



The Other 69 Percent: Designing a Performance-Based Pay System for Teachers of Non-Tested Subjects and Grades, English Language Learners, and Students with Disabilities

Cynthia D. Prince
Vanderbilt University
2008 TIF Grantee Meeting
June 5-6, 2008

Purpose of this session

To discuss challenges involved in designing appropriate measures of performance/productivity to identify and reward effective teachers of:

1. non-tested subjects (e.g., art, music, physical education, foreign languages);
2. non-tested grades (particularly high school and pre-kindergarten to Grade 2);
3. English language learners; and
4. students with disabilities.

A few caveats

1. Teacher opinion on this subject is far from uniform.
2. None of the potential solutions that will be presented is perfect. All have both advantages and disadvantages and will require trade-offs.
3. However, early attempts and the lessons learned from them are certainly worth considering when attempting to design a performance-pay program that includes all teachers.

Different points of view on what is fair and how different teachers should be rewarded

1. Teacher rewards should be aligned with the most important instructional priorities of the school (generally, student learning gains in the core subjects).

vs.

2. Incentive systems should encourage every teacher to excel in his or her particular subject, not just those who teach reading and math.

CECR activities to find workable solutions to these challenges

1. Established an advisory group of researchers and practitioners with expertise in teacher compensation, measurement, education policy, special ed, and ELLs.
2. Convened the Working Group in May 2008 in Washington, DC, to provide guidance and recommendations to CECR.
3. We are now developing a module for the CECR guidebook on this topic that will incorporate the Working Group's ideas and input.

Members of the Working Group

- Chris Barr, Columbus (OH) City Schools
- Ardrej Beijulin, Eagle County (CO) Public Schools
- James DiPerna, Pennsylvania State University
- Lynn Malarz, National Education Association
- Howard Nelson, American Federation of Teachers
- Chidi Onyia, Lynwood (CA) Unified School District
- Gary Ritter, University of Arkansas
- Laura Snyder, North Carolina Department of Public Instruction
- Martha Thurlow, University of Minnesota

Teachers of Non-Tested Subjects

Why is it particularly difficult to measure the productivity of teachers of non-tested subjects?

1. Students' prior school experiences may influence their performance in some content areas more than others. As a result, it is more difficult in some subjects than in others to obtain reliable estimates of teachers' contributions to their students' performance. (Ballou, 2002)
 - a. One study in a large TN district revealed that 20% of math teachers were recognizably better or worse than average.
 - b. However, the percentage fell to 10% in language arts instruction and to about 5% among reading teachers.

Why is it particularly difficult to measure the productivity of teachers of non-tested subjects? (cont.)

2. Very limited research (one study) has examined the extent to which teachers of other subjects contribute to student achievement gains in reading and math. (Koedel, 2007)
 - a. In San Diego, social studies teachers – not just math teachers – contributed to student gains in mathematics achievement.
 - b. In addition, math teachers – not just English teachers – contributed to student gains in reading.

What are some potential ways to measure the performance of teachers of non-tested subjects?

1. Non-core teachers are eligible for **BONUSES BASED ON SCHOOLWIDE PERFORMANCE ONLY.**
2. Non-core teachers are eligible for **SOME, BUT NOT ALL,** of the individual performance incentives that teachers of core subjects can earn.

What are some potential ways to measure the performance of teachers of non-tested subjects?
(cont.)

3. No student test scores are used to determine non-core teachers' eligibility for rewards.
4. States or school districts adopt or create new student tests to assess teacher performance in the non-core subjects.

Examples of states/districts using Approach #1

Non-core teachers are eligible for BONUSES
BASED ON SCHOOLWIDE PERFORMANCE
ONLY.

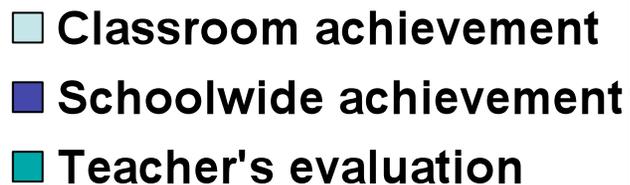
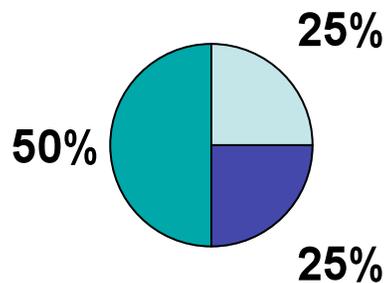
Examples:

- Alaska
- North Carolina
- TAP schools (e.g., Eagle County Public Schools)

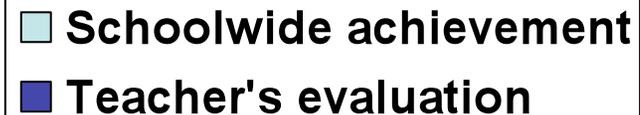
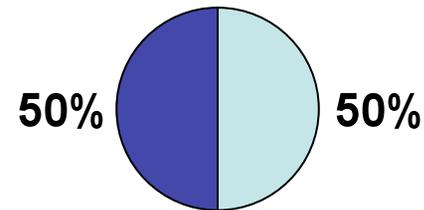
Variation: South Carolina TAP permits non-core teachers to emphasize either math or reading and tie their pay to student gains in one of those subject areas.

TAP example

Criteria for rewarding math and reading teachers



Criteria for rewarding teachers of non- tested subjects



Examples of states/districts using Approach #2

Non-core teachers are eligible for **SOME, BUT NOT ALL**, of the performance incentives that teachers of core subjects can earn.

Example:

- Houston

Houston's approach

1. Beginning in 2000, Houston paid bonuses to all teachers in “Exemplary” or “Recognized” schools.
2. However, this approach only rewarded teachers at high-performing schools, so HISD began rewarding individual teacher performance in 2006.
3. Strand 1: Schoolwide improvement in reading and math on TAKS. All faculty and staff in school eligible for bonus.

Strand 2: Compared student progress on the Stanford and Aprenda tests in one school to progress made by students in similar schools across HISD. Non-core teachers eligible for smaller bonuses than core teachers.

Strand 3: Year-to-year student progress in reading and math on TAKS. Only core teachers eligible.

Examples of states/districts using Approach #3

No student test scores are used to determine non-core teachers' eligibility for rewards. Instead, non-test criteria are used, such as classroom observations, acquisition of additional knowledge and skills, or assumption of additional roles or responsibilities.

Example:

- Colonial School District near Philadelphia
Variation: In Denver, non-core teachers design their own performance goals in their own area of specialization, and their rewards are based on attainment of these goals.

The Colonial School District's approach

1. Established mandatory pay-for-performance program in 1999-2000.
2. Based performance awards for core teachers strictly on student test scores on the TerraNova and the statewide achievement test.
3. Hired a consultant to establish criteria for judging teachers of non-tested subjects, with input from other sources.
4. Developed a separate evaluation system for evaluating teacher groups by grade/team/department at the elementary/middle school/high school levels.
5. Used pupil and parent surveys to evaluate the performance of non-teaching staff (e.g., nurses, counselors, and librarians).

Examples of states/districts using Approach #4

States or school districts adopt or create new student tests to assess teacher performance in the non-core subjects.

Example:

- Florida

Florida's approach

1. In Feb. 2006, FL adopted the mandatory E-Comp system. Districts were required to pay bonuses of at least 5% to at least the top 10% of teachers.
2. Reading and math teachers would be judged on student improvement on the FCAT (31% of FL teachers).
3. For teachers of other subjects (the other 69%), districts were required to develop standardized measures of performance (based on student learning gains, but not necessarily paper-and-pencil exams).
4. The new exams would be used to track student progress and determine teacher bonuses, but would not affect high school graduation or school grades.

Teachers of Non-Tested Grades

Why is it particularly difficult to measure the productivity of teachers of pre-K to Grade 2?

1. Assessment instruments appropriate for the early grades (e.g., DIBELS, MAP) tend to be adaptive tests designed to provide diagnostic feedback for instructional purposes, not to assess academic content mastery.
2. We don't know if it is appropriate to use these tests to measure teacher productivity because there is no solid research on the technical adequacy of using diagnostic tests for this purpose.
3. Potential for measurement error is high because assessments rely on very few observations.
4. Though value-added measurement experts recommend at least 3 yrs of standardized achievement test data to determine expected growth, most districts don't start testing until Grade 3.

What are some potential ways to measure the performance of **teachers of pre-K to Grade 2?**

1. Use DIBELS or MAP tests to assess performance at the early grades – risky.
2. Buy off-grade tests from vendor.
3. Create developmentally appropriate rubric to assess social/emotional development, motor development, etc. of young children.
4. Base rewards of teachers in pre-K to Grade 2 on schoolwide achievement gains and observed evaluations of their own classroom performance.

Why is it particularly difficult to measure the productivity of high school teachers?

1. No defined scope and sequence of curriculum at the HS level. Difficult to interpret standardized test results when you have radical differences in delivered curriculum.
2. Some HS teachers teach only in non-tested grades (e.g., 11th & 12th grade calculus). Individual teacher reward systems would exclude these teachers, who may be some of the top teachers in the school.
3. End-of-course tests need pretests to assess gains in performance.
4. ACT and SAT tests are designed to measure college preparedness. Not aligned to state standards – how curriculum sensitive are they?
5. Not certain whether use of MAP test is appropriate because it was designed for a different purpose.

What are some potential ways to measure the performance of high school **teachers**?

1. Develop end-of-course tests.
2. Use ACT or SAT tests.
3. Use MAP test to fill in non-tested grades and subjects.
4. Buy off-grade tests from vendor.
5. Base rewards of high school teachers on departmentwide performance, rather than individual classroom performance.

Teachers of English Language Learners

Why is it particularly difficult to measure the productivity of teachers of English language learners?

1. Highly mobile populations of students tend to have a lot of missing data.
2. Standardized achievement tests normed on populations of native English speakers may not be valid measures of ELL students' academic content knowledge. Tests may simply be measuring English proficiency.
3. Difficult to calculate expected gain when language of the test (and the test form itself) changes as students become increasingly proficient in English.
4. May be difficult to develop a districtwide teacher compensation policy because of variability in the concentration of English language learners across schools.

What are some potential ways to measure the performance of **teachers of English language learners**?

1. Base rewards of teachers of ELLs on schoolwide achievement gains and observed evaluations of their own classroom performance.
2. Consider using gains in English proficiency as an additional performance measure for teachers of ELLs.

Teachers of Students with Disabilities

Why is it particularly difficult to measure the productivity of teachers of students with disabilities?

1. In many schools, the numbers of students with disabilities are not in big enough numbers in any combination of disability categories.
2. Although the majority of students with disabilities take the district test, they receive test accommodations. Technically, you can't assert that results are on the same scale when you alter the test.
3. A small proportion of students with severe cognitive disabilities take alternative assessments, rather than the district test.

What are some potential ways to measure the performance of **teachers of students with disabilities**?

1. Base rewards of teachers of students with disabilities on schoolwide achievement gains and observed evaluations of their own classroom performance.

Lessons Learned – TIF Grantee Panelists

- Holly Bayonas and Emily Scott
Guilford County, North Carolina
- Jennifer O'Brien
Houston
- Tanly Cabrera and Joe Perez
Hillsborough County, Florida