

Key Issue:

Using Value-Added Models to Identify and Support Highly Effective Teachers

All resources contained within the TQ Tips & Tools documents have been reviewed for their quality, relevance, and utility by TQ Center staff and three content-area experts. These experts usually have a policy, practice or research background. The strategies and resources are provided to help regional comprehensive center and state education agency staff to be aware of the initiatives, programs or activities taking place in other settings. Our provision of the links to these resources should not be considered an endorsement but a qualified suggestion that they be considered as an option to study and/or pursue given the needs and context of the inquiring region, state, or district. Evidence of the impact of initiatives, programs or activities is provided where available or appropriate.

May 2008

Laura Goe, Ph.D.
ETS



1100 17th Street NW, Suite 500
Washington, DC 20036-4632
877-322-8700 • 202-223-6690
www.tqsource.org

Copyright © 2008 National Comprehensive Center for Teacher Quality, sponsored under government cooperative agreement number S283B050051. All rights reserved.

This work was originally produced in whole or in part by the National Comprehensive Center for Teacher Quality with funds from the U.S. Department of Education under cooperative agreement number S283B050051. The content does not necessarily reflect the position or policy of the Department of Education, nor does mention or visual representation of trade names, commercial products, or organizations imply endorsement by the federal government.

The National Comprehensive Center for Teacher Quality is a collaborative effort of Education Commission of the States, ETS, Learning Point Associates, and Vanderbilt University.

2875_05/08



Contents

	Page
Scenario.....	1
Definition of Value-Added Measures.....	2
Benefits of Using Value-Added Measures	3
Tips and Cautions	5
Strategy 1: Use Value-Added Measures as One Way to Identify Effective and Ineffective Teachers	8
Strategy 2: Use Value-Added Scores as a Part of a Teacher’s Performance Assessment.....	11
Strategy 3: Use Value-Added Scores to Help Make Differentiated Pay Decisions	12
Strategy 4: Use Value-Added Measures in Teacher Certification Decisions	15
Strategy 5: Use Value-Added Measures to Improve Professional Development	16
Strategy 6: Use Value-Added Measures to Improve Preservice Preparation Programs	18
Strategy 7: Use Value-Added Models to Determine the Impact of Teacher Characteristics and Teaching Practices on Student Outcomes	20
Real-Life Example: Using Value-Added Data to Improve Teaching in Dallas	22
References.....	24

Scenario

Dr. Richards, a state superintendent of schools, was appointed to lead a task force mandated by the governor and legislature to design a performance-based system of teacher assessment. State lawmakers discussed using the system to ground a statewide pay-for-performance plan, target struggling teachers for mandatory remediation (or termination), and create incentives and other policies to help equalize the distribution of highly qualified, effective teachers throughout the state. Although it was not yet under discussion in his state, Dr. Richards knew from talking to his colleagues around the country that some of them were beginning to use student achievement data as the basis for evaluating the effectiveness of teacher preparation programs as well as state-level professional development efforts.

Dr. Richards was aware that the task force was asked to consider a politically sensitive and complex issue. Many teachers, for example, were opposed to using growth scores—scores that are generated by analyzing students' standardized test scores linked to their teachers—as a basis for decisions that would affect their careers. They demanded the use of “authentic assessments” such as teacher observations. On the other hand, many political leaders insisted that observation scores were too subjective, too cumbersome, and too unreliable as the basis for the kinds of policies they wanted to implement; instead, these political leaders were strong advocates for using growth data, and many favored the adoption of *value-added methods*—using statistical models with longitudinal student test-score information to determine what contributions specific teachers are making to their students' learning. Dr. Richards knew that value-added models were very complex but understood that they were basically the difference between students' actual test scores and their predicted test scores.

As a former middle school teacher in Tennessee (which pioneered the use of value-added methods in education) and now as the official overseeing K–12 education in his state, Dr. Richards could appreciate both sides of the issue. There was no question in his mind, however, that he and the task force would have to recommend the use of some type of quantitative data about teachers, though not necessarily as the sole basis for the policy decisions. Dr. Richards believed that value-added methods would likely be the way to go, but he was aware that the task force needed to learn about assessment and accountability, performance pay systems that other states and districts had adopted, and especially about the promises and challenges of using value-added approaches before making their final recommendations.

The first task force meeting was in four weeks. Dr. Richards wanted to collect as much information as he could to share with the broad group of stakeholders who were its members: teachers, school administrators, business leaders, state education officials, researchers, and teacher union representatives. Where should he begin to look for it?

Definition of Value-Added Measures

A value-added measure is the “contribution of various factors toward growth in student achievement” (Goldhaber & Anthony, 2003, p. 38). According to leading researchers in the field, value-added models can be thought of as “a collection of complex statistical techniques that use multiple years of students’ test score data to estimate the effects of individual schools or teachers” (McCaffrey, Lockwood, Koretz, & Hamilton, 2003, p. xi). There are two main ways in which value-added models are used in practice. The first is to evaluate schools for accountability purposes, and the second is to evaluate teachers in terms of their effectiveness relative to other teachers. For a helpful discussion of these two applications of value-added models, see the publication *Evaluating Value-Added: Findings and Recommendations From the NASBE Study Group of Value-Added Assessments* (National Association of State Boards of Education, 2005).

In its most simple form, the value-added measure as it is used for evaluating teachers is calculated as follows: Students’ previous test scores are used to create *predicted* test scores for a given year. The difference between the predicted and actual test scores are growth scores. Teachers’ contribution to students’ learning is determined by looking at the average of all of their students’ growth scores. The teachers are then ranked against other teachers within a district (or other unit of interest) according to how much they contributed to students’ growth, and this ranking is their value-added “score.” In some value-added models, only students’ prior achievement scores are used in the calculation; other models include students’ gender, race, and socioeconomic background; still others include information about teachers’ experience. With a value-added measure, teachers whose students performed about as well as predicted are considered “average” teachers, those whose students performed much better than predicted are considered “above average” or “highly effective,” and those whose students performed worse than expected are considered “below average.”

Benefits of Using Value-Added Measures

Value-added measures are relatively objective, since they attempt to consider only teachers' contributions to student learning.

Unlike classroom observations (which may be influenced by the observers' own beliefs about good teaching, the appearance of the classroom, the students' behavior, and other factors), value-added scores are free from the subjective judgments and impressions of evaluators. With proper training, observer bias can be minimized but not eliminated; with value-added models, there is no observer—only scores.

Value-added measures provide a useful way to look for evidence about which teacher qualifications and characteristics matter for student learning.

Reviewing value-added test scores for teachers and linking those scores to other information—such as a teacher's highly qualified status, past experience, and degrees—can reveal interesting information. For example, in several research syntheses, it has been determined that teachers with either certification in mathematics or a strong mathematics background contributed significantly more to their students' achievement test scores than did teachers without a strong mathematics background or without a mathematics certification—although there was no strong, consistent evidence of the impact of certification, experience, or educational background on student achievement in other subjects such as language arts, social studies, and science (Goe, 2007; Rice, 2003; Wayne & Youngs, 2003).

Analyzing value-added data is relatively inexpensive compared with other means of assessing teachers.

Creating the necessary infrastructure (such as unique identifiers for teacher and students, and data systems that link students with teachers over time) to collect and analyze data may be expensive initially, but after that, the cost of analyzing the data is relatively low compared to the cost of collecting and analyzing classroom observation data or developing and evaluating teacher portfolios. (See the Key Issue titled *Using Performance-Based Assessments to Identify and Support Highly Effective Teachers*, available at <http://www2.tqsource.org/strategies/het.asp>.) States currently are working to collect more data at the student- and teacher-level in order to comply with the demands of the No Child Left Behind (NCLB) Act demands. Thus, some of the work already should be accomplished. Sorting, linking, and warehousing this data for reporting purposes as well as for use in value-added teacher evaluation is difficult and initially costly, but linked student-teacher data can be profoundly useful for making evidence-based decisions and thus worth the investment.

Value-added measures focus exclusively on student learning—not on teaching practices that may or may not be linked to positive outcomes for students.

It is possible for teachers to be given a high score when evaluated with a teacher observation instrument but still have average or below-average impact on their students' learning. Observation instruments can rate teachers on their use of teaching practices that fit with experts'

beliefs about good teaching, but empirical evidence that specific teaching practices improve student learning is lacking (see Goe, 2007 for a review of the literature on teacher quality and teacher practices). This result may be in part because of the difficulty of measuring differences in teaching practices as they relate to standardized achievement outcomes (see Valli et al., 2004, for a discussion of these difficulties). Because value-added measures focus only on actual student progress, the extent to which teachers' practices reflect an instructional ideal is not part of the equation. In this view, teacher effectiveness is based on confidence that student test scores are valid and reliable indicators of student learning.

Value-added measures identify highly successful classrooms and teachers, creating opportunities to learn from those teachers.

There are considerable differences among teachers, even within the same school, in terms of how much their students learn (Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). It can be extremely helpful to new or struggling teachers to observe how highly effective teachers teach, but identification of outstanding teachers is often based on their reputation or the extent to which their practices reflect experts' conceptions of accomplished teaching. Value-added measures offer a way to identify those highly effective teachers whose practices actually contribute the most to student learning gains, thus creating a "learning lab" for colleagues and researchers. It would be particularly valuable to identify and examine the classrooms of teachers who are highly effective—as measured by value-added scores—with students who are at-risk for poor educational outcomes. Providing preservice teachers an opportunity to learn from these effective teachers or giving new or struggling teachers opportunities to observe in these classrooms could provide valuable professional development.

Tips and Cautions

Even though value-added models are useful in tracking student progress over time, there are limitations and complexities involving these methods and the resultant data, particularly when used as the sole measure of teacher effectiveness. Educators and policymakers should give careful consideration to these concerns before committing to using value-added methods in ways that may not be warranted, given the current state of our understanding about the methodology and theory. The bottom line is that value-added data are limited in what they can tell about teacher quality and effectiveness; these data are more useful when supplemented by other measures and sources of evidence.

Following are several limitations to the use of value-added measures of teacher effectiveness. Officials should use extreme caution when making high-stakes decisions about teachers using these measures.

- **Difficulty in Determining Impact.** Value-added measures alone are insufficient to determine the impact of specific teaching practices on student progress and to guide instructional improvement. There is little that teachers (or administrators) can learn simply from seeing teachers' value-added scores because these numbers give no indication of what the teacher might be doing right or wrong that impacts student learning. Value-added scores can provide some direction for teacher learning, however. For example, knowing that a teacher's students are demonstrating slower growth in understanding fractions but faster growth in understanding multidigit multiplication helps the teacher know that his or her instruction of fractions needs to improve, but this situation provides no information on *how* to improve in that area. For beginning teachers, in particular, it is important to ensure that they are given detailed and specific information about the strengths and weaknesses of their instructional practices and targeted suggestions for improvement. Observations or other forms of performance-based assessments conducted by experts, followed by discussion of the results, are most likely to be helpful to new or struggling teachers who need guidance.
- **Difficulty in Isolating the Contributions of an Individual Teacher.** Many researchers have expressed concerns that it is difficult, if not impossible, to isolate an individual teacher's contributions to student achievement from a number of other factors that can affect it: classroom and school characteristics, students' peers, student mobility, curriculum quality, access to materials and resources, other teachers (e.g., Braun, 2005; Glass, 2004; Kupermintz, 2003; McCaffrey, Koretz, Lockwood, & Hamilton, 2004). Even the specific grade level can be a factor. Teachers who teach students in the first year of middle school or high school may have lower value-added scores because there is a "building change effect" that tends to lower achievement gains dramatically for some students (Sanders & Horn, 1998). Thus, comparing teachers across grades may be problematic. Moreover, comparing teachers across grades assumes that tests have sound "vertical scales" (i.e., that what is being measured is essentially the same across grade levels).
- **Difficulties With Methodological Issues.** Methodological issues in generating value-added scores can compromise teachers' value-added scores and thus their classification as effective or ineffective (for discussions of these issues, see Amrein-Beardsley, 2008;

Glass, 2004; McCaffrey, Koretz, et al, 2004). Some research suggests that students who have a series of effective teachers will have better achievement in subsequent years (Sanders & Horn, 1998), so some of the gains that students experience in a given year may be the effects of prior years of effective teachers and thus not attributable to the current teacher. The effects of bad teaching also may be cumulative, and a low score for a teacher in a low-performing school may in part reflect poor teaching by others in previous years. Although some value-added models attempt to control for students' prior teachers' contributions, it is not clear how well the models succeed in isolating the current teacher effects.

- **Incomplete Student Data and Small Sample Sizes.** Incomplete student data as well as small sample sizes may skew a teacher's value-added score. Given that the most challenging schools often have highly mobile student populations, it is likely that value-added estimates are affected. The more students' scores are available, the more likely it is that the score will accurately reflect a teacher's contribution. A teacher's score is likely to be more accurate when calculated with 20 students' achievement growth information than with only 10 students' growth information. Teachers with fewer students' growth information used in calculations of value-added scores may be disadvantaged since their scores may appear to be too high or too low depending on random fluctuation.
- **Relativity.** Teachers' value-added scores are not absolute or based on independent criteria but are relative—that is, dependent upon the teacher comparison group and the particular value-added method employed. This situation complicates efforts to use value-added measures as the basis for *statewide* evaluation or compensation systems. Teachers' value-added scores reflect a ranking based on the mean score of the comparison group, which is almost always limited to the teachers in a particular district (Kupermintz, 2003). The result is that an “average” teacher in a district with a high mean value-added score (a district where students have made notable achievement gains) may be as good as, or even better than, a “highly effective” teacher in a district with a low mean score (a district where students have failed to make substantial achievement gains). In addition, value-added models used by different states or districts may be built on very different assumptions. For example, models often differ on the length of time a student must be in a school for their examination score to count in a teacher's value-added calculation. This means that the value-added score received by the teacher of a particular class in one district could be higher or lower if the teacher taught the very same class in another district solely because the districts use different value-added models.
- **Parameters of Good Teaching.** The success of teachers based on value-added data does not necessarily reflect accomplished teaching. For example, a teacher whose class shows excellent gains in a given year and who thus has a high value-added score may be either (1) good at helping students explore and learn a range of worthwhile knowledge within a broad curriculum, or (2) good at preparing students for the standardized test upon which the value-added score is based. Thus, a teacher who focuses on a narrow curriculum that is closely aligned with a standardized test may score higher than a gifted and inspiring teacher whose students receive a broader, richer curriculum that includes material that is not tested. In addition, many worry that teacher appraisal systems based solely on value-added models may provide incentives for “teaching-to-the-test” or even cheating.

- **Inability to Use Value-Added Methods.** A great many teachers cannot be evaluated using value-added methods. Assessment and instructional policies in states or districts can result in limitations in the scope and legitimate application of value-added data. Although effective early elementary teachers are critical, those whose students are not tested cannot be assessed using value-added scores, since the scores rely on students' previous testing records. School subjects that are not tested cannot be used to generate value-added scores, so teachers of art, music, physical education, and so on will not receive value-added scores. Teachers with less traditional assignments (e.g., those who work in pull-out programs) may not spend enough time with students to include the students in their value-added score. Team teaching or other arrangements in which instruction is shared among teachers also make it difficult to identify the specific contributions—and thus the value-added scores—of individual teachers. In addition, value-added models used in some states and districts explicitly exclude first-year teachers. These factors pose a challenge to the evenhandedness of statewide or districtwide programs that evaluate or compensate teachers on the basis of value-added data. They also make it difficult to use value-added data fairly and accurately in making school-level decisions, such as the awarding of whole-school performance bonuses.
- **Privacy Issues.** Unique teacher and student identifiers are essential to create links between teachers and students for calculating value-added scores. A growing number of states have linked student-teacher data that can be used for value-added methods, but many states struggle with privacy issues related to using these identifiers. Some states, such as California, have only recently begun to establish policies and procedures to assign unique identifiers and use them for administrative and research purposes. The Data Quality Campaign (<http://www.dataqualitycampaign.org>) has worked diligently to advocate for the development of statewide longitudinal data systems. It offers many useful ideas about how to design and develop such systems and also provides advocacy materials for the support and adoption of unique identifiers.
- **Inadequacy of Standardized Tests.** Finally, standardized tests are incomplete measures of student learning. Some exams have better psychometric properties than others, some are more rigorous than others, and some require higher order thinking and some do not. Tests will inevitably be limited in their coverage of any subject matter domain. Moreover, teachers contribute to other valued student outcomes that are more difficult to measure—for example, socioemotional wellness, civic engagement, moral character, open-mindedness, and motivation for continued learning. A teacher appraisal system based solely on value-added models would exclude these other important contributions.

Strategy 1: Use Value-Added Measures as One Way to Identify Effective and Ineffective Teachers

With value-added methods, teachers can be placed on a continuum based upon how much more or less their students gained on their achievement scores than the students' prior achievement results would have predicted. Where to draw the "cut points" in that distribution for designating "effective" and "ineffective" teachers, however, is a matter that requires a decision involving many stakeholders, including representatives of teacher organizations. There has been a history of resistance among teachers to using value-added scores as a measure of teacher effectiveness because of the limitations and methodological complexities of value-added assessment. As an example, Tennessee introduced value-added assessment in the 1980s; however, until recently these scores were simply given to teachers and their principals without any expectations of further action or any stakes, such as compensation decisions, attached to them. Currently, they may be considered as one component in the overall evaluation of a teacher.

In a number of districts and states, value-added scores exist but are not yet being used in ways that affect or engage teachers. For example, McCaffrey & Hamilton (2007) found that Pennsylvania schools were given access to classroom-level value-added data but little use was made of the data to change instruction. In Ohio, teachers in some schools also receive value-added information about their classrooms and teachers can use the data to see whether they are more effective with high-achieving students or low-achieving students; however, whether they use that information to actually change instruction is not clear. Also, there is little other information in value-added scores that can be used instructionally. In other places, various kinds of agreements have been achieved that allow the use of value-added scores for limited administrative purposes. One modest but important way that value-added scores have been used in Tennessee is to determine the distribution of effective teachers within the state (Tennessee Department of Education, 2007). Using this information, Tennessee can target technical assistance to those districts that need to find ways to improve the effectiveness of their teachers. Similarly, value-added scores can be used to identify clusters of particularly effective teachers within a district or even within individual schools.

Resource 1: CCSSO's *Growth Models Primer*

Blank, R., & Cavell, L. (2005, March 16). *Growth models primer* [Memorandum to chief state school officers]. Washington, DC: Council of Chief State School Officers. Retrieved May 9, 2008, from <http://www.ccsso.org/content/pdfs/Growthmemo.pdf>

This four-page memo was designed to bring chief state school officers up-to-speed on the value-added approach. The authors state, "It discusses growth models in relation to status and improvement models and summarizes the following: (1) types of growth models and their purpose(s); (2) the advantages and disadvantages of growth models; (3) challenges to implementation and resource requirements; (4) use of growth models under No Child Left Behind (NCLB); and (5) policy questions that may shape state decision making regarding the use of growth models for accountability" (p. 1).

Resource 2: *Evaluating Value-Added*

National Association of State Boards of Education. (2005). *Evaluating value-added: Findings and recommendations from the NASBE Study Group on value-added assessments*. Alexandria, VA: Author.

This summary provides a useful description of the different ways that value-added measures can be used, including school accountability, teacher accountability and evaluation, school improvement, teacher training, and as a component of adequate yearly progress calculations. The summary also provides the group's findings and recommendations on both the uses of and limitations of value-added measures.

Resource 3: *Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models*

Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service. Retrieved May 9, 2008, from <http://www.ets.org/Media/Research/pdf/PICVAM.pdf>

Although Braun is a well-known statistician, he approaches the topic of using value-added to evaluate teachers from a policy-oriented point of view. The result is a very accessible and easy-to-read report that should help clarify some of the limitations of this method. Of note, Braun discusses why what is measured with value-added is actually "classroom effects" rather than "teacher effects."

Resource 4: *The Promise and Peril of Using Value-Added Modeling to Measure Teacher Effectiveness*

McCaffrey, D. F., Koretz, D., Lockwood, J. R., & Hamilton, L. S. (2004). *The promise and peril of using value-added modeling to measure teacher effectiveness* (Research Brief No. RB-9050-EDU). Santa Monica, CA: RAND Corporation. Retrieved May 9, 2008, from http://www.rand.org/pubs/research_briefs/2005/RAND_RB9050.pdf

The RAND researchers have conducted a great deal of research in this area and are considered among the leading experts in the country. This research brief does a good job of describing and summarizing both technical and practical issues involved in using value-added measures, including how various statistical strategies yield different results, problems related to inadequate data, issues related to using achievement test scores as outcomes, and sampling error. It also provides a quick "bottom line" paragraph that cautions against using value-added for high-stakes decisions regarding teachers.

Resource 5: *A Review of Value-Added Models*

Hibpshman, T. L. (2004). *A review of value-added models*. Frankfort: Kentucky Education Professional Standards Board. Retrieved May 9, 2008, from <http://www.kyepsb.net/documents/Stats/Journals/Heterogeneity%20of%20regression.pdf>

This review does a great job of summarizing the important information about the basic types of models used for value-added measures as well as what researchers have concluded about the differences among the different models. The author also asks, and answers, the question “What is the best model, and how should we use it in Kentucky?” For states asking themselves similar questions, this review may be very helpful. The only drawback is that in spite of the author’s attempts at making the statistical discussions “friendly,” it is still somewhat technical.

Resource 6: *Sizing Up Test Scores*

Ballou, D. (2002). Sizing up test scores. *Education Next*, 2(2). Retrieved May 9, 2008, from <http://www.hoover.org/publications/ednext/3365706.html>

Education Next published an “opinion” piece by Dale Ballou in 2002 that is still highly relevant when considering using value-added measures for teacher accountability. Ballou describes three considerations that make value-added assessments problematic for high-stakes teacher evaluation, although he does suggest that value-added results can be useful to inform policy discussions. The three considerations are: (1) methods of testing that are not completely accurate for measuring student gains; (2) other factors beside teacher or school quality that may influence student gains; and (3) different student ability levels, which make results difficult to interpret.

Resource 7: Identifying Effective Teachers in Tennessee

Tennessee Department of Education. (2007). *Tennessee’s most effective teachers: Are they assigned to the schools that need them most?* (Research Brief). Nashville, TN: Author. Retrieved May 9, 2008, from http://tennessee.gov/education/nclb/doc/TeacherEffectiveness2007_03.pdf

The state of Tennessee has used its Value-Added Assessment System (TVAAS) to investigate whether there are disparities in the effectiveness of teachers between the state’s high-poverty and more affluent schools. This summary of the study shows how the state conducted this analysis and provides insight into how TVAAS works. This study overview should be particularly helpful to states that are considering using value-added for research purposes and want to know what research questions other states have attempted to answer.

Resource 8: Benwood Initiative, Chattanooga, Tennessee

The Achievement Alliance. (n.d.). *It’s being done: The Benwood Initiative, Chattanooga, Tennessee*. Washington, DC: Author. Retrieved May 9, 2008, from <http://www.achievementalliance.org/files/Benwood.pdf>

The Hamilton County School District in Tennessee, which includes the city of Chattanooga, took advantage of the state’s Value-Added Assessment System to develop a major school reform initiative in the district’s poorest performing schools, called the Benwood schools. Using value-added data, officials identify those teachers in the districts who are the most effective and offer opportunities to earn salary bonuses if they will take positions in the Benwood schools.

Strategy 2: Use Value-Added Scores as a Part of a Teacher's Performance Assessment

Value-added scores should not be the only measure of teachers' effectiveness, quality, or performance. A number of states and districts use value-added data as an important part of a more comprehensive teacher performance assessment system, perhaps because they recognize that these measures should not be used in isolation from other indicators of teaching quality and effectiveness. For example, teachers contribute to students' learning, growth, social skills, self-esteem, and citizenship, but these important contributions are not measured by standardized achievements tests. Moreover, teachers contribute to the overall school climate and effectiveness by collaborating with their colleagues, conducting action research, and providing other professional services. Many performance-based assessment systems have been designed to capture evidence of teachers' contribution to both their students' development as well as to their school and profession. (These systems are described in the Key Issue titled *Using Performance-Based Assessments to Identify and Support Highly Effective Teachers*, available at <http://www2.tqsource.org/strategies/het.asp>.) The following resources describe several performance-based assessment systems that have a value-added component.

Resource 9: Teacher Advancement Program

Website: <http://www.talentedteachers.org/tap.taf?page=whatistap>

For districts and states thinking about designs for compensation systems related to teacher effectiveness, the Teacher Advancement Program (TAP) provides a useful real-world example. TAP is an alternative model of teacher recruitment, career development, evaluation, and compensation that was developed by the Milken Family Foundation and is now operated by the National Institute for Excellence in Teaching. In TAP, the value-added gains of a teacher's pupils are one of the factors included in the evaluation of both individual teachers and the school as a whole; these gains also are a factor in the determination of the teacher's total compensation package. Professional development is a key part of the TAP strategy. TAP has been implemented in more than 60 districts and 180 schools throughout the United States.

Resource 10: Houston's ASPIRE Program

Website: <http://portal.battelleforkids.org/ASPIRE/Home.html?sflang=en>

The Houston Independent School District has launched a new, comprehensive, continuous improvement effort to align teaching practice, evaluation, professional development, career advancement, and recognition. Called ASPIRE ("Accelerating Student Progress. Increasing Results and Expectations"), the initiative uses value-added data as a core indicator of teachers' strengths and weaknesses and also as an indicator of the effectiveness of the various programs and practices the district uses to improve the quality of instruction. Teachers who are identified as successful on the basis of their value-added scores will receive monetary rewards as part of Houston's differentiated pay plan, and they also will be considered for career advancement opportunities.

Strategy 3: Use Value-Added Scores to Help Make Differentiated Pay Decisions

Using value-added scores as a basis for teacher compensation systems poses significant challenges, particularly when such compensation is teacher based rather than school based. Although few districts use value-added scores alone to determine salary bonuses, some districts have reached agreements with teachers and teachers' representatives to use value-added scores in combination with other measures to differentiate teacher pay. One thing most of these approved pay plans have in common is some way to include teachers who do not receive value-added scores. In some states and districts, bonuses are awarded on a whole-school basis to those schools in which the student population's average growth exceeds value-added projections. Many teachers prefer these schoolwide awards because this approach eliminates any sense that teachers within a school are competing with or comparing themselves to one another. Schoolwide awards raise other issues of fairness, however, since the basis of the awards is the performance of teachers in the core subjects that are tested, rather than the performance of all teachers.

Resource 11: Center for Educator Compensation Reform

Website: <http://cecr.ed.gov/>

The Center for Educator Compensation Reform (CECR) is a partnership of five organizations funded by the U.S. Department of Education to raise awareness of alternative and effective strategies for educator compensation reform. The website has information, tools, and resources to support federal Teacher Incentive Fund (TIF) grantees and other educators, stakeholders, and policymakers considering compensation reform. Several of the profiled TIF grantees use models that promote differentiated staffing opportunities in which teachers are recognized and rewarded for serving as lead teacher, mentor, and other positions.

Resource 12: National Center on Performance Incentives

Website: <http://www.performanceincentives.org/index.asp>

This center, housed at Vanderbilt University in Tennessee, seeks to find answers to the question "Do financial incentives for teachers, administrators, and schools affect the quality of teaching and learning?" Through research initiatives and evaluations of existing pay-for-performance plans, the center seeks to inform education policymakers and ultimately contribute to better teaching and learning. Besides policy briefs and publications, it provides useful definitions of various pay-for performance initiatives at its webpage on national incentive-pay initiatives (http://www.performanceincentives.org/news_events/detail.asp?pageaction=ViewSinglePublic&LinkID=46&ModuleID=28&NEWSPID=1).

Resource 13: Teacher Advancement Program

Website: <http://www.talentedteachers.org/tap.taf?page=whatistap>

As described in Resource 9, the Teacher Advancement Program (TAP) is an alternative model of teacher recruitment, career development, evaluation, and compensation that was developed by the Milken Family Foundation and is now operated by the National Institute for Excellence in

Teaching. In TAP, the value-added gains of a teacher's pupils are one of the factors included in the evaluation of both individual teachers and the school as a whole. These gains also are a factor in the determination of the teacher's total compensation package. TAP has now been implemented in more than 60 districts and 180 schools throughout the United States.

Resource 14: Florida's STAR Plan

Florida Department of Education. (2007). *Florida performance pay guidance: 2007–2008 and beyond*. Tallahassee, FL: Author. Retrieved May 9, 2008, from <http://www.fldoe.org/PerformancePay/pdfs/MeritAwardProgram.pdf>

STAR stands for "Special Teachers Are Rewarded." It is a Florida statewide plan that enables districts opting in to reward outstanding teachers based on their documented classroom performance. One of the principal factors considered in teachers' eligibility for the bonus is the value-added gains of their students on statewide assessments. Florida is one of the few states whose statewide longitudinal data system is sufficiently advanced to conduct value-added comparisons anywhere in the state.

Resource 15: Benwood Initiative, Chattanooga, Tennessee

The Achievement Alliance. (n.d.). *It's being done: The Benwood Initiative, Chattanooga, Tennessee*. Washington, DC: Author. Retrieved May 9, 2008, from <http://www.achievementalliance.org/files/Benwood.pdf>

Using value-added data, officials identify Hamilton County School District teachers who are the most effective and offer them the opportunity to earn salary bonuses if they will teach in the Benwood schools, the district's poorest performing schools. If these teachers succeed in having high value-added achievement gains for their students, Chattanooga's mayor and the district award them significant bonuses that can total up to \$6,000.

Resource 16: North Carolina's ABCs of Public Education

Website: <http://abcs.ncpublicschools.org/abcs/>

North Carolina has implemented a school accountability system that uses students' value-added achievement gains, aggregated by school, to identify schools that are effective or ineffective. If a school either meets or exceeds state-determined expectations for student achievement growth, its entire instructional staff, including teacher assistants, is awarded a performance bonus.

Resource 17: Dallas's Performance Pay Program

Dallas Independent School District. (2008). *Performance pay manual for teachers and campus employees*. Dallas, TX: Author. Retrieved May 9, 2008, from <http://www.dallasisd.org/performancepay/pdf/PerfPayManual0708.pdf>

The Dallas Independent School District is in the process of implementing a new performance pay system for teachers that uses value-added data and a number of other factors to determine the

amount of incentive pay that teachers will receive. The value-added data include a score both for an individual teacher's classroom performance and for the performance of a teacher's entire school, and bonuses are awarded to both outstanding individual teachers and outstanding schools.

Resource 18: Mission Possible, Guilford County School District, North Carolina

Website: http://www.gcsnc.net/depts/mission_possible/background.htm

In 2005–06, the Guilford County School District in North Carolina implemented a teacher compensation bonus program called Mission Possible in its lowest performing schools. Teachers in some subjects are offered significant annual bonuses that range up to \$10,000 for Algebra I teachers if they agree to teach in the Mission Possible schools. In addition, many teachers in the schools are eligible for performance bonuses, which are determined solely on the basis of value-added data. Teachers whose students exceed the average growth for the district receive a \$2,500 bonus for the year, and teachers whose students show achievement growth greater than 1.5 times the district average receive a \$4,000 bonus.

Strategy 4: Use Value-Added Measures in Teacher Certification Decisions

One potential use of value-added measures that is particularly troubling for teachers is making value-added scores part of the evidence in making licensure or certification decisions. Because such usage has high-stakes implications, it is unlikely to be adopted without considerable conflict. Research on the validity and reliability of value-added scores for high-stakes decisions is still in its infancy, and prominent researchers in the field warn against such high-stakes uses of value-added results (e.g., Braun, 2005; McCaffrey, Koretz, et al., 2004), so their use in licensure and certification decisions may be premature. However, at least one state—Georgia—currently uses value-added data in decisions about voluntary advanced-level certification.

Resource 19: Georgia Master Teacher Certification Program

Website: <http://www.georgiamasterteachers.org/MasterTeacher.aspx> and http://www.gamasterteachers.org/MT_about.aspx

This program provides a Master Teacher Certificate to Georgia public school “clear renewable” certified teachers with at least three years of experience who consistently demonstrate excellence in the classroom that is linked to gains in student achievement. Candidates who teach in subjects that are tested on the statewide student assessment are required to provide sufficient identifying information so that state officials can use student assessment results to calculate the learning gains of the students in the classes that teachers have taught. The first cohort of Georgia Master Teacher Certification program applicants were selected in 2006.

Strategy 5: Use Value-Added Measures to Improve Professional Development

Effective professional development should improve teachers' instruction and, in turn, improve student learning and student achievement test scores. Tracking teachers who have participated in specific professional development opportunities or who have gone through specific trainings can be useful in trying to determine the contribution of these experiences to improved instruction. Classroom observations, journals, portfolios, samples of teachers' assignments, and other evidence can be used to determine if teachers' practices changed as a result of their professional development, but these changes may or may not translate into improved student learning. Value-added data may be able to reveal whether these professional development experiences truly have a positive impact on student achievement.

Resource 20: Houston's ASPIRE Program

Website: <http://portal.battelleforkids.org/ASPIRE/Home.html?sflang=en>

As noted previously, the Houston Independent School District launched a new, comprehensive continuous improvement effort to align teaching practice, evaluation, professional development, career advancement, and recognition. Called ASPIRE ("Accelerating Student Progress. Increasing Results and Expectations"), the initiative will use value-added data as a core indicator of teachers' strengths and weaknesses and also of the effectiveness of the various programs and practices that the district may use in efforts to improve the quality of instruction.

Resource 21: Analysis of Professional Development in Dallas

Babu, S., & Smith, G. (1995). *Can staff make a difference in school effectiveness?* Dallas, TX: Dallas Public Schools. Retrieved May 9, 2008, from http://www.dallasisd.org/inside_disd/depts/evalacct/research/articles/Babu-Can-Staff-Development-Make-a-Difference-1995.pdf

Not long after Dallas implemented its value-added system in the early 1990s, the district undertook a series of studies to see if the system could be used to determine the practices of schools and their faculty that value-added methods identified as effective. One of the subjects of study was the professional development practices, and researchers looked for differences in staff development between high-performing and low-performing schools. The study of professional development was inconclusive, but the methodology used in the study is instructive and points to what might be possible using more powerful statistical tools.

Resource 22: REL Southwest's Issues and Answers Report on Effective Professional Development

Yoon, K. S., Duncan, T., Lee, S. W-Y, Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007–No. 33). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional

Assistance, Regional Educational Laboratory Southwest. Retrieved May 9, 2008, from http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf

REL Southwest's 2007 review of 1,300 studies that examine the impact of professional development on teacher effectiveness found only nine that meet the What Works Clearinghouse's rigorous standards of evidence. The nine studies, nevertheless, showed that high-quality professional development can impact teacher effectiveness. The report provides recommendations for better evaluating the link between professional development and student achievement.

Strategy 6: Use Value-Added Measures to Improve Preservice Preparation Programs

A number of efforts around the country are used to assess the effectiveness of teacher preparation programs—whether traditional or alternative certification programs—by linking what graduates learned in the programs with the performance of their K–12 pupils in the graduates’ first or second years of teaching. This area of research is important because there is little hard evidence indicating how the strategies that teachers learn in their programs translate into effective teaching. Unfortunately, documenting this relationship is difficult because it is not easy to separate out the impact of what’s taught in a program from the knowledge, skills, and other characteristics that teacher candidates bring with them to their preparation experience and that often influence which preparation programs they choose. In addition, teachers tend to go to certain districts and even certain schools when they graduate from their preparation programs. The match between teacher and school also may contribute to a teachers’ eventual effectiveness, and it is difficult to determine how much of the effectiveness came from the preparation program and how much came from a teaching assignment that was a good fit.

Resource 23: Louisiana’s Teacher Preparation Accountability System

Louisiana Board of Regents. (2003). *Revised blue ribbon commission on teacher quality / teacher preparation accountability system*. Baton Rouge, LA: Author. Retrieved May 9, 2008, from <http://www.regents.state.la.us/Academic/TE/overview.pdf>

As part of its teacher preparation accountability system, the state of Louisiana has decided to focus on the teaching outcomes of teacher preparation program graduates. Using a pilot project, researchers have validated a model that uses value-added K–12 student achievement data and eventually will be employed to assess and compare the effectiveness of graduates of all teacher preparation programs in the state. The pilot project already has revealed significant differences between some programs in the effectiveness of the teachers they graduate, which may signal significant differences in the quality of the programs if the program impact can truly be separated out from the other factors that can possibly influence the results obtained. The Louisiana Board of Regents is now considering adding the value-added program assessment to its program accountability system, which rates teacher preparation programs from exemplary to low-performing. An institution’s performance on this indicator would count for one third of its total score.

Resource 24: Teachers for a New Era

Website: <http://www.teachersforanewera.org/>

Teachers for a New Era is a multicampus effort, developed by the Carnegie Corporation of New York, which involves a core of 11 U.S. institutions and includes both undergraduate and postbaccalaureate teacher preparation programs. One of the key features of the initiative is that it seeks to use the value-added achievement gains of the K–12 pupils of preparation program graduates to confirm the effectiveness of the programs and, by implication, a general design for teacher education that all programs have been asked to adopt.

Resource 25: Ohio Teacher Quality Partnership

Overview: http://tqi.uwsa.edu/events/2007/TQP_Overview.doc

Ohio is engaged in a statewide effort, involving all 50 institutions that prepare teachers, to use K–12 value-added achievement data to identify effective preparation programs. The eventual aim of this Teacher Quality Partnership is to generate information for program improvement; currently, the work is primarily focused on developing methodologically rigorous research protocols and conducting pilot studies. Eventually, using value-added and other data, the project hopes to reveal what characteristics of teacher preparation programs—curriculum, mission, faculty, students, and other characteristics—account for the teaching effectiveness of their graduates.

Resource 26: Teacher Pathways Project

Website: <http://www.teacherpolicyresearch.org/TeacherPathwaysProject/tabid/81/Default.aspx>

The Teacher Pathways Project is one piece of a research partnership between Stanford University and the State University of New York at Albany. The project focuses on K–12 schools in New York City and the preparation programs that serve them, and it has used value-added achievement data to try to identify the relative effectiveness of the different programs or types of pathway (such as alternative route, traditional route) The project also has attempted to identify the key structural differences between programs identified as more effective and less effective based on the value-added scores of their graduates, but to date the studies have found more similarities than differences among them.

Strategy 7: Use Value-Added Models to Determine the Impact of Teacher Characteristics and Teaching Practices on Student Outcomes

Researchers are very interested in finding strong, clear evidence about which characteristics of teachers have the greatest positive effect on student achievement. A number of studies have investigated the impact of characteristics such as race and gender and of qualifications such as credentials, degrees, course taking, college selectivity, SAT or ACT scores, licensure or certification status, and teaching experience. A number of other studies, usually involving observations of teaching, have tried to determine the link between specific teaching practices and student learning gains. With the exception of teaching experience and knowledge of mathematics, neither the studies on teacher qualifications and characteristics nor those on teaching practices have been able to demonstrate a consistent relationship between these factors and student achievement (Goe, 2007).

With more districts and states beginning to collect value-added scores, however, it will become increasingly possible to tie those scores to teacher practices and characteristics and to identify stronger connections to student achievement. States and districts that want to conduct this type of research need to maintain longitudinal data files with unique identifiers for every student and every teacher. Because the various teacher factors are all likely to have different degrees of impact depending upon the grade level and subjects taught, student demographics, school characteristics, and geographic location, gathering information from many locations and across many contexts will help identify which factors matter the most.

It also is important to note that student outcomes other than achievement gains can be an important source of data about students' academic success, particularly for at-risk students. Teachers who are effective at helping at-risk students to stay in school, pass to the next grade, attend classes regularly, and stay out of trouble are contributing in ways that may be even more important than small gains on standardized tests. Unlike value-added measures, these types of measures are more likely to be done at the school level. (Because this Key Issue focuses on value-added models, it does not include any studies or actual practices that use such nonachievement outcomes to identify teachers who are effective in producing them. For a review of the research on measuring teacher effectiveness that includes measures other than value-added measures, see Goe, Bell, and Little, in press.)

Resource 27: *The Link Between Teacher Quality and Student Outcomes: A Research Synthesis*

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved May 9, 2008, from <http://www.tqsource.org/publications/LinkBetweenTQandStudentOutcomes.pdf>

This research synthesis is a comprehensive summary and analysis of the research on teacher quality, as determined by looking at how teacher qualifications, characteristics, and practices are

linked with student achievement in recent research. Only a small number of the studies included use value-added models, while most use less sophisticated measures.

Resource 28: Data Quality Campaign

Website: <http://www.dataqualitycampaign.org/>

The Data Quality Campaign is a partnership between a number of nationally prominent organizations and associations that seeks “to encourage and support state policymakers to improve the collection, availability and use of high-quality education data and to implement state longitudinal data systems to improve student achievement.” Its website has a great deal of information about data, assessment, and accountability as well as helpful resources.

Real-Life Example: Using Value-Added Data to Improve Teaching in Dallas

The Dallas Independent School District has a 15-year history of using value-added data to analyze the performance of the district's 160,000 students. An important part of its improvement strategy involves a commitment to improve instruction. Dallas has become one of the most aggressive U.S. districts in its use of value-added data to assess the performance of its 7,000 teachers and take appropriate action to enhance their effectiveness.

One of the keys to the district's ability to move forward in using value-added data has been the creation of a District Accountability Task Force, a committee of 27 individuals who are appointed by the board of education to oversee the district's accountability system. Representing multiple stakeholders, the task force includes parents, teachers, school administrators, business leaders, and school district officials who review and approve the selection and relative weighting of the variables that go into the accountability system and the performance rewards that are tied to it. The task force is responsible for ensuring that the system operates fairly and that no schools have an unfair advantage or disadvantage that would help or hurt them in the accountability calculations and the financial rewards tied to them.

The impetus for developing the education accountability system in Dallas was litigation in the 1980s contesting that educational opportunities for white students and minority students were not equal in the district. In response, the district implemented a system to measure the relative effectiveness of its schools, and it eventually created a commission that recommended the adoption of a value-added approach.

The Dallas value-added methodology has similarities to the Tennessee Value-Added Assessment System developed by William Sanders, but there are important differences. Most notably, the Dallas system tries to factor into the value-added calculation a number of what it calls "fairness variables," statistical adjustments for considerations such as ethnicity, socioeconomic status, prior student achievement, and school characteristics such as size and population density. The Tennessee system does not do this; rather, when estimating the value-added measure for an individual teacher, all scores (both past and future) for a particular student are used—meaning that controlling for individual student characteristics is largely unnecessary.

Initially, the Dallas value-added system was focused principally on identifying and motivating effective schools. Using a "school effectiveness index," the district calculated the performance scores of all its schools and then ranked the schools in order of effectiveness. Beginning in the 1990–91 school year, teachers and principals in the most effective schools were each given \$1,000 bonuses, and secretaries and custodians each received \$500. A few years later, the incentive policy was broadened to offer reduced bonuses to schools that performed well but not at the highest level.

It was clear from the beginning, however, that school effectiveness and student performance were tied to teacher effectiveness. The value-added data and analysis system that Dallas had implemented pointed clearly to the fact that students' academic achievement was fundamentally affected by the effectiveness of their teachers. Thus, in 1994–95, Dallas developed "teacher

effectiveness indexes” using value-added analysis. The results of the indexes at first were used only for internal school planning purposes. Soon, however, they became the basis for teacher evaluations that were used to identify the lowest performing teachers. These teachers were then required to take appropriate professional development based on a more detailed analysis of the weakness in their teaching practices.

Over the years, the district has developed a number of different teacher-related policies and initiatives based on the value-added data results. It has looked at the practices of the teachers in the schools identified as highest performing through the value-added data and has tried to learn what made them successful. One lesson was that the most effective teachers focused more than other teachers on higher order skills and challenging subject matter. These teachers had strong knowledge of the subjects they taught, avoided approaches that emphasized drill and repetition, and covered the entire curriculum.

The district also has attempted to use value-added data to guide professional development and align it with school improvement priorities. This effort has involved identifying high-performing and low-performing schools to see if there were differences in staff development practices that might account for the differences in school and teacher effectiveness. Beyond professional development, the district has tried to determine if there were differences between effective and ineffective schools in instructional practices, staff cohesiveness, administrative leadership and support, and a number of other school and student factors. There were some differences in professional development and instruction between high-performing and low-performing schools, but it was not clear if these differences could account for the variation in school performance because so many factors could play a role.

The Dallas district introduced two notable new-teacher initiatives tied to its value-added system starting in the 2007–08 school year. It launched an effort to improve the quality of teachers in its lowest performing schools by offering a \$6,000 annual salary bonus to teachers identified as highly effective who agreed to teach in them. Also, it began to pilot its Pay for Performance program in 59 schools. The pay program combines value-added data and a number of other factors to determine the amount of incentive pay that teachers will receive. The value-added data include a score both for an individual teacher’s classroom performance and for the performance of a teacher’s entire school, and bonuses are awarded to both outstanding individual teachers and outstanding schools.

It is difficult to determine the impact on student achievement of all of the value-added initiatives that Dallas has developed and implemented over the years in an effort to enhance the effectiveness of its teachers and their career development and compensation opportunities. Many other changes also have been introduced in the district in recent years, some due to the implementation of the No Child Behind Act, making it difficult to know what the specific impact is of any individual policies. What is certain, however, is that Dallas is breaking new ground in using value-added data to shape teacher policies and programs. The rest of the country will be watching to see how these efforts develop and hoping there ultimately will be an opportunity to assess the nature of their impact. Whatever happens and however successful or unsuccessful they might be, it is certain that policymakers, educators, and everyone else stand to learn a great deal from them.

References

- Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher*, 37(2), 65–75.
- Braun, H. I. (2005). Value-added modeling: What does due diligence require? In R. W. Lissitz (Ed.), *Value added models in education: Theory and applications* (pp. 19–39). Maple Grove: JAM Press.
- Glass, G. V. (2004). *Teacher evaluation* (Policy Brief). Tempe: Arizona State University, Education Policy Studies Laboratory. Retrieved May 9, 2008, from <http://epsl.asu.edu/epru/documents/EPSTL-0401-112-EPRU.doc>
- Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved May 9, 2008, from <http://www.tqsource.org/publications/LinkBetweenTQandStudentOutcomes.pdf>
- Goe, L., Bell, C., & Little, O. (in press). *Value-added models and other approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Goldhaber, D., & Anthony, E. (2003). *Teacher quality and student achievement* (Urban Diversity Series No. 115). New York: ERIC Clearinghouse on Urban Education. Retrieved May 9, 2008, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/1e/03.pdf
- Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis*, 25(3), 287–298.
- McCaffrey, D. F., & Hamilton, L. S. (2007). *Value-added assessment in practice: Lessons from the Pennsylvania Value-Added Assessment System Pilot Project* (Technical Report TR-506). Santa Monica, CA: RAND Corporation. Retrieved May 9, 2008, from http://rand.org/pubs/technical_reports/2007/RAND_TR506.pdf
- McCaffrey, D. F., Koretz, D., Lockwood, J. R., & Hamilton, L. S. (2004). *The promise and peril of using value-added modeling to measure teacher effectiveness* (Research Brief No. RB-9050-EDU). Santa Monica, CA: RAND Corporation. Retrieved May 9, 2008, from http://www.rand.org/pubs/research_briefs/2005/RAND_RB9050.pdf
- McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND Corporation. Retrieved May 9, 2008, from http://www.rand.org/pubs/monographs/2004/RAND_MG158.pdf

- National Association of State Boards of Education. (2005). *Evaluating value-added: Findings and recommendations from the NASBE Study Group on value-added assessments*. Alexandria, VA: Author.
- Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes* (Report). Washington, DC: Economic Policy Institute.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94(2), 247–252.
- Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247–256.
- Tennessee Department of Education. (2007). *Tennessee's most effective teachers: Are they assigned to the schools that need them most?* (Research Brief). Nashville, TN: Author. Retrieved May 9, 2008, from http://tennessee.gov/education/nclb/doc/TeacherEffectiveness2007_03.pdf
- Valli, L., Croninger, R., Alexander, P., Chambliss, M., Graeber, A., & Price, J. (2004, April 15). *A study of high quality teaching: Mathematics and reading*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73(1), 89–122.