

Looming Questions in Performance Pay

Reformers must define teacher and student performance and agree on the goals of schooling when proposing a performance pay scheme.

By Donald B. Gratz

The many current proposals regarding pay for performance address different problems and are built on different assumptions. As the debate proceeds, both proponents and opponents of the various schemes must answer three big questions: What is the definition of teacher performance? What is the definition of student performance? and What are the goals of schooling?

DEFINING TEACHER PERFORMANCE

Performance-based compensation requires a definition of teacher performance. What is the organizational bottom line, and how do teachers help achieve it? Some incentives are directed at specific teacher behaviors, such as participating in training, teaching in underperforming schools, mentoring younger teachers, or earning National Board certification. These may relate to school or district goals, be research-driven, or be generally believed to support better student outcomes.

The federal government's plan, in contrast, judges teacher performance on student test scores, which teachers can influence but not control. This approach has been historically unsuccessful, is based on questionable assumptions, and often leads to significant negative consequences. Defining teacher performance is not easy; measuring it is even harder. But failing to define teacher performance completely or rewarding only a narrow slice will ill serve students and teachers and will undermine public support.

DEFINING STUDENT PERFORMANCE

A system that bases teacher compensation on student performance also requires a definition of student performance. The proposal championed by government leaders focuses on students' standardized test scores. Although the current Administration promotes measures of student *growth*, rather than the simple cut scores of its predecessor, problems with standardized tests remain:

- More than half of America's teaching force is not covered by standardized tests, including specialists

DONALD B. GRATZ is professor, chair of the Education Department, and director of the master's program at Curry College in suburban Boston, Mass. He assisted with Denver's Pay for Performance Pilot in its early stages and wrote *The Peril and Promise of Performance Pay* (Rowman & Littlefield Education, 2009).



in media, art, gym, and special education; nurses and counselors; teachers of young children; and teachers of many secondary courses.

- Most testing programs cover only a few subjects — reading, math, and sometimes science — encouraging a dramatic narrowing of the academic curriculum.
- Student outcomes — such as critical thinking, communication, teamwork, creativity, and healthy social, emotional, and physical growth — are pushed aside.
- The time and resources devoted to this limited range of tests are already substantial. As more tests are added, the cost in dollars and time will increase, further reducing resources available for other activities.
- Most important, test scores are weak indicators of student success, and the pursuit of higher scores undercuts the educational process in other areas.

At best, over-reliance on standardized tests provides a limited description of any student's ability; at worst, it corrupts the educational process. "Our research informs us that high-stakes testing is hurting students, teachers, and schools," say Berliner and Nichols. "By restricting the education of our young people and substituting for it training for performing well on high-stakes examinations, we are turning America into a nation of test takers, abandoning our heritage as a nation of thinkers, dreamers and doers" (2007: 48).

Teachers and parents support standardized tests, but not to the exclusion of other measures and other topics (Bushaw and Gallup 2008). Assessment experts support their use only as one tool among many. As the National Research Council cautions, "Tests are not perfect. Test questions are a sample of possible questions that could be asked in a given area." A test score "is not an exact measure. . . no single test score can be considered a definitive measure of a student's knowledge." Furthermore, educational decisions "should not be made solely or automatically on the basis of a single test score. Other relevant information about the student's knowledge and skills should also be taken into account" (Heubert and Hauser 1999: 3). Still, the practice enjoys considerable support among policy makers, economists, and the press, for whom simplicity and ease of explanation appear to trump actual results.

But if not standardized tests, then what? If student results drive teacher results, what student results should be considered? Student results should be based on school goals, as discussed below. My recommendation: Start the discussion by broadly

defining student *success*, as determined by school goals, not student *achievement* or *performance*.

THE GOALS OF SCHOOLING

Most organizations link employee performance to organizational goals. If our goal for schools is high test scores, that's what we should measure. But if our goals include art, music, history, citizenship, critical thinking, communication skills, the ability to work with others, and lifelong learning, these must also be included. Do schools address just the intellectual domain, or should they seek to encourage children to grow socially, emotionally, and physically as well? Whatever we decide, it is hypocritical to claim a range of goals but to provide incentives for achieving only a few.

Not all goals are quantifiable, but no important goal should be ignored. Quantifiable goals often distort the assessment process, but performances, demonstrations, portfolios, and behavior scales can approximate many goals, as they do in other spheres (Gratz 2009). Such measures work if all parties agree that they reasonably address stated goals — for example, incentives to teach in inner-city schools. Although a national discussion around school goals is needed, school and district goals should be set at the local level, so parents, teachers, and citizens can be involved.

WHY PERFORMANCE PAY?

The reason usually given for reforming teacher pay is that U.S. schools are in a crisis. This case for crisis is most often presented in economic terms:

- **Productivity:** Schools aren't graduating workers with the skills to be productive. This lack of productivity will harm the workers themselves and may lead to economic collapse.
- **Undereducation:** Undereducated workers pay less in taxes and consume more in resources than workers with college degrees. The cost to the country in services and lost taxes is a significant drag on our economy.
- **International Tests:** American students underperform on international tests, particularly in math and science. This may cause economic decline, as American companies move overseas, outsource, or hire foreigners to fill important jobs.

Productivity: Predictions of disaster based on low

At best, over-reliance on standardized tests provides a limited description of any student's ability; at worst, it corrupts the educational process.

Too much attention diverted to comparison draws resources, attention, and time from student learning.

productivity have been sounded since *A Nation at Risk* in 1983, but American workers have actually been more productive since the 1990s than in any period in history. Unfortunately, this productivity hasn't increased their economic success. Instead, most of the gains have accrued to the wealthiest few, and the income gap between rich and poor has increased (Mishel and Rothstein 2007). Even though American schools produce workers capable of high productivity, this no longer guarantees economically successful workers. Our economic system is the source of this problem, not our education system.

The Cost of Undereducation: Undereducated workers pay less in taxes and cost more in services, so some

critics blame the achievement gap for U.S. economic weakness. A recent McKinsey report claims that, "If the United States had closed the racial achievement gap. . . by 1998, GDP in 2008 would have been between \$310 billion and \$525 billion higher, or roughly 2 to 4% of GDP" (McKinsey & Co. 2009: 17). But high school dropouts pay lower taxes because they earn less, and they earn less because of the jobs they hold. If the gap had been closed, and if all of today's high school dropouts and graduates had college degrees, there wouldn't be nearly enough good jobs for them. Instead, if we had closed the achievement gap in 1998, huge numbers of college-educated workers would be un- or under-employed today.

In fact, if more college-educated workers glutted the labor market, their earning power would likely drop. Skilled jobs would pay less than now, and skilled workers who couldn't find good jobs would take the lower-paying jobs now filled by unskilled workers.

Individuals who complete college usually earn more than high school graduates. But without a large increase in high-paying jobs, there would be no change in the economy.

International Tests: Finally, American students compare well on international tests with students from other industrialized nations. These nations score well but aren't often first, in part because of their broad diversity. However, the more important point is that, for countries in the top 50% of the economic scale — including all the industrialized nations — there is simply *no correlation* between test scores and the country's economic success. The countries with the highest scores don't have the strongest economies, and doing well on tests doesn't

predict economic strength (Tienken 2008). As for outsourcing, companies hire foreign workers because they can pay them less, not because American workers aren't available (Rotberg 2008).

Accurate predictors of national economic strength, such as the World Economic Forum in Switzerland, measure a broad range of factors — including, for example, creativity and innovation, in which the United States excels (Tienken 2008; Bracey 2008). If we want to promote economic success, we should focus on factors that actually correlate with that success, not tests.

Teacher Motivation: Policy makers appear to believe that the primary causes of low test scores are weak school and district leadership, a weak curriculum, and poor teaching. To address a weak curriculum, they propose higher standards and high-stakes tests. To address leadership and teaching, they propose sanctions for poor results and performance-based compensation.

If poor teaching causes low student test scores, what causes poor teaching? Test-based compensation plans suggest that teacher motivation is the primary cause, and financial incentives are the primary solution. The assumptions implied in this reasoning are troubling:

- Many teachers aren't trying hard enough because they aren't motivated.
- These teachers know what to do, but they don't do it because they lack a financial incentive.
- Financial incentives are more important to teachers than student success.

These are unlikely conclusions. While teachers want to be paid professionally, the evidence suggests that they aren't motivated primarily by financial rewards. If they were, why would they enter teaching?

Studies of workplace motivation consistently downplay the value of financial incentives. Indeed, management consultants put such incentives last on their lists of effective tools for organizational improvement, after "approval; trust, respect, and high expectations; loyalty, given that it may be received; job enrichment; and good communications" (Grimes 2006: 56). Professionals are motivated by positive working environments, the respect of their colleagues, the ability to apply their professional knowledge and skills in solving problems, and the opportunity to contribute to their immediate and larger communities.

Thus the constant barrage of criticism and the demeaning assumptions hardly seem motivational. Potential unintended consequences of this approach include *reducing* teacher motivation, driving creative

The danger is that the concept of school improvement will be hijacked by a federal education department awash in money but distant from local aspirations and realities.

Contrary to popular belief, many businesses issue merit raises based on supervisor or peer evaluations, but few rely on quantitative measures alone.

teachers from the field, and turning away potential new teachers.

The Achievement Gap: Another reason frequently cited for reforming teacher pay is the achievement gap between middle-class and poor children. This problem is one on which schools can and should take positive action. This problem, not international test scores, is worthy of concern.

But the problem is much broader than the schools. As unemployment skyrocketed, the income gap grew. In 2009, one in four American children was eligible for food stamps (DeParle and Gebeloff 2009). The causes of the achievement gap are too complex to explore in this brief essay, but schools alone can't eliminate it. Indeed, student achievement tracks family income more closely than any other factor.

Beware of solutions that blame schools alone for this gap; such accusations divert attention from the more serious and intractable issues in our health, human service, and economic systems. America's average earnings are similar to those in European countries, but we do far less to redistribute income equitably and provide far fewer social supports (Sawhill 2006). The United States also lags behind other countries in supporting children and families. In UNICEF's 2007 survey of child well-being, for example, we ranked next to last (Hargreaves and Shirley 2008). These kinds of factors must be addressed in any serious effort to eliminate the achievement gap.

Why are these points important? Few will dispute that improvements are needed in many schools or that some are truly in crisis. But if the problem is misdiagnosed, the solution is unlikely to work. And if we fear an epidemic when the problem is more isolated, time and resources will be drawn from other significant issues. We need to avoid the rush to judgment and do the job right. We must base our efforts on accurate assumptions and appropriate goals.

IMPLEMENTATION ISSUES

Implementation issues are huge in organizational change, especially in large and complex organizations. There are far too many to address here, but a few less often discussed issues are worth noting:

Assessment for Accountability or for Improvement: Assessments can be used either to compare students, teachers, and schools to one another or to help students learn. The difference matters. Portfolios, presentations, and rubric-driven assessments are good measures of student success and help promote learning, but they're impractical for comparison and accountability. Standardized tests compare students, teachers, and schools easily (if not always accurately), but they're too slow and not specific enough to use in helping students learn. Standardized tests

may have their place, but too much attention diverted to comparison draws resources, attention, and time from student learning

Complex Implementation: Linking multiple student results on multiple tests to multiple teachers is a daunting task. The many systems needed to support a performance pay plan are difficult to assemble. In addition, organizations don't build substantive change simply by announcing something new or adopting the practices of another. Successful change agents understand the need to build a constituency for change within the organization.

Systemic Approach: Changes in compensation are systemic reforms. To support ProComp, Denver realigned district operations, including curriculum, instruction, assessment, professional development, and personnel. Furthermore, school practices were tightly linked to district practices. Changes in school operations without related changes in district operations are often unsustainable. The system, not just some parts, needs to change.

CAMPBELL'S LAW

Public policy approaches built on incentives are not new, and they have often produced serious unintended consequences. Consider Campbell's Law, which states that, "the more any quantitative social indicator is used for social decision making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it was intended to monitor" (Nichols and Berliner 2007: 25-26). Attempts to measure such results as hospital mortality rates and to establish crime quotas for police have yielded distorted and sometimes dangerous results (Rothstein 2009). In education, cheating, teaching to the test, encouraging poor students to stay home, eliminating recess, and other inappropriate activities escalate when test scores are emphasized.

Contrary to popular belief, many businesses issue merit raises based on supervisor or peer evaluations, but few rely on quantitative measures alone. Most jobs are too complex and multifaceted to reduce solely to numbers.

Quantitative measures of performance may improve the particular result targeted, such as raising test scores, but the unintended consequences can be significant. "How much gain in reading and math scores is necessary to offset the goal distortion — less art, music, physical education, science, history, character building — that inevitably results from rewarding teachers or schools for score gains only in math

The evidence suggests that teachers aren't motivated primarily by financial rewards. If they were, why would they enter teaching?

and reading?” asks Rothstein. “How much misidentification of high- or low-performing teachers or schools is tolerable in order to improve their average performance? How much curricular corruption are we willing to endure. . . ?” (Rothstein 2009: 105).

Failing to define teacher performance completely or rewarding only a narrow slice will ill serve students and teachers and will undermine public support.

Start the discussion by broadly defining student *success*, as determined by school goals, not student *achievement* or *performance*.

WHAT TO DO

In a normal planning process, planners identify problems based on evidence and analysis, consider the potential causes, and design interventions to address these causes. They try to predict and avoid obstacles, monitor the results of intervention — both intended and unintended — and make needed adjustments. In such planning, there is clear alignment between goals, main problems, identified causes, and proposed solutions.

In addition to alignment, broad involvement is critical. Solutions that don’t address the underlying problems won’t succeed, and solutions that don’t engage workers who must implement them won’t be fully implemented. Change agents understand this need to build a constituency for change from the start (Kotter 1996; Pfeffer and Sutton 2006).

Attempts to impose alterations in professional behavior rarely succeed. They generate compliance, but not real change. That is why “proven programs” often fail when they’re simply adopted from other districts. Those who try to implement performance pay or any other major change would do well to en-

gage stakeholders in the entire process.

Many forms of teacher compensation are under discussion, including incentives to teach in hard-to-serve schools and hard-to-fill positions and incentives to earn National Board certification. Some districts are experimenting with mentor teachers, master teachers, and curriculum specialists, rewarding experienced teachers for sharing their expertise. Denver’s ProComp includes several of these approaches and also rewards teachers for meeting classroom objectives that they set themselves (with principal approval), thus engaging teacher interest and expertise.

The danger is that the concept of school improvement will be hijacked by a federal education department awash in money but distant from local aspirations and realities. The good news is that all parties are discussing the value of education and focusing on improving the teaching and learning process. The potentially positive steps include:

- Developing voluntary national standards by expert groups in specific disciplines — supported by but apart from the government — and including skills needed in the 21st century;
- Discussing goals for students, which may, in time, return us to a more balanced view of how different students can achieve success — the educational “bottom line”;
- Considering how educators add to this educational bottom line and how differentiated compensation might enhance the profession; and
- Beginning to move away from standardized tests (some major policy initiatives to the contrary).

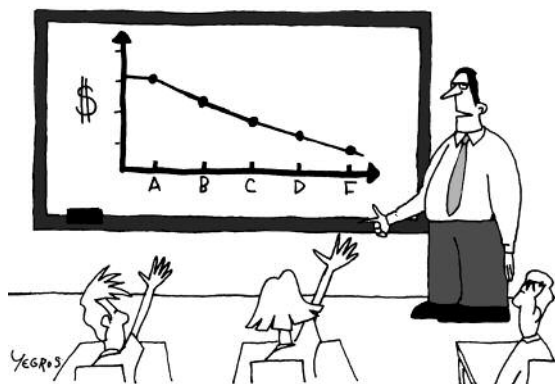
However, as long as policy makers and the press confuse student achievement with standardized test scores, we won’t improve the education of our children — even if test scores rise. When we decide to pursue a broader view of success, when we engage parents, citizens, and educators in shaping the visions and strategies in their own districts, there is the possibility of real improvement. Alternative forms of compensation may be a part — but only a part — of that change. **K**

REFERENCES

Berliner, David C., and Sharon L. Nichols. “High Stakes Testing Is Putting the Nation at Risk.” *Education Week*, March 12, 2007: 48.

Bracey, Gerald W. “The 18th Bracey Report on the Condition of Public Education: Schools-Are-Awful Bloc Still Busy in

Teaching Math in a Merit Pay World



“Let’s say you guys really, really disappoint me in your next quiz. What would my salary be then?”

2008." *Phi Delta Kappan* 90, no. 2 (October 2008): 108-109.

Bushaw, William J., and Alec M. Gallup, "Americans Speak Out: Are Educators and Policy Makers Listening?" *Phi Delta Kappan* 90, no. 1 (September 2008): 7-31.

DeParle, Jason, and Robert Gebeloff. "Food Stamp Use Soars, and Stigma Fades." *The New York Times*, Nov. 29, 2009. www.nytimes.com/2009/11/29/us/29foodstamps.html?_r=1.

Gratz, Donald B. *The Peril and Promise of Performance Pay: Making Education Compensation Work*. Lanham, Md.: Rowman & Littlefield, 2009.

Grimes, Cliff F. "Historical Perspectives." In *Employee Motivation, the Organizational Environment and Productivity*. London: Accel-Team, 2006.

Hargreaves, Andy, and Dennis Shirley. "Beyond Standardization: Powerful New Principles for Improvement." *Phi Delta Kappan* 90, no. 2 (October 2008): 135-143.

Heubert, Jay P., and Robert M. Hauser, eds. *High Stakes: Testing for Tracking, Promotion, and Grading*. Washington, D.C.: National Academy Press, 1999.

Kotter, John P. *Leading Change*. Cambridge, Mass.: Harvard Business School Press, 1996.

McKinsey & Co., Social Sector Office. "Introduction." In *The*

Economic Impact of the Achievement Gap in America's Schools. Washington, D.C.: McKinsey & Co., April 2009.

Mishel, Lawrence, and Richard Rothstein. "False Alarm." *Phi Delta Kappan* 88, no. 10 (June 2007): 737.

Nichols, Sharon L., and David C. Berliner. *Collateral Damage: How High-Stakes Testing Corrupts America's Schools*. Cambridge, Mass.: Harvard University Press, 2007.

Pfeffer, Jeffrey, and Robert Sutton. *Hard Facts, Dangerous Half Truths, and Total Nonsense: Profiting from Evidence-Based Management*. Cambridge, Mass.: Harvard Business School Press, 2006.

Rotberg, Iris C. "Quick Fixes, Test Scores and the Global Economy." *Education Week*, June 11, 2008: 27.

Rothstein, Richard. "The Influence of Scholarship and Experience in Other Fields on Teacher Compensation Reform." In *Performance Incentives*, ed. Matthew G. Springer. Washington, D.C.: Brookings Institution Press, 2009.

Sawhill, Isabel. *Opportunity in America: The Role of Education*. Princeton, N.J.: Brookings Institute, 2006, 1-3.

Tienken, Christopher H. "Rankings of International Achievement Test Performance and Economic Strength: Correlation or Conjecture?" *International Journal of Education Policy & Leadership* 3, no. 4 (April 25, 2008). www.ijepi.org.

We know your time is valuable.

Cut through the clutter of online resources with PDKConnect,
an elite community just for PDK members.

www.pdkconnect.org

