



# IssueBrief

March 2008

*This issue brief was made possible with the generous support of MetLife Foundation.*

## Measuring and Improving the Effectiveness of High School Teachers

Most education reformers agree that effective teaching is defined by improving student learning, but they disagree on how to measure teacher effectiveness and how to use those measurements to improve teaching. Thus far, most of the policy debate on teacher effectiveness has focused on using test scores to implement merit pay or to fire teachers, but those strategies alone will not lift teacher performance on a large scale. The best way to improve teacher effectiveness is to provide teachers with support and guidance that are grounded in effectiveness—that is, which uses effectiveness data to enhance professional development and teacher education, strengthen evaluations and career development, and revamp accountability policies to reward and encourage student learning. Only then can staffing, pay, or any other high school reform effort advance the primary goal of improving student achievement for college and work readiness.

In order to improve high school teaching, educators and policymakers must first invest in solid, objective ways to measure a teacher's effectiveness. Currently, many experts believe that the best method is to use “value-added” analysis, a statistical method described in more detail later in this brief. However, value-added analysis is not a perfect measure, and it works best when supplemented with other measures of student learning and of teacher knowledge and skill. With robust, multiple measures of teacher effectiveness, complemented by targeted professional development, high-quality evaluations, and smart accountability, educators and policymakers can indeed use effectiveness measures to improve the quality of high school teaching.

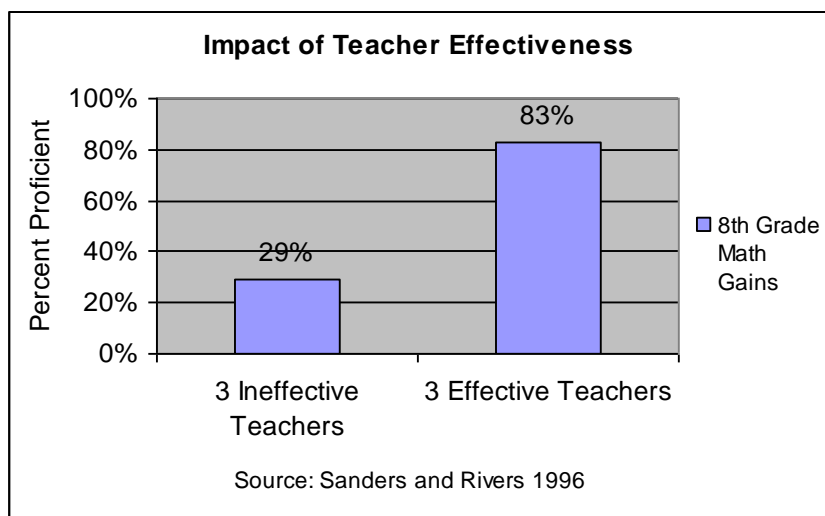
### How Should High School Teacher Effectiveness Be Defined?

For the purposes of this brief, high school teacher effectiveness is defined as demonstrating contributions to growth in student learning. Good high school teachers accomplish other things, including motivating and engaging students, acquiring new knowledge and skills, and collaborating with colleagues. But those accomplishments best serve their purpose when they lead teachers to improve student achievement. That said, student learning can be measured in many ways, including but not limited to gains made on standardized tests, improvement on periodic classroom assessments, performance on end-of-course exams such as Advanced Placement or International Baccalaureate assessments, or even standardized portfolios of student work. Regardless of the assessment instrument, teacher effectiveness is demonstrated when student learning improves, and on this point educators and policymakers agree (Darling-Hammond 2007; Gordon et al. 2006). Furthermore, there is consensus that learning in *high school* progresses toward the goal of graduation with the knowledge and skills needed to succeed in college, work, and as a citizen (Corbett and Huebner 2008). Therefore, the definition of high school teacher effectiveness must focus on growth in student learning that ensures students are successful after high school (Alliance for Excellent Education 2007).

Defining teacher effectiveness in terms of student outcomes, such as learning gains and college or work readiness, frees educators and policymakers to focus on which teachers are getting results and why. When data on student achievement is available, that data creates less need to rely on proxies for teacher effectiveness. For years, educators and policymakers have measured and rewarded teacher quality based on qualifications such as having achieved years of experience, certification status, or advanced degrees, and teachers are generally rewarded based on those criteria. Qualifications serve as quality controls and can sometimes predict student achievement. Nevertheless, qualifications are approximations of effectiveness, rather than direct measures of student learning (Clotfelter et al. 2007; Goldhaber and Brewer 2000; Goldhaber 2006; Kane and Staiger 2005; Hanushek et al. 2005; Jepsen and Rivkin 2002).

Research is clear that what a teacher does in the classroom is a far greater predictor of student success than anything else, and students who consistently get effective teachers benefit exponentially (Gordon et al. 2006). Researchers in

Tennessee have found that students given the most effective teachers for three years in a row made over twice the gains of comparable students assigned to the least effective teachers (Sanders and Rivers 1996). Researchers have even found that effective teachers have such a significant impact on a student's ability to learn that teaching can offset learning challenges such as low income levels and achievement gaps (Rivkin et al. 2002; Clotfelter et al.



2007). By showing the magnitude of difference in student outcomes as a result of teaching, such research underscores that effectiveness, more than any other indicator of teacher quality, is the area in which policymakers and educators must focus their attention in order to improve student achievement.

## How Should High School Teacher Effectiveness Be Measured?

Measuring teacher effectiveness is a thorny issue—methodologically, practically, and politically. A teacher's impact on student achievement can be measured in many ways, with different results depending on the learning measured or the instrument used to measure it. Naturally, when policymakers suggest determining teachers' pay or career status based on one imperfect measurement, conflict and concern arise. Thus research and good practice indicate that high school teacher effectiveness should be measured using student learning gains in multiple ways, supplemented by other relevant indicators of effectiveness such as a teacher's acquisition of knowledge and skill that lead to student achievement.

### Student Learning Gains: The Essential Measure of Teacher Effectiveness

Many reformers argue for using standardized assessments to measure student learning and, thus, teacher effectiveness. To be sure, tests have limitations as measures of student learning: they are not



explicitly designed to measure teacher quality, test scores have margins of error, and some tests do not align to curriculum standards (Braun 2005; Gore 2007; Elmore 2002). However, for policy purposes, standardized tests are currently the best objective and quantifiable measure of student learning available. While not perfect, tests are certainly better measures than what is currently used to judge and to reward teachers—haphazard and irregular evaluations and qualifications, which are proxies for effectiveness. Ultimately, measures of student learning put the focus squarely on the end result of what effective teaching should accomplish, leaving the means up to the teacher and to the school.

However, judging teachers by the performance of students on a *single* test is not an accurate measure of what teachers contribute to student learning. High-achieving students may hit an achievement target no matter what, while low-achieving students may take more than one year to hit the same target. Thus, the best way to measure teacher effectiveness is to measure the amount of growth a student makes over time, demonstrated on several assessments. That is why advocates for effectiveness measures argue that “value-added” is the most objective and fair way to assess teaching, and that it should be supplemented with other means of measuring student learning gains (Harris 2007; Gordon et al. 2006).

### **Using Value-added to Measure Teacher Effectiveness in High Schools**

Value-added is a complex statistical method for determining the impact a teacher or school—versus other factors, including income level, prior achievement, and school characteristics—makes on student achievement. Taking such factors into account, value-added analysis estimates the academic growth a student is expected to make for the year and compares it to how the student actually performs on standardized assessments. Students who make greater gains than expected are judged to have teachers who “added value” whereas students who did less well than anticipated have teachers who did not, with many students and teachers falling in the middle range.<sup>1</sup> Interestingly, value-added may be most accurate in identifying top- and low-performing teachers, rather than in distinguishing between the majority of teachers who cluster in the middle (Gordon et al. 2006; but see Braun 2005). Currently, then, the best use of effectiveness measures may be to serve as a flag to investigate which teachers need in-depth assistance or to be counseled out of the profession and which top-performing teachers might serve as mentors or master teachers.

**Determining a teacher’s added value** at the high school level is more complicated than at lower grade levels, so it must be done with caution. For example, at the high school level, student achievement, especially in the humanities, is influenced by multiple teachers, so special statistical methods must be used to determine a student’s achievement and a teacher’s contribution. To determine value-added, high schools must first administer yearly assessments that are aligned with one another on a common scale for comparison purposes, or they must administer end-of-course exams in each subject for which teacher effectiveness is to be measured. Most high schools do not have standardized end-of-course exams or administer standardized tests every year. Therefore, value-added will mostly capture the impact of teachers in core academic subjects, such as English, math, science, and social studies, for which standardized assessments are available, while additional assessments will have to be created for other subjects. Second, assessment results must be compared with students’ prior levels of achievement to gauge their progress over the year, using statistical analysis to isolate the teacher’s

---

<sup>1</sup> For a technical overview of value-added and its usefulness for policymaking, see Harris 2007, Braun 2005, and McCaffrey et al. 2003. For a lay-friendly introduction, see Carey 2004, Ballou 2002, and Stewart 2006.



contribution. While end-of-course exams are not usually aligned from one test to the next, researchers in Dallas find that prior student achievement in math and reading, subject-specific achievement, and course-taking patterns can help to establish a valid and reliable baseline from which to estimate expected academic growth for each course (Babu and Mendro 2003).

In terms of **using value-added data**, statistical information must be translated into a user-friendly document that educators clearly understand and can use to assess their practices. Moreover, additional information about classroom practice must be gleaned to help a teacher identify practices that contributed to learning, and what can be done to improve. Such information can come from a variety of sources, including periodic classroom assessments, benchmark exams, or even classroom observations and quality evaluations. Since value-added data can be highly technical, teachers and principals need training and school structures that support their use of the data. Because of most high schools' large size and complicated scheduling, they are not automatically set up to provide teachers time to collaborate around student achievement data. The move toward transforming large high schools into smaller learning communities has offered one strategy for addressing the time and collaboration challenges, though structural reforms must be made a priority in order to maximize the benefits of measuring teacher effectiveness (Supovitz and Christman 2003).

### **The Pros and Cons of Value-added Analysis**

The promise of value-added analysis is that it isolates teacher effectiveness in an objective and comparable way, and it grounds conversations about teacher improvement in data on student outcomes. However, value-added has some limitations that are important to keep in mind (McCaffrey et al. 2003; Harris 2007).

At its best, value-added analysis can only say that a student, and thus a teacher, exceeded or fell short of expectations; it cannot determine which teaching practices led to improved student achievement or which practices must change—a big reason why it must be supplemented with other data. Other current limitations for policymakers to consider include the lack of yearly and aligned tests in high school grades, the lack of standardized tests for many subjects, missing data, and the confounding impact of school factors like nonrandom assignment of students to teachers (Braun 2005). This last concern is a frequently cited objection to value-added because teachers willing to work with more challenged students may experience less progress than teachers with already high-achieving students. In addition, many critics point out the limits of any standardized test to document the full range of student learning. Since value-added relies on test scores, it is subject to the margins of error and statistical limitations of any tool used to measure learning (McCaffrey et al. 2003; Elmore 2002).

All of that said, when value-added is done carefully and supplemented with other measures of student learning, it can generate reliable data that can be used to help high school teachers improve (Aaronson et al. 2007; Battelle for Kids 2008). Fortunately, value-added analysis at the high school level can compensate for the factors listed above, even the nonrandom assignment of teachers, by using complex statistical methods such as school fixed effects and calculating test score dispersions for actual versus simulated (random) teacher assignments (Aaronson et al. 2007; Clotfelter et al. 2004). Moreover, value-added is already in use at the high school level in multiple locations across the country. Researchers in Chicago produced reliable, valid, and stable estimates of high school math teachers' value-added (Aaronson et al. 2007). Over the past decade, high schools in Tennessee, Ohio, North Carolina, Texas, and Colorado have generated high school value-added data to improve teaching



(Bratton et al. 1996; Babu and Mendro 2003). And researchers in Ohio, New York, and Wisconsin are currently developing stronger value-added models and programs for using value-added data to improve teaching in high schools (Battelle for Kids 2008; Council of Great City Schools 2008).

### **Other Ways to Measure Student Learning**

As the methodology of value-added analysis improves and its limitations are worked through, other measures of student learning gains should be used to get a full picture of what happens in the high school classroom (Harris 2007; Gordon et al. 2006). Some high school teachers administer pretests at the beginning of the year or the start of a unit and then administer a post-test at the end, measuring students' growth in learning along the way. Interim assessments (also known as benchmark exams), aligned with state accountability tests or even periodic classroom (formative) assessments, can provide more frequent effectiveness data than annual tests. They also provide richer information on what skills or topics students are or are not mastering (Perie et al. 2007). In this way, benchmark exams and formative assessments chart a course for student and teacher improvement. Formative assessments have the added benefit of being tied directly to individual teachers and their classroom practice (Darling-Hammond 2007).

### **Other Measures of Effectiveness: Teacher Knowledge and Skill**

Measures of teacher knowledge and skill can be helpful in supplementing measures of student learning gains in determining high school teacher effectiveness. The key is to ensure that supplemental qualities can indeed be measured and linked to student outcomes in ways that help teachers improve their practice. While measuring teachers' knowledge or skill is one step removed from quantifiable gains in student achievement, some sets of knowledge and skill have been linked to student achievement and student outcomes. What is listed below, while not exhaustive, represents what can be measured in order to improve high school teacher effectiveness:

- Effective high school teachers must know both their subject and methods for teaching that subject, particularly for math and science (Walsh and Tracy 2004; Allen 2003; Monk 1994). Quantitative research is mixed on the impact of content knowledge and pedagogy for subjects other than math, but all high school subjects are complex enough that a solid grasp of content and teaching methods is required for effective teaching in all subjects (Darling-Hammond 2006).
- Student engagement is the key to learning and part of good pedagogy in upper grades, more so than in earlier grades (National Research Council 2004). The ability to motivate students is a baseline skill that effective high school teachers need to deliver content and to prepare students for college, and it can be measured in ways that inform instruction (Conley 2007).
- Given the changing demographics of U.S. schools, effective high school teachers must be prepared to work successfully with diverse student populations such as English language learners or students of differing cultural backgrounds (Ladson-Billings 1999; Wenglinsky 2002; Short and Fitzsimmons 2007). Enrollment rates for English language learners have increased exponentially in the last ten years, with the fastest rate of growth happening in secondary schools (NCELA 2007; Capps et al. 2005). The same is true for immigrant students; growth of this population is largest at the secondary level, and the majority of these students are new arrivals (Capps et al. 2005).
- Effective high school teachers must know how to impart literacy skills in their discipline. Literacy instruction is essential to teacher effectiveness in the upper grades because large numbers of older students still struggle to read and to write proficiently, and all students need ongoing literacy training that prepares them for college and work (Biancarosa and Snow 2006; Heller and Greenleaf





2007). Better literacy skills improve student achievement in math, science, and social studies, as well as English; and they improve college enrollment, college persistence, and college grade point averages (GPA) (ACT 2006).

- Since the ultimate outcome of high school should be students' readiness for college and work, high school teacher effectiveness should include the ability to prepare students for those increasingly similar challenges (Alliance for Excellent Education 2007). College and work readiness can be measured by how well students perform after high school and by the extent to which teachers employ strategies associated with higher-order tasks related to college and work (National Research Council 2000; Wenglinsky 2002).

The key to any point above is that it be measurable and link knowledge and skill to student outcomes. For example, Oklahoma now provides feedback to all high schools on the performance of their students who attend postsecondary institutions in-state, including ACT scores, college enrollment rates, college GPA, and remediation rates (Oklahoma State Regents for Higher Education 2008). As state data systems improve in scope and quality, more states can and should follow suit (Data Quality Campaign 2008).

## **How Can High School Teacher Effectiveness Be Improved?**

A primary goal of measuring high school teacher effectiveness is to improve the knowledge and skills of teachers so that they improve student achievement. Some reformers advocate effectiveness measures solely to implement merit pay or to fire teachers based on test scores. Merit pay may be a useful reform when done well, and some chronically ineffective teachers should be dismissed. But caution is needed when using test scores or value-added analysis to evaluate teaching. At this point, effectiveness measures may best be used to inform improvement strategies such as enhancing pre-service and in-service training, strengthening evaluations and career development, and revamping accountability policies (Darling-Hammond 2007).

### **Enhancing Professional Development and Teacher Preparation**

The most immediate use for effectiveness measures is to target and strengthen *professional development*, including evaluating which professional development programs are the most productive in enhancing teacher effectiveness. High school teachers armed with specific data on student progress and who receive constructive feedback from robust observations of their instruction can markedly improve student achievement (Wenglinsky 2002; Killion 2002). If value-added works best in identifying top- and low-performing teachers, then its use in professional development lies in leveraging the expertise of top-performers to improve the skill of low-performers, with the caveat that some chronically low-performers may need to be counseled out of the profession. When the strengths of effective teachers in a school are directed toward improving the practice of other teachers, professional development fosters collaboration and builds capacity within a school, thus reducing the need for outside “experts” (McLaughlin and Talbert 2001; Hirsh and Killion 2007).

All teachers can benefit from professional development focused on improving effectiveness. Research is clear that one-day workshops are inadequate to improve teaching (Garet et al. 2001). But staff development that happens regularly, in the building where the teacher works—and that is driven by clear goals, useful data, and teacher input—does improve effectiveness (Supovitz and Christman 2003; Education Trust 2005; Garet et al. 2001). Research also shows that, at the high school level, teacher



effectiveness improves when teachers collaborate as part of learning communities—groups of teachers working together (rather than apart as most teachers do) to improve student achievement and to build a culture of shared responsibility for learning (Hirsh and Killion 2007). So, at the high school level, enhancing professional development relies on reorganizing the school schedule to provide weekly, if not daily, time for teacher learning and common planning focused on using student learning data from multiple sources—from value-added data to actual student work—rather than coming up with money to send teachers to workshops. In addition to dedicated time, high school teachers need training in the use of effectiveness data and strong leaders who actively support them and focus common planning time on student outcomes (Supovitz and Klein 2003; Supovitz and Christman 2003).

Effectiveness measures also have the potential to enhance *teacher education* programs. Current projects in Louisiana, Ohio, Virginia, and at eleven teacher preparation institutions across the country examine teacher effectiveness data to strengthen their teacher preparation programs (Noell 2004; Teacher Quality Partnership 2008; Data Quality Campaign 2007; Carnegie Corporation of New York 2008). Louisiana uses value-added data to discern the effectiveness of candidates graduating from its teacher preparation programs (Noell et al. 2007). Ohio’s Teacher Quality Partnership includes every teacher education provider in the state, all of whom have agreed to report on their graduates’ effectiveness. That data will result in a longitudinal analysis of teachers’ impact on student achievement due to be released in 2009 (Teacher Quality Partnership 2008). Most of these programs have only preliminary results that need refinement, but over time, they can kick-start tough, but necessary, conversations on which programs produce effective teachers, which do not, and what changes must be made accordingly.

### **Professional Development Driven by Teachers and Effectiveness Data**

In 1998, barely 70 percent of students at Norview High School in Norfolk, VA, passed the state reading exam, and only 30 percent passed algebra or geometry. In every subject, African American students performed less well than their white peers. In response, Norview teachers mobilized to improve achievement. Teachers were grouped into teams by subject area, adopted shared curriculum guides and common assessments, and met regularly as teams around assessment data in order to review student progress. To evaluate their effectiveness as teachers, teams focused on three central questions: “What am I teaching well?”; “What am I not teaching well?”; and even “Why do your students perform better than mine?” Struggling teachers observed successful teachers in the classroom. Six years later, the results were clear: Norview raised achievement and narrowed gaps. In 2004, 93 percent of students passed the reading exam, 94 percent passed the algebra exam, and just under 90 percent passed the geometry exam. In reading and algebra, Norview no longer had a black-white achievement gap. Norview continues to post high achievement, to narrow gaps, and to outpace the average performance of high schools across the state. In 2006, all subgroups of students made Adequate Yearly Progress, 90 percent of students passed the reading exam, and 87 percent of students passed the math exam.

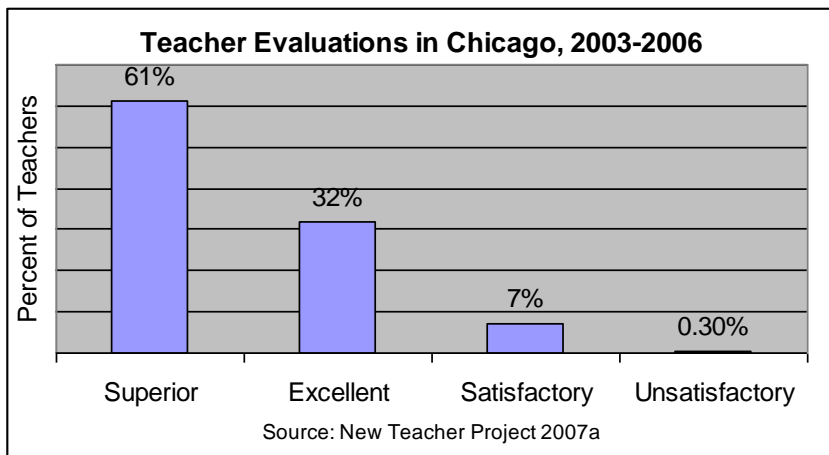
For more information, go to [www.all4ed.org/publications/ReadingNext/NorfolkReadingCaseStudy.pdf](http://www.all4ed.org/publications/ReadingNext/NorfolkReadingCaseStudy.pdf).

### **Strengthening Evaluations and Career Development**

In essence, *teacher evaluations* are assessments of teachers’ work based on classroom observations done by a principal or administrator. However, most evaluation tools are poorly constructed, and they are administered haphazardly. Thus, evaluations are generally dismissed as ways to improve effectiveness. But when taken seriously and executed thoughtfully, evaluations can chart a course for improving effectiveness and for advancing a teacher’s career (Toch and Rothman 2008).



Under the current approach, most evaluations are given too infrequently and are too inadequate to improve effectiveness (Toch and Rothman 2008). Only fourteen states require annual teacher evaluations, and only twenty-nine states require that evaluations, whenever they happen, even include classroom observations (National Council on Teacher Quality 2007). When evaluations do occur they



are often meaningless. Principals are rarely trained to use evaluations as a way to improve teaching, and few have ample time or expertise to evaluate every teacher. However, these are not the only reasons that evaluations mean little. In a study of Chicago public schools, principals admitted that they inflate evaluation ratings because they do not value the instrument and because they want to avoid the cumbersome grievance or dismissal process (New Teacher

Project 2007a). Therefore, from 2003 to 2006, 93 percent of Chicago teachers were rated as “excellent” or “superior” while less than 1 percent were deemed “unsatisfactory.” During that time, 79 percent of the city’s 87 failing schools did not issue a single unsatisfactory rating (New Teacher Project 2007a). Similar practices occur across the country (New Teacher Project 2007b, New Teacher Project 2007c, Education Daily 2008).

On the other hand, meaningful evaluation instruments do exist, and they hold promise for identifying and improving effective teaching. The best evaluations have explicit standards for the instruction to be assessed and clear rubrics for assessing it. They use multiple measures to gauge effectiveness, including but not limited to student learning gains. Good evaluations take place several times throughout the year and are administered by multiple evaluators, some of whom are peers and some of whom are administrators. The best evaluators are those who come from the same subject area or grade level as the teacher being evaluated. But regardless of their position, evaluators should be rigorously trained and their work routinely reviewed for accuracy (Toch and Rothman 2008).

Evaluations are most worthwhile when they are used to improve teaching. Thus, evaluations must be credible and reliable enough to merit teachers’ trust. Mentor teachers can then use data from evaluations to guide low-performing teachers in devising an improvement plan based on student learning data and standards of instruction. To enhance evaluations, administrators can provide common planning time for teachers to review student work or release time for struggling teachers to observe effective teachers. In this way, teachers begin to take shared responsibility for student learning (Hirsh and Killion 2007). School leaders must also ensure that collaboration focuses on instruction; research shows that, absent leadership, planning time degenerates to procedural matters (Supovitz and Christman 2003).

In addition, *career ladders* that focus on improving teaching through advancement in the profession are healthy companions to evaluations and effectiveness data. Career ladders provide new roles for teachers with additional pay and responsibilities as they increase their knowledge and skills (Chait 2007). Career ladders are favored and often led by teachers’ unions; the most widely known is the





Career in Teaching program in Rochester, NY. In this program, teachers advance from intern (new) to resident to professional to lead teacher, earning greater pay and more responsibility along the way (Koppich et al. 2002). Other notable forms of career development have been developed in Denver and as a component of schools participating in the Teacher Advancement Program (Denver Public Schools 2008; National Institute for Excellence in Teaching 2008).

Informed by data from effectiveness measures and evaluations, career ladders provide a way to recognize effective teaching, and they create an incentive for teachers to improve over time. Career ladders also offer teachers a way to advance in the profession without having to leave the classroom. Under most compensation schemes, the best way to increase salary substantially or “move up” is to become an administrator, which has led to a glut of educators holding degrees in educational administration but not actually filling positions (Mazzeo 2003).

### Using Evaluations and Career Development to Improve Teaching

The Teacher Advancement Program (TAP) is one promising approach to teacher evaluation and career development. TAP is used more than 180 schools, including approximately fifteen high schools in seven states. TAP's success is due not to an evaluation instrument alone, but rather to how it fits within a larger framework of professional development, career advancement, and differentiated pay. In TAP, teachers meet weekly in clusters, led by master and mentor teachers, to review student work and to improve instruction. Cluster meetings provide ongoing, job-embedded professional development, and the roles of master and mentor teachers provide effective teachers with a career ladder and additional pay. Four to six times a year, teachers are evaluated by multiple certified evaluators in their classroom using a sophisticated observation tool. Once a year teachers are given a value-added score based on student achievement gains to supplement their evaluations. However, it should be noted that TAP high schools provide value-added scores for the whole school rather than for individual teachers for two reasons: 1) value-added for individual teachers is more difficult to determine at the high school level and 2) to underscore that student growth rises and falls on teamwork. Thus, in TAP, feedback on student learning gains works in tandem with regular classroom observations to provide a robust evaluation of teaching, and effective teachers are rewarded with additional pay. Most importantly, professional development and evaluations guide a teacher to becoming more effective. Research has shown that TAP improves student learning and teaching practice, though a recent study finds less success in the upper grades (Solmon et al. 2007; Schacter et al. 2002; Springer et al. 2008). As the TAP model is refined for high schools, researchers and educators agree that the components of TAP—multiple career paths, ongoing professional development, instructional accountability, and performance-based compensation—are key to improving teacher effectiveness at the high school level (see Darling-Hammond 2007; Toch and Rothman 2008).

For more information, go to [www.talentedteachers.org](http://www.talentedteachers.org).

### Revamping Accountability Policies

Current teacher accountability policies do not focus on teacher effectiveness. The vast majority of teachers receive tenure *pro forma* after several years in the profession, without having to demonstrate tangibly that they have improved student achievement (Gordon et al. 2006). Federal law holds schools accountable for ensuring that teachers are highly qualified, but that provision focuses on teacher qualifications rather than actual impact on students in the classroom.<sup>2</sup> In contrast, if multiple, robust

<sup>2</sup> Under the No Child Left Behind Act, new high school teachers are deemed “highly qualified” when they 1) hold a bachelor’s degree, 2) earn state certification, and 3) demonstrate competency in each subject they teach. Veteran teachers must meet all three standards, though they can demonstrate subject matter competency in different ways, such as by meeting high objective uniform state standards of evaluation (HOUSSE).



measures of student learning are used to gauge teacher effectiveness—along with measures of acquired knowledge and skill and quality evaluations—then surely tenure and accountability decisions should be made, in large part, based on classroom effectiveness rather than exclusively on accruing qualifications. To date, proposals to use effectiveness measures to hold individual teachers accountable have been met with mixed reactions by teachers and administrators (Commission on No Child Left Behind 2007; Gordon et al. 2006). The bulk of opposition appears to result from skepticism that test scores can reliably measure effectiveness (Hoff 2007).

Therefore, it is clear that trust must be built in effectiveness measures so that teachers use them to improve their practice, and the right incentives must encourage this outcome. Teachers must value the measures as both accurate and helpful. Research supports using a combination of external (e.g., high stakes accountability or merit pay) and internal incentives (e.g., asking teachers to reflect on their personal goals) to stimulate improvement in teaching (Marks and Wright 2002). For most good teachers, their internal sense of accountability probably drives change more than external demands of an accountability system (Elmore 2002). However, some teachers do not have the motivation to improve on their own. In addition, research specifically on high school accountability suggests that pressure from the outside can foster change within schools and classrooms. For example, states with stronger accountability policies are more likely to stimulate change in curriculum and instruction in low-performing high schools, even as the bulk of motivation for change came from individual teachers (Geortz and Massell 2005). This finding confirms other research that suggests teacher collaboration must be focused on improving student achievement by outside forces such as strong school leaders (Supovitz and Christman 2003), even though it is teacher collaboration which leads to gains in student learning (Hirsh and Killion 2007).

Principals also need the right mix of incentives to leverage effectiveness measures for improvement—perhaps being held accountable for student performance but given flexibility in hiring and firing teachers (Toch and Rothman 2008). Two important studies find evidence that principals can be accurate evaluators of teacher effectiveness, though factors like accountability for student achievement and authority over staffing matters must be considered (Jacob and Lefgren 2005; Harris and Sass 2007). Ideally, accountability policy provides incentives for principals to take evaluations seriously while hiring policy allows them to leverage evaluations to improve staffing (Toch and Rothman 2008).

In sum, effectiveness measures may work best to revamp accountability policies when responsibility for student growth primarily rests on the shoulders of schools as a whole, since multiple teachers contribute to student learning, particularly at the high school level. Furthermore, cooperation between, rather than competition among, teachers is needed to improve student achievement on a large scale. Nevertheless, the success of programs such as TAP suggests that individual incentives—robust evaluation or extra pay—combined with professional support are crucial factors in increasing teacher effectiveness, even if the bulk of accountability falls on whole schools.

## **Where Do We Go From Here?**

Most educators and policymakers agree that growth of student achievement toward the goal of college and work readiness defines high school teacher effectiveness. Where the policy discussion turns contentious is in how to measure effectiveness and how to use effectiveness measures. The effectiveness measures described above, while far from perfect, hold great promise for zeroing in on



student learning growth and charting a path for improving student outcomes. However, policymakers and educators must take some clear steps to ensure that these measures are accurate and actionable.

First, measures of student learning must improve. Value-added analysis can be strengthened, especially at the high school level, as student assessments themselves are enriched, and as the statistical methods for determining value-added improve. Further, states and districts must enhance their data systems so that yearly student data is accessible and so that student data can be tied to teacher data in ways that produce effectiveness measures on a large scale. Right now only fifteen states have the ability to link student data to teachers in order to perform value-added analysis (Data Quality Campaign 2007). Because value-added or test scores should be supplemented with alternate measures of student learning (e.g., benchmark exams or classroom assessments) alternate measures must also be strengthened and new ones developed, such as standardized portfolios of actual student work.

Second, policymakers and educators must develop and strengthen teacher effectiveness measures that assess knowledge, skill, and classroom practice. Few valid and reliable measures of teacher knowledge and skill exist, while most teacher evaluation tools are inadequate. Learning from research and promising practice can help in this area, guiding policymakers to invest in quality assessments. Sophisticated measures of teacher effectiveness and high quality evaluations are costly, in time and money, but when executed well, they are useful and essential to improve teaching practice. Given the number of studies that show teacher effectiveness leads to better student outcomes, a sizable investment in evaluations and other teacher quality reforms would certainly be worthwhile.

Third, policymakers and educators must improve the school structures that allow effectiveness measures to improve teaching. The execution and culture of professional development must markedly change. Too much professional development continues to rely on one-shot workshops instead of using effectiveness data, common planning time, and strong leadership to target staff development. Evaluations and career ladders must be enhanced using effectiveness measures to identify top-performing teachers, recognize and reward their work, and harness their expertise to build the capacity of low-performing teachers. In terms of accountability, following the lead of innovative practice, policymakers must identify the right mix of incentives and consequences that fosters better teaching that leads to better student achievement.

In the end, improving the effectiveness of high school teachers rises and falls on the quality of effectiveness measures and the policies related to them that encourage or inhibit improved teaching. Perfect measures and definitive improvement policies do not yet exist, but as standards continue to rise and as education becomes increasingly important in a globalized economy, policymakers and educators must together find new and better ways to measure and to improve high school teacher effectiveness.

## MetLife Foundation

The Alliance for Excellent Education is grateful to MetLife Foundation for its generous financial support for the development of this brief. The findings and conclusions presented are those of the Alliance and do not necessarily represent the views of the funder.



## References

- Aaronson, D., L. Barrow, and W. Sander. 2007. *Teachers and student achievement in the Chicago public high schools*. *Journal of Labor Economics* 25, no. 1.
- ACT, Inc. 2006. *Reading between the lines*. Iowa City, IA: Author.
- Allen, M. 2003. *Eight questions on teacher preparation*. Denver, CO: Education Commission of the States.
- Alliance for Excellent Education. 2007. *High school teaching for the twenty-first century: Preparing students for college*. Washington, DC: Author.
- . 2004. *Tapping the potential: Retaining and developing high-quality new teachers*. Washington, DC: Author.
- Babu, S. and R. Mendro. 2003. *Teacher accountability: HLM-based teacher effectiveness indices in the investigation of teacher effects on student achievement in a state assessment program*. Presented at the American Education Research Association Annual Meeting.
- Ballou, D. 2002. *Sizing up test scores*. *Education Next*. Summer 2002.
- Battelle for Kids. 2008. *High school value added project*. [http://www.battelleforkids.org/home/value\\_added/va\\_highsch](http://www.battelleforkids.org/home/value_added/va_highsch) (accessed February 17, 2008).
- Biancarosa, C., and C. E. Snow. 2006. *Reading next*. 2nd ed. Washington, DC: Alliance for Excellent Education.
- Bratton, S., S. Horn, S. Wright. 1996. *Using and interpreting Tennessee's value-added assessment system*. Knoxville: University of Tennessee.
- Braun, H.I. 2005. *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service.
- Capps, R., M. Fix, J. Murray, J. Ost, J. Passel, and S. Herwanto. 2005. *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Washington, DC: Urban Institute.
- Carey, K. 2004. *The real value of teachers*. Washington, DC: The Education Trust.
- Carnegie Corporation of New York. 2008. *Teachers for a new era*. <http://www.teachersforanewera.org/> (accessed February 17, 2008).
- Chait, R. 2007. *Current state policies that reform teacher pay*. Washington, DC: Center for American Progress.
- Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor. 2007. *Teacher credentials and student achievement in high school: A cross-subject analysis with fixed effects*. Washington, DC: Urban Institute.
- Clotfelter, C.T., H.F. Ladd, and J.L. Vigdor. 2004. *Teacher sorting, teacher shopping, and the assessment of teacher effectiveness*. Unpublished manuscript, Public Policy Studies, Duke University.
- Commission on No Child Left Behind. 2007. *Beyond NCLB: Fulfilling the promise to our nation's children*. Washington, DC: The Aspen Institute.
- Conley, D. 2007. *Toward a more comprehensive conception of college readiness*. Eugene, OR: Educational Policy Improvement Center.
- Corbett, G. C., and T. A. Huebner. 2008. *Rethinking high school: Preparing students for success in college, career, and life*. San Francisco: WestEd.
- Council of the Great City Schools. 2008. *First round of senior urban education research fellowships announced*. Press Release. September 18, 2007. [www.cgcs.org/pressrelease/Research\\_Fellowship.pdf](http://www.cgcs.org/pressrelease/Research_Fellowship.pdf) (accessed March 7, 2008).
- Darling-Hammond, L. 2007. *Recognizing and enhancing teacher effectiveness: A policymaker's guide*. In L. Darling-Hammond and C. D. Prince (eds.), *Strengthening teacher quality in high-need schools—policy and practice*. Washington, DC: The Council of Chief State School Officers.
- . 2006. No child left behind and high school reform. *Harvard Educational Review* 76, no.4.
- Data Quality Campaign. 2008. *Developing and supporting P–20 education data systems*. Austin, TX: National Center for Educational Accountability.
- . 2007. *Benefits of and lessons learned from linking teacher and student data*. Austin, TX: National Center for Educational Accountability.
- Denver Public Schools. 2008. *ProComp*. <http://denverprocomp.org/> (accessed February 17, 2008).
- Education Daily. 2008. Teacher evaluations deemed dismal. January 9, 2008.
- The Education Trust 2005. *Gaining traction, gaining ground*. Washington, DC: Author.



- Elmore, R. 2002. *Testing trap*. *Harvard Magazine*. September–October 2002.
- Garet, M. S., A. C. Porter, L. Desimone, B. B. Birman, and K. S. Yoon. 2001. What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal* 38, no. 4.
- Geortz, M. and D. Massell. 2005. *Holding high hopes: How high schools respond to state accountability policies*. Philadelphia: Consortium on Policy Research in Education.
- Goldhaber, D. 2006. *Everyone's doing it, but what does teacher testing tell us about teacher effectiveness?* Washington, DC: Urban Institute.
- Goldhaber, D. and D. Brewer. 2000. *Does teacher certification matter? High school teacher certification status and student achievement*. *Educational Evaluation and Policy Analysis* 22 (2).
- Gordon, R., T.J. Kane, and D.O. Staiger. 2006. *Identifying effective teachers using performance on the job*. Washington, DC: The Brookings Institution.
- Gore, L. 2007. *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Hanushek, E. A., J. F. Kain, and S. G. Rivkin. 2005. *Teachers, schools, and academic achievement*. *Econometrica* 73 (2)
- Harris, D. N. 2007. *The policy uses and "policy validity" of value-added and other teacher quality measures*. Madison, WI: Teacher Quality Research.
- Harris, D.N., and T.R. Sass. 2007. *What makes for a good teacher and who can tell?* Paper presented at the 2007 summer workshop of the National Bureau of Economic Research (NBER).
- Heller, R. and C. Greenleaf. 2007. *Literacy instruction in the content areas: Getting to the core of middle and high school improvement*. Washington, DC: Alliance for Excellent Education.
- Hirsh, S., and J. Killion. 2007. *The learning educator: A new era for professional learning*. Oxford, OH: National Staff Development Council.
- Hoff, D. 2007. Panel report is latest Rx for NCLB. *Education Week* 26, no. 24. Bethesda, MD: Editorial Projects in Education.
- Jacob, B. A., and L. Lefgren. 2005. *Principals as agents: Subjective performance measurement in education*. [http://ksgnotes1.harvard.edu/Research/wpaper.nsf/rwp/RWP05-040/\\$File/rwp\\_05\\_040\\_jacob.pdf](http://ksgnotes1.harvard.edu/Research/wpaper.nsf/rwp/RWP05-040/$File/rwp_05_040_jacob.pdf) (accessed on January 31, 2008).
- Jepsen, C. and S. Rivkin. 2002. *Class size reduction, teacher quality, and academic achievement in California public elementary schools*. San Francisco: Public Policy Institute of California.
- Kane, T. and D.O. Staiger. 2005. *Using imperfect information to identify effective teachers*. Unpublished Paper. Los Angeles: School of Public Affairs, University of California—Los Angeles.
- Killion, J. 2002. *What works in the high school: Results-based staff development*. Oxford, OH: National Staff Development Council.
- Koppich, J., C. Asher, and C. Kerchner. 2002. *Developing careers, building a profession: The Rochester career in teaching plan*. Washington, DC: National Commission on Teaching & America's Future.
- Ladson-Billings, G. 1999. *Preparing teachers for diversity*. In L. Darling-Hammond and G. Sykes (eds.) *Teaching as the learning profession*. San Francisco: Jossey-Bass.
- Marks, A. and R. Wright. 2002. *Defining intrinsic and extrinsic motivators of continuing professional education*. Paper presented at the Annual Meeting of the Eastern Educational Research Association.
- Mazzeo, C. 2003. *Improving teaching and learning by improving school leadership*. Washington, DC: National Governors Association.
- McCaffrey, D.F., J. R. Lockwood, D. M. Koretz, and L. S. Hamilton. 2003. *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND Corporation.
- McLaughlin, M., and J. Talbert. 2001. *Professional communities and the work of high school teaching*. Chicago: University of Chicago Press.
- Monk, D. 1994. Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review* 12(2).
- National Clearinghouse for English Language Acquisition. 2007. *The growing numbers of Limited English Proficient students: 1995/96–2005/06*. Washington, DC: U.S. Department of Education.
- National Council on Teacher Quality. 2007. *State teacher policy yearbook*. Washington, DC: Author.
- National Institute for Excellence in Teaching. 2008. *Multiple career paths*. <http://www.talentedteachers.org/tap.taf?page=element1> (accessed February 17, 2008).





- National Research Council. 2004. *Engaging schools*. Washington, DC: National Academies Press.
- . 2000. *How people learn*. Washington, DC: National Academies Press.
- New Teacher Project. 2007a. *Hiring, assignment, and transfer in Chicago public schools*. New York: Author.
- New Teacher Project. 2007b. *Hiring, assignment, and transfer in Portland public schools*. New York: Author.
- New Teacher Project. 2007c. *Hiring, assignment, and transfer in Milwaukee public schools*. New York: Author.
- Noell, G. H., B. A. Porter, and R. M. Patt. 2007. *Value-added assessment of teacher preparation in Louisiana: 2004–2006*. <http://www.regents.state.la.us/Academic/TE/2007/VAA%20TPP%20Technical%20Report%2010-24-2007.pdf> (accessed March 13, 2008).
- Noell, G. H. 2004. *Technical report of: Assessing teacher preparation program effectiveness: A pilot examination of value added approaches*. [http://asa.regents.state.la.us/TE/technical\\_report.pdf](http://asa.regents.state.la.us/TE/technical_report.pdf) (accessed on January 31, 2008).
- Oklahoma State Regents for Higher Education 2008. *High school indicators project*. <http://www.okhighered.org/studies-reports/indicators/college-going/> (accessed February 17, 2008).
- Perie, M., S. Marion, B. Gong, and J. Wurtzel. 2007. *The role of interim assessments in a comprehensive assessment system: A policy brief*. Washington, DC: The Aspen Institute.
- Rivkin, S. G., E. A. Hanushek, J. F. Kain. 2002. *Teachers, Schools and Academic Achievement*. Dallas: University of Texas-Dallas Texas Schools Project.
- Sanders, W. and Rivers, J. 1996. *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville: University of Tennessee Value-added Research and Assessment Center.
- Schacter, J., T. Schiff, Y. Thum, C. Fagnano, M. Bendotti, L. Solmon, K. Firetag, and L. Milken. 2002. *The impact of the Teacher Advancement Program on student achievement, teacher attitudes, and job satisfaction*. Santa Monica, CA: National Institute for Excellence in Teaching.
- Short, D. and S. Fitzsimmons. 2007. *Double the work: Challenges and solutions to acquiring language and academic literacy for adolescent English language learners, a report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Solmon, L., L. T. White, D. Cohen, and D. Wood. 2007. *The effectiveness of the Teacher Advancement Program*. Santa Monica, CA: National Institute for Excellence in Teaching.
- Springer, M., D. Ballou, and A. Peng. 2008. *Impact of the Teacher Advancement Program on student test score gains: Findings from an independent appraisal*. Nashville, TN: National Center on Performance Incentives.
- Stewart, B.E. 2006. *Value-added modeling*. New York: Carnegie Corporation of New York.
- Supovitz, J.A., and J. B. Christman. 2003. *Developing communities of instructional practice*. Philadelphia: Consortium for Policy Research in Education.
- Supovitz, J.A. and V. Klein. 2003. *Mapping a course for improved student learning*. Philadelphia: Consortium for Policy Research in Education.
- Teacher Quality Partnership. 2008. *Teacher quality partnership*. <http://www.teacherqualitypartnership.org/> (accessed January 24, 2008).
- Toch, T., and R. Rothman. 2008. *Rush to judgment: Teacher evaluation in public education*. Washington, DC: Education Sector.
- Walsh, K. and C. Tracy. 2004. *Increasing the odds*. Washington, DC: National Council on Teacher Quality.
- Wenglinsky, H. 2002. The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives* 10 (12).

