be well trained, and able to produce reliable scores across teachers and subject areas. And focus groups during the first year of implementation would also be wise, to get immediate input on how implementation is going and to identify any new "glitches" or "issues" that emerge and need to be fixed.

A sub-issue here is to decide if all teachers will be included in the new pay plan, or if teachers can volunteer to be a part, which gives the option to teachers of staying on an old schedule. Although it is difficult to operate two different salary schedules for any organization, having teachers opt into the new system might reduce resistance to the program in the initial years. If the program works as intended, if teachers who opt in like the program and see high pay, their satisfaction could induce more teachers to opt in over time. Choosing to participate in the new program is a key element of the new Denver ProComp program.

3. Evaluation and change

Finally, after at least a year of full implementation, a more formal evaluation of each program element should be inaugurated. For many pay elements, the first step in evaluation would be more formative, and seek teacher attitudes towards the new pay elements; standard research instruments and procedures are available for this stage (e.g., Heneman & Milanowski, 2003). A second step would be to determine the reliability and validity of teacher performance assessment scores; again, instruments and procedures are available for this more complex evaluative task (Milanowski, Kimball and Odden, 2005). A third step would be to redesign the human resources (HR) system for a state and local education agency so that all elements of the system relate to the goals, elements and aspects of the new pay system, including teacher

recruitment, selection, induction, assignment, evaluation and pay over time (Heneman & Milanowski, 2004). A teacher compensation system with a new vision of instruction, a performance evaluation and a new pay structure entails huge change in pieces of the human resources system; if the entire HR system is not restructured to reinforce these elements, the power of the new pay system can be eroded and the HR system and the elements of the new pay system can be "pushing" in different directions. Finally, states and districts should be open to making adjustments to each new pay element each year if glitches, rough edges or problems emerge; it is better to improve the system as implementation occurs over time than to leave problems identified but not resolved.

Of course, the last step in evaluation and assessment is to go back to base one and ask whether the new system is working:

- Is the state or district able to recruit and retain high quality teachers, overall, in urban and rural districts, in subjects where there had been shortages, and in high poverty schools? In other words, are the pay levels sufficiently high (and working conditions sufficiently good)?
- Has sufficient professional development been provided both for teachers to learn and use the instructional strategies embodied in the new vision of instruction embodied in the new teacher evaluation system and do those instructional strategies actually produce large improvements in student academic achievement?
- Has student achievement improved, in sufficient magnitudes, and for all groups of students?
- Are teachers and administrators pleased with the operation of the new systems that undergird the new pay elements (the analyses of student achievement for the performance bonus programs and of the performance evaluations for the knowledge and skills pay elements), and if not, what aspects of those systems need further refinement?
- Are policymakers and the public satisfied with the new pay system and the performance results?

Hopefully, the questions will have a positive answer. If not, the problem could be with the pay system as well as other parts of the state and local education system. Remember, a new pay structure is designed to support and reinforce a state's or district's strategic education goals, but a changed compensation structure alone will not renew an education system. Clear goals are required, good leadership is needed, sufficient new training will be crucial, an aligned HR system is important, and other working conditions impact whether overall goals are met and whether the more specific goals for the

compensation system are met. The point of the above bullets, however, is to say that a restructured compensation system is rationalized on the basis of its helping to attain key organizational goals, so at some point in determining whether the compensation changes have been successful, the issue of organizational goal attainment needs to be raised. Hopefully, the education system will be able to say, after a period of years, that the system's strategic goals have been accomplished—student academic achievement is much higher—and the teacher compensation changes have helped—there is a quality teacher in every classroom, teachers' instructional practice has improved, and teachers see that their individual students are learning more than they used to and know that it is because of their new and more powerful instructional strategies.

4. Cost

There are no magic new sources of funds to finance new approaches to teacher compensation. Raising beginning and average salaries, providing wage premiums for several circumstances and creating performance award programs will cost more money. New state or local appropriations will be needed to provide money for these new teacher salary elements, either through natural growth in general fund revenues or through tax increases.

For example, in 2001 Arizona passed a half penny sales increase, mandating that a major portion of the new revenues could be used only if local districts designed new "performance pay" structures. Unfortunately, there was a wide variety of such programs created, many not rigorously performance oriented, and many state policymakers have been disappointed with the results. Minnesota is providing $260 per pupil for local school districts that design new performance pay programs, that meet certain criteria in a state proposal. We would argue that the criteria are quite broad and will allow rigorous as well as less rigorous programs. Nevertheless, both of these states provided districts with money "up front" to design and implement new forms of teacher compensation.

Denver passed a multi-million dollar increase to fund its ProComp program through a special tax increase referendum.

We would argue that long lasting forms of new compensation are those that reallocate substantial dollars in the current salary structure. If all new pay elements are contingent upon special categorical dollars, the likelihood of their demise in hard fiscal times is high, and local districts and teachers

might not see them as a permanent part of the new pay structure. But if the new pay system is funded with substantial funds already in the system, then teachers might see them as more permanent structures and unlikely to go away when the budget is tight.

Having said that, it is difficult politically to get teachers to agree to reallocate substantial portions of dollars in the current salary schedule into a new one, which is different and makes if more difficult to earn pay increases and sometimes just get pay bonuses. But we would argue that funding new salary systems via salary dollar reallocation is the best route to solidifying the new salary structure in the future. Otherwise what states and districts will create are salary add-ons which, given past history, will be jettisoned when the economy slumps and public dollars drop.

Nevertheless, states and local districts should view the new federal Teacher Incentive Fund program as a strategic opportunity to move foward on the new teacher compensation agenda. We would advise them to structure a program that requires as much reallocation of local salary dollars as possible into any proposed new system to strengthen its chances of remaining in existence over time. We would advise against state and local efforts to use just TIF funds to finance new salary components; we would predict that such salary innovations would be dropped when the TIF funds themselves phase out.

How can the costs of bonus programs be estimated?

Costs of such programs can be estimated with a few assumptions. If the average teacher salary is $50,000, and the target bonus is 6 percent or $3000 per teacher, and it is assumed that half the schools (or schools with half the teachers) qualify for the bonus, then a good guesstimate of the cost is $1500 per teacher. Total costs are $1500 times the number of teachers (licensed individuals paid on the teacher salary schedule) in the system. Clearly costs will vary if the bonus is larger (or smaller), if the average salary is larger (or smaller), if a higher (or lower) percentage of teachers qualify, and if the bonus awards between threshold, target and above target are not evenly distributed, but the above rule can be used to make a decent estimate of costs, given the assumptions. Costs will also be higher if administrators and classified staff in schools are included in the program; those added costs can be estimated using the same above simple formula.

What if program costs exceed budgets?

Teachers are very skeptical about whether states or districts will actually pay out performance awards that have been earned. This was true in the fifth year of the Kentucky program, even though appropriations for the program were made and put in a "lock box" before every academic year began. And indeed, California dropped its program when the budget became tight. There are three options for providing the awards when actual costs exceed estimated or budgeted revenues.

The first, and best approach, is to increase the budget. Indeed, during the first year of North Carolina's program, many more schools qualified than had been assumed. In response, the legislature went back and added tens of millions to the program, and paid the full award to each teacher in every school that qualified, thereby solidifying teacher trust in the program.

A second, and acceptable but second best, approach is to reduce the award levels. This maintains the budget and provides an award to everyone who earned it. But by lowering the award level, it will create some skepticism about the program for the next year.

A third but not recommended approach is, after the fact, to hike the bar for award qualification to a level that keeps the award payments within budget. This keeps the budget intact but changes the rules of the game and "stiffs" teachers in schools that meet initial performance improvement targets but do meet the higher, after the fact, targets. This approach would also likely create more skepticism across the education system, rather than bolster everyone's working to improve student performance.

Whatever approach is taken, it is probably best to identify the policy for this problem at the beginning of each year. Then at least everyone will know what happens when the cost of paying all who qualify for the award exceeds the initial budget. But the best policy is to adjust the budget to what it would take to pay everyone the full award for producing improvement. After all, producing more improvement than expected should be celebrated, even if it costs money, rather than dampened by reducing the award level or not giving anything to those who meet rigorous performance improvement targets.

References

Allegretto, Sylvia A., Sean P. Corcoran, & Larry Mishel. (2004). *How Does Teacher Pay Compare? Methodological Challenges and Answers.* Washington, DC: Economic Policy Institute.

Azordegan, Jennifer, Byrnett, Patrick, Campbell, Kelsey, Greenman, Josh & Coulter, Tricia. (2005). *Diversifying Teacher Compensation.* Issue Paper. Denver, CO: Education Commission of the States.

Clotfelter, Charles, & Helen Ladd. (1996). Recognizing and Rewarding Success in Public Schools. In H. Ladd (Ed.), *Holding School Accountable* (pp. 23-64). Washington, DC: The Brookings Institution.

Cohen, David K., & Heather C. Hill. (2001). *Learning Policy: When State Education Reform Works.* New Haven, CT: Yale University Press.

Danielson, Charlotte. (1996). *Enhancing Professional Practice: A Framework for Teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Goldhaber, Dan. (2002). Teacher Quality and Teacher Pay Structure: What Do We Know and What Are the Options? *Georgetown Public Policy Review,* 7(2), 81-94.

Goldhaber, Dan, & Dan Player. (2005). What Different Benchmarks Suggest About How Financially Attractive it is to Teach in Public Schools. *Journal of Education Finance,* 30(3), 211-230.

Hanushek, W. A., Kain, J. F., & Rivkin, S. G. (2004). The Revolving Door: Factors Effecting Teacher Turnover. In W. J. Fowler (Ed.), *Developments in School Finance: 2003* (pp. 5-16): U.S. Dept of Education, National Center for Education Statistics.

Heneman, Herbert G. III, & Anthony T. Milanowski. (2003). Continuing Assessment of Teacher Reactions to a Standards-Based Teacher Evaluation System. *Journal of Personnel Evaluation in Education,* 17(2), 171-195.

Heneman, Herbert G. III, & Anthony T. Milanowski. (2004). Alignment of Human Resource Practices and Teacher Performance Competency. *Peabody Journal of Education,* 79(4), 108-125.

Imazeki, Jennifer. (2005). Assessing the Use of Econometric Analysis in Estimating the Costs of Meeting Sate Education Accountability Standards: Lessons from Texas. *Peabody Journal of Education,* 80(3), 96.

Kelley, Carolyn, Herbert Heneman III, & Anthony Milanowski. (2000). *School-Based Performance Award Programs, Teacher Motivation, and School Performance: Findings from a Study of Three Programs* (CPRE Research Report Series RR-44). Philadelphia: Graduate School of Education, Consortium for Policy Research in Education.

Lankford, Hamilton, Susanna Loeb & James Wyckoff. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis,* 24(1), 37-62.

Lawler III, E. E. (2000). *Rewarding Excellence: Pay Strategies for the New Economy.* San Francisco:

McCaffrey, Daniel F., J.R. Lockwood, Daniel Koretz, & Laura Hamilton. (2003) *Evaluating Value-Added Models for Teacher Accountability.* Santa Monica, CA: RAND Corporation.

Milanowski, Anthony. (1999). Measurement Error or Meaningful Change? The Consistency of School Achievement in Two School-Based Award Programs. *Journal of Personnel Evaluation in Education.* 12(4), 343-363.

Milanowski, Anthony. (2002). *The Varieties of Knowledge and Skill-Based Pay Design: A Comparison of Seven New Pay Systems for K-12 Teachers* (CPRE Research Report Series RR-050). Philadelphia: University of Pennsylvania, Graduate School of Education, Consortium for Policy Research in Education.

Milanowski, A. T. (2003). An Exploration of the Pay Levels Needed to Attract Students with Mathematics, Science and Technology Skills to a Career in K-12 Teaching. *Education Policy Analysis Archives.* Retrieved May 18, 2004, from the World Wide Web: http://epaa.asu.edu/epaa/v11n50/

Milanowski, Anthony T. (2004). The Criterion-related Validity of the Performance Assessment System in Cincinnati. *Peabody Journal of Education,* 79(4), 33-53.

Milanowski, Anthony T., & Kimball, Steve M. (2005). *The Relationship Between Teacher Expertise and Student Achievement: A Synthesis of Three Years of Data.* Paper presented at the American Educational Research Association Meeting. Montreal, Quebec, Canada.

Milanowski, Anthony T., Steven M. Kimball, & Allan Odden. (2005). Teacher Accountability Measures and Links to Learning. In L. Stiefel & A. E. Schwartz & R. Rubenstein & J. Zabel (Eds.). *Measuring School Performance and Efficiency: Implications for Practice and Research* (pp. 137-161). Larchmont, NY: Eye on Education.

Monk, D. H. (1994). Subject Area Preparation of Secondary Math and Science Teachers and Student Achievement. *Economics of Education Review,* 13(2), 125-145.

Murnane, Richard, & David Cohen. (1986). Merit Pay and the Evaluation Problem: Why Some Merit Pay Plans Fail and a Few Survive. *Harvard Educational Review,* 56(1), 1-17.

National Commission on Excellence in Education. (1983). *A Nation At-Risk: The Imperative for Educational Reform.* Washington, DC: U.S. Department of Education.

Odden, Allan & Carolyn Kelley. (2002). *Paying Teachers for What They Know and Can Do: New and Smarter Compensation Strategies to Improve Student Achievement.* Thousand Oaks, CA: Corwin Press

Odden, Allan, Eileen Kellor, Herbert Heneman, & Anthony Milanowski. (1999). *School-Based Performance Award Programs: Design and Administration Issues Synthesized from Eight Programs.* University of Wisconsin-Madison, Wisconsin Center for Education Research, Consortium for Policy Research in Education. Retrieved, from the World Wide Web: http://www.wcer.wisc.edu/cpre/tcomp

Odden, Allan, Carolyn Kelley, Herbert G. Heneman III, & Anthony Milanowski. (2001). *Enhancing Teacher Quality through Knowledge and Skills-Based Pay* (CPRE Policy Brief RB-34). Philadelphia: University of Pennsylvania, Graduate School of Education, Consortium for Policy Research in Education.

Odden, A. & Wallace, M.C. (2007). *Ho to Achieve World Class Teacher Compensation.* St. Paul: Freeload Press. Download at www.freeloadpress.com.

Perie, Marianne, Moran, Rebecca, Lutkus, Anthony D., & Tirre, William. (2005). *NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics* (NCES 2005-464). Washington, DC: US Department of Education, National Center for Education Statistics.

Wright, S.P., S.P. Horn, & W.L. Sanders. (1997). Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. *Journal of Personnel Evaluation in Education,* 11(1), 57-67.

Youngs, Peter, Allan Odden, & Andrew Porter (2003). State Policy Related to Teacher Licensure. *Educational Policy,* 17(2), 217-236.