



What's Happening

July 2015

Alternative student growth measures for teacher evaluation: Implementation experiences of early-adopting districts

Moira McCullough
Brittany English
Megan Hague Angus
Brian Gill
Mathematica Policy Research

Key findings

- Evaluation systems that include alternative measures of student achievement growth showed a wider range of teacher performance than previous evaluation systems that lack measures of student growth.
- Evidence is limited on the reliability and validity of alternative student growth measures used in early-adopting districts—especially student learning objectives (customized goals set by teachers).
- Among districts that used student learning objectives, the most frequently reported benefit was increased collaboration, whereas alternative assessment-based value-added models of student growth were perceived as fairer than student learning objectives.
- Alternative student growth measures came with financial costs and implementation challenges, especially related to teacher time, test administration, and rigor in evaluating performance.

U.S. Department of Education

Arne Duncan, *Secretary*

Institute of Education Sciences

Ruth Neild, *Deputy Director for Policy and Research*
Delegated Duties of the Director

National Center for Education Evaluation and Regional Assistance

Joy Lesnick, *Acting Commissioner*
Amy Johnson, *Action Editor*
Felicia Sanders, *Project Officer*

REL 2015–093

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

July 2015

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0006 by Regional Educational Laboratory Mid-Atlantic administered by ICF International. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

McCullough, M., English, B., Angus, M. H., & Gill, B. (2015). *Alternative student growth measures for teacher evaluation: Implementation experiences of early-adopting districts* (REL 2015–093). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

This report is available on the Regional Educational Laboratory website at <http://ies.ed.gov/ncee/edlabs>.

Summary

Throughout the country, school districts are scrambling to adhere to new state requirements for teacher evaluation. More than 40 states have mandated that some measure of student achievement growth be included in teacher evaluations. Annual state assessments are commonly used for this purpose (using statistical techniques known as value-added models or student growth models) but typically cover only grades 3–8 and one high school grade and have minimal coverage of subjects other than math and reading.

To address this limitation, many states and districts are developing alternative ways to measure student growth in grades and subjects not covered by state assessments. These alternative student growth measures fall into two broad categories:

- Alternative assessment-based value-added models (or student growth models), which use statistical methods to measure a teacher's contribution to students' achievement growth on end-of-course assessments or commercially available tests.
- Student learning objectives, which are customized goals set by a teacher (or a team of teachers) at the beginning of the school year with the approval of the principal. Student learning objectives do not involve statistical modeling.

These alternative student growth measures have only recently begun to be implemented widely. Thus, information is limited on how the measures can be used to evaluate teachers and on the costs and benefits associated with implementing the measures. This study examines implementation of alternative student growth measures in a sample of eight school districts that were early adopters of the measures. It builds on an earlier Regional Educational Laboratory Mid-Atlantic report that described the two types of alternative student growth measures—alternative assessment-based value-added models and student learning objectives—in the early-adopting districts (Gill, English, Furgeson, & McCullough, 2014). This report incorporates perspectives from multiple stakeholders in the districts and more closely examines how the measures were used, their costs and benefits, and common implementation challenges and solutions.

Key findings include:

- The early-adopting districts were using or preparing to use the alternative student growth measures alongside other measures of teacher effectiveness in formal evaluations.
- Across the districts, evaluation systems that include alternative student growth measures showed a wider range of teacher performance than previous evaluation systems that lack measures of student growth.
- Evidence is limited on the reliability and validity of alternative student growth measures used in the districts—especially student learning objectives.
- Alternative student growth measures—especially student learning objectives—were used for many purposes other than teacher evaluation, at the district and school levels.
- Among districts that used student learning objectives, the most frequently reported benefit was increased collaboration, whereas alternative assessment-based value-added models were perceived as fairer than student learning objectives.
- Alternative student growth measures came with financial costs and implementation challenges, especially related to teacher time, test administration, and rigor in evaluating performance.

Contents

Summary	i
Why this study?	1
What the study examined	2
What the study found	3
How districts incorporate alternative student growth measures in teacher evaluation	3
How alternative student growth measures differentiate teacher performance	4
How alternative student growth measures are used outside of teacher evaluation	8
Benefits of alternative student growth measures	13
Costs and challenges of alternative student growth measures and how they were addressed	15
Implications and limitations of the study	19
Appendix A. Study sample selection, data collection, and analysis	A-1
Appendix B. How the sample districts incorporate alternative student growth measures into teacher evaluation systems	B-1
Appendix C. Benefits of using alternative student growth measures for student populations with special needs	C-1
Appendix D. Contextual factors influencing implementation: funding, policies and politics, and teachers' unions	D-1
Appendix E. District administrator interview protocol	E-1
Appendix F. Principal interview protocol	F-1
Appendix G. Teacher interview protocol	G-1
Appendix H. Teachers' union representative interview protocol	H-1
References	Ref-1
Boxes	
1 Definitions of alternative student outcome measures and approaches to measuring teacher performance	2
2 Using student learning objectives for high-stakes purposes and instructional improvement	12
B1 Adding student learning objectives alongside alternative assessment-based value-added measures	B-5
D1 Culture of data use and implementation of student learning objectives	D-2
Tables	
1 Districts and stakeholder respondents	3
2 Reported correlations of alternative student growth measures with other performance measures	6

3	Informal uses for alternative student growth measures, by school-level user	10
4	Benefits of alternative student growth measures	13
A1	Stakeholder respondent samples, by district	A-2
B1	Measures used or proposed for use in teacher evaluation systems, as of 2013/14 school year	B-4

Why this study?

Throughout the country, school districts are scrambling to adhere to new state requirements for teacher evaluation. More than 40 states have mandated that some measure of student achievement growth be included in teacher evaluations. Annual state assessments are commonly used for this purpose (using statistical techniques known as value-added models or student growth models). These value-added models can be used for only a subset of teachers, because state assessment regimes typically include only grades 3–8 and one high school grade and have minimal coverage of subjects other than math and reading. As a result, student growth measures based on state assessments are not available for most teachers in most districts (Whitehurst, 2013).

To address this limitation, many states and districts are developing alternative ways to measure student growth in grades and subjects not covered by state assessments (see box 1 for definitions of these measures). These alternative student growth measures fall into two broad categories:

- Alternative assessment–based value-added models (or student growth models), which use statistical methods to measure a teacher’s contribution to students’ achievement growth on assessments other than those in math and reading in grades 3–8, such as end-of-course assessments or commercially available tests.
- Student learning objectives, which are customized goals set by a teacher (or a team of teachers) at the beginning of the school year with the approval of the principal and which can vary across classrooms based on school-level goals, the aims of the course, and students’ abilities. Student learning objectives might use standardized assessments to measure growth but do not involve statistical modeling.

These alternative student growth measures have only recently begun to be implemented widely. Thus, information is limited on how the measures can be used to evaluate teachers and on the costs and benefits associated with implementing the measures (see Gill, Bruch, & Booker, 2013, for a review of the literature on the reliability and validity of alternative student growth measures). Based on a request from Regional Educational Laboratory (REL) Mid-Atlantic’s Teacher Evaluation Research Alliance, this study examines implementation of alternative student growth measures in a sample of eight school districts that were early adopters of the measures. It provides important information on key implementation considerations not only to the REL Mid-Atlantic Region—where all five states have committed to including student growth measures in teacher evaluation—but also to states and districts across the country.

This report is the second of two reports from a study of alternative student growth measures. The first report used pilot interviews with one administrator in each early-adopting sample district to examine the selection, development, and use of the measures and to report the district’s perspective on implementation (Gill et al., 2014). Among the findings of the first report were:

- To promote consistency in student learning objectives, districts sometimes required that student learning objectives be approved by the district office.
- Districts that used student learning objectives chose them as a teacher-guided method of assessing student growth, whereas districts that used an alternative assessment–based value-added model were motivated to use them partly to take advantage of existing assessments.

This study examines implementation of alternative student growth measures in eight school districts that were early adopters of the measures

Box 1. Definitions of alternative student outcome measures and approaches to measuring teacher performance

End-of-course assessment. An assessment designed by a state or district to assess student mastery of the content covered in a course.

Reliability. How well differences among teachers' ratings on a measure at one point in time reflect true differences in teachers' skills that would also be observed at a different time (test-retest reliability).

Student growth models. Statistical methods (such as the student growth percentile method) that measure achievement growth among a teacher's students. Their purposes are similar to those of value-added models, and this report treats the two approaches as functionally equivalent.

Student learning objectives. Classroom-specific achievement growth targets set by individual teachers and approved by principals or other administrators.

Validity. The extent to which an alternative assessment-based value-added model or student learning objective measures the concept that it is intended to measure (teacher effectiveness at raising student achievement).

Value-added model. A statistical approach that estimates a teacher's impact on student achievement using students' prior achievement and other characteristics. The achievement measures in a value-added model can be any standardized assessment.

Widely used commercial test. A standardized, nationally normed assessment such as the Stanford Achievement Test or the Iowa Tests of Basic Skills.

- Student learning objectives required substantial teacher, principal, and district participation, but value-added models did not require as much work by teachers and principals to implement.
- Alternative student growth measures can be used for multiple purposes, including teacher evaluation, professional development, and performance-based compensation.

This report incorporates perspectives from multiple stakeholders in the districts and explores implementation in greater depth, examining how the measures were used, their costs and benefits, and common implementation challenges and solutions.

What the study examined

This report builds on a previous examination of how the two types of alternative student growth measures—alternative assessment-based value-added models and student learning objectives—were implemented to measure teacher performance in eight early-adopting school districts (Gill et al., 2014). This report addresses five research questions:

- How are alternative student growth measures used to evaluate teachers? In particular, how are alternative student growth measures incorporated alongside state assessment-based value-added models, measures of professional practice, and other elements of evaluation systems?

Table 1. Districts and stakeholder respondents

District	Measure implemented at time of sample selection	Years of implementation, as of 2013/14	District size ^a	Types of stakeholders interviewed
A	Student learning objectives	4	Large	District
B	Student learning objectives	More than 5	Large	District
C	Student learning objectives	5	Medium	District, principal, teacher
D	Student learning objectives	More than 5	Large	District, principal, teacher, teachers' union
E	Alternative assessment–based value-added model	4	Large	District, teachers' union
F	Alternative assessment–based value-added model	4	Small	District, principal, teacher, teachers' union
G	Alternative assessment–based value-added model	3	Medium	District, teachers' union
H	Alternative assessment–based value-added model	4	Medium	District, principal, teacher, teachers' union

a. *Small* indicates fewer than 10,000 students, *medium* indicates 10,000–50,000 students, and *large* indicates more than 50,000 students.

Note: Because many of the districts requested anonymity, this report does not identify them.

Source: Interviews with district administrators supplemented by a review of documents on district websites.

District priorities, as well as state policies dictating measures that must be included in teacher evaluation systems, influence the pace and scope of implementation of alternative student growth measures

- How do alternative student growth measures compare with ratings based on classroom observation and ratings based on state assessment–based value-added models in differentiating teachers?
- For what purposes other than teacher evaluation are alternative student growth measures used?
- What are the perceived benefits of using each type of alternative student growth measure?
- What are the perceived drawbacks of using each type of alternative student growth measure, and what costs do the measures impose on teachers, principals, and districts? What challenges arise when alternative student growth measures are implemented, and how have experienced districts addressed them?

This report is based on interviews with district administrators, principals, teachers, and teachers' union representatives in the eight districts, supplemented by document reviews (see appendix A for further description of the research methods).

The study examined the experiences of four districts that were using an alternative value-added model and four districts that were using student learning objectives. Information about each district and the stakeholders interviewed is provided in table 1.

What the study found

This section presents the study findings, grouped by research question.

How districts incorporate alternative student growth measures in teacher evaluation

District priorities, as well as state policies dictating measures that must be included in teacher evaluation systems, influence the pace and scope of implementation of alternative

student growth measures. While some districts prioritized incorporating alternative student growth measures to align with strategic priorities around identifying effective teachers, others were compelled to use alternative student growth measures sooner than planned because of state legislation.

Among the eight sample districts, three included one or more alternative student growth measures in their teacher evaluation system during 2013/14, and four were preparing to do so. The remaining district had previously used student learning objectives in teacher evaluation but had stopped by the 2013/14 school year. See appendix B for a detailed description of how the sample districts are integrating alternative student growth measures into teacher evaluation systems.

How alternative student growth measures differentiate teacher performance

This section examines evidence from the sample districts on the extent to which alternative student growth measures differentiate among teachers, the extent to which teachers receive consistent scores on the measures across years, and the relationships between teachers' scores on the alternative student growth measures and their scores on other performance measures.

Both types of alternative student growth measures showed a wider range of teacher performance than ratings based on classroom observation, although binary student learning objectives showed limited differentiation. In all four districts that used student learning objectives, a majority of teachers met their student learning objective goals. Across the four districts the proportion of teachers meeting at least one student learning objective goal in 2012/13 ranged from 68 percent to 87 percent; in one district 80 percent of teachers met two of two goals. All the respondents from districts that used student learning objectives noted that the proportion of teachers meeting student learning objective goals (a binary measure) was lower than the proportion rated as satisfactory under comparable binary (satisfactory or unsatisfactory) measures based on classroom observation that were used in previous years.

All the respondents from districts that used student learning objectives noted that the proportion of teachers meeting student learning objective goals was lower than the proportion rated as satisfactory under comparable binary measures that were used in previous years based on classroom observation

Three of the four districts that used student learning objectives measured performance on a binary scale of “met” or “not met,” which inherently limited differentiation among teachers (similar to previous systems that used only “satisfactory” and “unsatisfactory” performance categories). The remaining district assessed student learning objective performance on a 1–4 categorical scale. Two of the three districts that used a binary scale to measure student learning objective performance in 2013/14 were planning to implement a continuous or categorical scale. (The remaining district no longer used student learning objectives for any purpose.) In two of the three districts that were using or planning to use a continuous or categorical scale, teachers set rating thresholds at the beginning of the school year for their own student learning objectives; those thresholds define rating categories based on the percentage of students meeting growth targets. In the third district the proportion of a teacher's students meeting growth targets was converted to a student learning objective score on a continuous 1–4 scale (for example, if 80 percent of students met growth targets, the student learning objective score would be 80/25, or 3.2).

In the two districts where principals reported on distributions of performance on student learning objectives in their schools, the range of performance was wider when

a four-category scale was used than when a binary scale was. One of the two districts was piloting four-category student learning objective scores. Principals reported that an estimated 95–100 percent of teachers in their schools had met student learning objective goals under the previously used binary scale. Those principals were not yet able to report on distributions for the current year, but one principal anticipated a higher proportion of teachers not meeting student learning objective goals, due in part to the categorical system and in part to increased rigor in goal setting. A district administrator also indicated that the four-category scoring system was adopted to increase the district’s ability to differentiate teacher performance. In the other district, which was already assessing student learning objectives using a four-category scale (with 1 as the lowest performance category), one principal reported that most teachers received a 3 or a 4. Another principal estimated that 50 percent of teachers received a 4, 30 percent received a 3, 20 percent received a 2, and none received a 1. The extent of differentiation of performance under the current system was perceived to have increased relative to the extent of differentiation when the district had used a binary scoring system.

In each of the four districts that used an alternative assessment-based value-added model, teacher evaluation scores showed a wider range of teacher effectiveness than previous scores that used the same performance categories but did not incorporate growth measures

In each of the four districts that used an alternative assessment-based value-added model, teacher evaluation scores under the value-added model showed a wider range of teacher effectiveness than previous scores that used the same performance categories but did not incorporate growth measures. In each district the scores from the value-added model follow a normal, bell-shaped distribution.

In one district, when results from the end-of-course assessment-based value-added model were incorporated into teacher evaluation, an estimated 15 percent of teachers scored at the highest level on a four-point scale, 9 percent of teachers scored at the lowest level, and the remaining teachers scored in the middle two performance categories. On average across grades and subjects, scores for more than a third of teachers who were subject to the end-of-course assessment-based value-added model could be statistically distinguished from the district average.

Across each of the other three districts that used an alternative assessment-based value-added model, incorporating scores from the model corresponded with an increase in the proportion of teachers at the lower end of the distribution of performance ratings:

- In one district the proportion of teachers identified as “need[ing] improvement” (the second-to-lowest rating on a four-point scale) increased when results from the end-of-course assessment-based value-added model were incorporated into teacher evaluation. The proportion of teachers identified as “highly effective” (the highest rating) remained unchanged. Teachers who did not have results from the value-added model in their evaluation because student data were missing tended to receive the highest scores; in fact, many of these teachers received perfect scores.
- In another district, including scores from the value-added model (both end-of-course assessment-based value-added models and state assessment-based value-added models) in evaluation ratings (in a pilot process) resulted in nearly a third of teachers receiving a rating of “needs improvement” or “ineffective”—the two lowest categories on a five-point scale. The proportion of teachers in the two lowest categories had previously been substantially smaller and typically included new teachers who were not retained.
- One district was unable to provide data on previous and current distributions of teacher performance, but teachers and principals reported a larger proportion of

teachers at the lower end of the distribution after incorporating a commercial test-based value-added model.

Evidence on the reliability and validity of alternative student growth measures in the districts was limited, and respondents' perceptions of the consistency and accuracy of alternative student growth measures were mixed. Although seven of the eight sample districts reported ongoing efforts to examine the reliability or validity of alternative student growth measures, only three cited documented evidence on the statistical properties of the measures. These documents provide some support for the validity of student learning objectives (that is, evidence that ratings based on student learning objectives are consistent with other measures of teacher effectiveness) and some evidence of the reliability and validity of end-of-course assessment-based value-added models. Among respondents in districts that used an alternative student growth measure, perceptions were mixed on whether the measure consistently and accurately captured teacher contributions to student learning.

Among respondents in districts that used an alternative student growth measure, perceptions were mixed on whether the measure consistently and accurately captured teacher contributions to student learning

Three of the four districts that used student learning objectives and two of the four districts that used an alternative assessment-based value-added model (both of which used an end-of-course assessment) cited positive correlations between the alternative student growth measure and at least one other measure of teacher performance (table 2). Both districts that used an end-of-course assessment-based value-added model reported a statistically significant positive correlation between the alternative student growth measure and observational measures of teachers' practice; however, in one of the districts the correlations were small (between 0.10 and 0.20). The districts that used student learning objectives reported similarly positive relationships between student learning objective performance and other student growth and achievement measures.

Table 2. Reported correlations of alternative student growth measures with other performance measures

District	Type of measure	Correlation with state assessment-based value-added model	Correlation with observational measures	Correlation with other measures (such as student achievement and surveys)
A	Student learning objective	Positive		
B	Student learning objective	Positive		
C	Student learning objective			
D	Student learning objective			Positive*
E	Alternative assessment-based value-added model		Positive*	
F	Alternative assessment-based value-added model			
G	Alternative assessment-based value-added model			
H	Alternative assessment-based value-added model	Positive*	Positive*	Positive*

*Statistically significant, as reported by district.

Source: District site visits and interviews.

In three of the four districts where principal interviews were completed—one of which used an alternative assessment–based value-added model and two of which used student learning objectives—principals perceived teacher performance on alternative student growth measures to be inconsistent with performance on other measures:

- In one district both principal respondents noted that the scores teachers received based on classroom observation often differed from the scores they received based on student learning objectives. However, district administrators noted that the district aimed for low correlations across the various measures used to assess teacher performance because it believes that each component captures a unique component of teacher performance.
- In another district a principal reported that the scores teachers received based on classroom observation were often higher than the scores they received based on the alternative assessment–based value-added model.
- In the third district principals' observations were mixed: two reported that student achievement was positively related to teacher attainment of student learning objective goals, but a third noted that even though most teachers met their student learning objective goals, student achievement was low and growth was stagnant.

Most principals, teachers, and teachers' union respondents in the four districts that used an alternative assessment–based value-added model were skeptical that the model was a valid measure of teacher performance

Evidence of the reliability of alternative student growth measures was limited primarily to the districts that used an alternative assessment–based value-added model. Of those four districts:

- Two reported that scores for individual teachers were consistent across years.
- One reported that scores for individual teachers were somewhat consistent across years.
- One reported that scores based on a value-added model varied across years (which was one reason the district used multiple years of test scores in its value-added model).

Only one of the four districts that used student learning objectives had examined the consistency of student learning objective performance across years; that district found that student learning objective attainment tended to improve for novice teachers and to remain stable for experienced teachers.

Overall, principals were more likely than teachers and teachers' union representatives to perceive alternative student growth measures as accurately capturing teacher contributions to student learning. Teachers and teachers' union representatives were more likely to find the measures inaccurate and to believe that the measures should not be used to evaluate teachers. In the two districts that used student learning objectives where principals were interviewed, teachers and principals alike believed that student learning objective performance depended on teachers' effort to tie student learning objective goals to specific standards and to use a measure appropriate for assessing those standards. Principals and teachers perceived student learning objectives to be inaccurate measures of teacher contributions to growth in student learning when treated as a compliance exercise.

Most principals, teachers, and teachers' union respondents in the four districts that used an alternative assessment–based value-added model were skeptical that the model was a valid measure of teacher performance. In two districts teachers and teachers' union representatives reported personal experiences with scores from the value-added model varying widely from one year to the next. Principals' and teachers' concerns about the accuracy of

the alternative assessment–based value-added model sometimes centered on the underlying student outcome measures. In two districts respondents were particularly concerned with a value-added model that was based on high school end-of-course assessments.

How alternative student growth measures are used outside of teacher evaluation

The use of alternative student growth measures by the sample districts evolved in response to stakeholder input, district priorities, and external forces such as changing state policies. This section explores formal and informal uses identified by stakeholders in all eight districts. In the two districts without principal, teacher, or teachers’ union respondents, only district-level uses are reported.

All districts had formally or informally used performance on alternative student growth measures to target professional development. The measures were used both formally at the district level and informally at the school level to identify teachers for professional development. Performance on the measures informed professional development in three ways:

- *Professional development formally targeted based on teachers’ overall performance.* In three districts teachers’ overall evaluation ratings, which incorporated the alternative student growth measures, determined when teachers were placed on improvement plans, which included targeted professional development. One of these districts used schoolwide performance—including student learning objective scores—to determine the allocation of instructional coaches to schools.
- *Professional development formally targeted using teachers’ performance on or goals for an alternative student growth measure.* Two districts systematically identified or had identified teachers for professional development using only their previous or targeted performance on an alternative student growth measure. One district targeted professional development based on scores from a value-added model. In this district professional development could be targeted to individual teachers, grade or subject teams or areas, or entire schools because of individual-, group-, or school-level performance under the value-added model. Another district used student learning objectives to reinforce professional development. For instance, principals used the student learning objective process to identify school-, grade-, or subject-level goals and then used professional development to help teachers achieve the goals.
- *Ongoing professional development informally provided based on teachers’ progress toward goals under the alternative student growth measure.* In two districts decisions on targeted professional development based on student learning objective performance were made informally at the school level. Principals in both districts used teachers’ progress toward goals for the alternative student growth measure to provide teachers with additional, ongoing support as needed.

Alternative student growth measures were used both formally at the district level and informally at the school level to identify teachers for professional development

In seven of the eight districts, performance-based compensation was or had been informed by performance on the alternative student growth measure. Seven districts provided teachers with bonuses or salary increases related to their performance on the alternative student growth measure. In three districts that used student learning objectives and one district that used an alternative assessment–based value-added model, performance-based compensation was initially the primary use of the measure. Administrators in these districts noted that the initial focus on rewards was intended to promote teacher familiarity with and support for the measures. In two districts alternative student growth measures had always been used for performance-based compensation and teacher evaluation simultaneously.

The structure of performance-based compensation varied by the type of measure implemented, as well as by its integration with other measures of teacher performance. Districts provided performance-based compensation to teachers in two ways:

- *End-of-year bonuses tied directly to performance on the measure.* In three of the four districts that used student learning objectives, performance-based compensation was provided through end-of-year bonuses tied directly to student learning objective performance at the individual, team, or school level. These bonuses were typically part of compensation systems that offered incentives for such factors as retention in high-need schools and subjects in addition to performance-based bonuses. In these districts use of the measure for compensation depended on the availability of funding, such as grants from the Teacher Incentive Fund or Race to the Top.
- *End-of-year bonuses or salary increases determined by overall performance on multiple measures of teacher effectiveness.* Three districts that used an alternative assessment-based value-added model and one district that used student learning objectives delivered performance-based compensation to teachers based on multiple measures of teacher effectiveness, including the alternative student growth measure. In three of these districts, compensation was determined based on a teacher's overall evaluation rating. The form of delivery of additional compensation varied. In two districts teachers' performance could result in base pay increases, and in the other district, teachers received end-of-year bonuses linked to their performance. In the fourth district experienced teachers could opt into additional compensation based on performance on multiple measures, and new teachers could receive salary increases based on a salary schedule that accounted for overall performance.

In four districts multiple measures of teacher performance, including alternative student growth measures, were increasingly affecting districts' retention decisions; in the remaining districts alternative student growth measures were not used in these decisions

In four districts performance on alternative student growth measures informed retention decisions. At least four districts had recently moved away from or were in the process of abandoning traditional tenure policies, following changing state policies. In districts with tenure, traditional factors such as teachers' years of experience and satisfactory performance drove tenure decisions; the alternative student growth measures were not yet included in teacher contracts. However, in four districts multiple measures of teacher performance, including alternative student growth measures, were increasingly affecting districts' retention decisions. In two districts that did not have tenure, a teacher's overall evaluation rating determined whether the district retained the teacher. Low performers, identified through multiple measures, might not be retained, and high performers might receive salary increases. Two other districts considered performance on alternative student growth measures when deciding whether to move probationary teachers to nonprobationary contracts. In the remaining four districts alternative student growth measures were not used in retention decisions.

In two districts that used student learning objectives, alternative student growth measures served as a means of ensuring curriculum and standards alignment. Two districts that used student learning objectives reported using the measures to assess alignment of the curriculum or standards across the district. The district offices were highly involved in the student learning objective process. For example, district policy might specify the assessments that teachers could use for their student learning objectives, which could ensure that instruction is geared toward the curriculum or standards. District administrators also indicated that teachers were able to deepen their understanding of standards by developing

their student learning objective goals and crafting assessments aligned with standards. In these two districts the student learning objective process (rather than student learning objective results) was used to promote districtwide instructional or achievement goals.

Principal-reported uses included fostering collaboration, targeting professional development, encouraging data-driven instruction, and building assessment capacity. A variety of informal uses for alternative student growth measures were identified across the six districts where principals, teachers, or teachers’ union representatives were interviewed (table 3). Fostering collaboration among teachers and establishing a culture of data use were key uses reported by principals from districts that used student learning objectives. In two of these districts, principals indicated that their districts emphasized data-driven instruction. Principals engaged in various activities to use the measures in this way, such as convening professional learning communities in which teachers examine performance on student learning objective measures to individualize instruction for their students. Within these communities, teachers established common goals, tracked student progress, and graded student work at the team level rather than the classroom level. This approach allowed principals to use the student learning objective process to set school-level goals.

Principals also used alternative student growth measures to determine teaching assignments

Principals also used alternative student growth measures, on their own or as part of an overall evaluation system, to determine teaching assignments. This was reported by district administrators, principals, and teachers’ union representatives in three districts that used an alternative assessment-based value-added model and three that used student learning objectives. For example, elementary school principals might consider departmentalizing instruction in a grade if one teacher excels at teaching math and another teacher in the grade excels at teaching English language arts. Performance on the measures also helped principals assign content areas in upper elementary grades. In two districts, district administrators and teachers’ union representatives suggested that results on the measure or overall evaluation ratings may eventually affect teaching assignments across schools as well as within schools. Indeed, one teachers’ union respondent regarded this use of alternative assessment-based value-added models as having the most potential to make a positive impact on student outcomes because value-added model subgroup analyses could indicate whether certain teachers are more effective with certain student populations and thus could be used to place teachers in schools accordingly. However, transfer policies negotiated as part of teachers’ contracts typically constrain the use of performance to adjust assignments across schools within a district.

Table 3. Informal uses for alternative student growth measures, by school-level user

Reported use	Alternative assessment-based value-added measures	Student learning objectives
Encouraging data-driven instruction		Principals, teachers
Targeting professional development	Principals	Principals
Determining teaching assignments	Principals	Principals
Fostering collaboration		Principals, teachers
Building assessment capacity		Principals, teachers

Source: District site visits and phone interviews with stakeholders.

Teachers and teachers' union officials in districts that used student learning objectives reported that the measures informed instructional practice, while teachers and teachers' union officials in districts that used value-added models reported that the measures had minimal influence on instructional practice. In five districts, including three districts implementing student learning objectives, either teachers or teachers' union representatives reported that teacher uses of alternative student growth measures (outside of those formally specified by district policy such as teacher evaluation) focused on increased attentiveness to instructional practice. However, the nature of the alternative student growth measures and the type and frequency of feedback received by teachers dictated how the teachers used the measures to change their instructional practice. A teachers' union representative in one district and teachers in two other districts that used student learning objectives reported using student learning objectives to inform instruction. School-level priorities and policies largely guided teachers in using the data. Additionally, teachers who were required to complete student learning objectives were directly involved in implementation and had continual access to data on student progress throughout the year because they are responsible for monitoring the data. The ongoing receipt of student-level data was embedded in the student learning objective process, which teachers indicated facilitated data use. Teachers in districts that used an alternative assessment-based value-added model did not report using results outside of reflecting on their performance over the last year. In two districts teachers reported that the value-added model increased attention to student growth but that the results were not intended to and did not provide the information necessary to inform instructional practice.

Teachers and principals reported that the student learning objective process produces a great deal of data on student achievement that teachers can use

As teachers and principals reported, the student learning objective process produces a great deal of data on student achievement that teachers can use. Teachers from one district that used student learning objectives described in detail how the student learning objective process affected their approaches to instructional change and goal setting for themselves, as well as for their students. Through student learning objectives, teachers obtained detailed information on student achievement. Rather than seeing only a final test score, teachers who administered district- or teacher-developed assessments to measure student progress toward targets for student learning objective growth could review student responses on individual assessment items. Seeing this level of detail allowed teachers to draw more precise conclusions about student achievement and, in turn, establish goals for individuals or groups of students. If many students struggled with the same assessment items, teachers were also able to identify specific areas of weakness in their own instructional practice. Student learning objectives equipped these teachers with the information necessary to differentiate instruction and to improve their instructional strategies, with the goal of increasing student growth and achievement, but teachers may need assistance in using the data.

Teachers in two other districts noted that the student learning objective process guided how they established goals for themselves and their students. For example, teachers from one district focused goal setting on student learning objective outcomes that were measured and monitored throughout the year rather than on summative, state assessments given at the end of the year. In doing so, they adjusted their lesson plans throughout the year to ensure that students continued to make progress on student learning objective outcomes. Teachers from both districts emphasized that student learning objectives caused them to be more deliberate about their goal-setting processes and more reflective on their teaching practice. Student learning objectives could also shape their planning for the following school year.

Although student learning objectives can generate tools to improve instructional practice, this use could be undermined by the additional use of student learning objectives in teacher evaluations (box 2). The attachment of high stakes can give teachers an incentive to “game the system.” District administrators struggled to establish processes that would maintain the use of student learning objectives to improve instruction while keeping the student learning objectives rigorous enough for use in evaluation.

Box 2. Using student learning objectives for high-stakes purposes and instructional improvement

Districts should be aware of the tensions that can arise when student learning objectives are used to meet multiple objectives. Using student learning objectives for evaluation or compensation can undermine their use in improving instructional practice.

Student learning objectives typically require teachers to establish goals, approved by their school leaders, at the beginning of the school year. These goals are then measured against student achievement on approved assessments, which can be purchased by the district or developed by individual teachers. Teachers are responsible for administering and scoring the assessments used in their student learning objectives. When high stakes, such as teachers’ evaluation ratings or additional compensation, are attached to student learning objective performance, teachers may be tempted to set low goals for themselves or to select less rigorous student assessments. In doing so, teachers would not receive the information necessary to make meaningful improvements to instruction. Respondents from all districts that used student learning objectives described ensuring a sufficient level of rigor as a key challenge of implementing student learning objectives (as detailed later in this report).

To address this tension, the sample districts made several changes to the student learning objective process in the hope that student learning objectives could be used for high-stakes purposes while helping improve instructional practice. One district was already using student learning objectives in its teacher evaluation system. Administrators from this district noted that initially teachers created their own assessments for use in student learning objectives, and the rigor of student learning objectives was inconsistent across teachers and schools. At the time, district administrators were focused on establishing the student learning objective process rather than centralizing assessments used in student learning objectives. The district later increased its focus on the assessments by purchasing more vendor assessments and specifying which assessments should be used for different grades and subjects. By doing so, district administrators hoped to improve the rigor of the measure in order to improve its use in teacher evaluation ratings, while providing teachers with accurate and consistent information about student achievement that could ultimately improve instruction.

One district that was using student learning objectives for performance-based compensation adopted a similar approach and specified at the district level which assessments teachers could use for their student learning objectives based on grade and subject taught. Another district that used student learning objectives—which had previously used student learning objectives only for instructional planning—was keenly aware of this tension. District administrators were thus trying to establish a process that would help maintain the use of student learning objectives to improve instruction while being rigorous enough for use in its teacher evaluation system. The districts aimed to strike a balance between empowering teachers in the district’s student learning objective process and centralizing the student learning objective process to ensure consistency and rigor for high-stakes use. Such procedures may help promote the rigor and integrity of the assessments and goals, but they cannot eliminate the inherent conflicts of interest created by using the measures for high-stakes purposes.

In contrast to student learning objectives, alternative outcome-based value-added models provide teachers with only end-of-the-year results. Respondents from districts that used an alternative assessment-based value-added model indicated that results minimally influenced instructional practice. Because teachers receiving value-added scores do not receive ongoing information regarding their current students' performance, they cannot make real-time adjustments to their instruction.

Benefits of alternative student growth measures

This section examines respondents' perceptions of the benefits of alternative student growth measures. Reporting reflects responses from all eight sample districts unless otherwise noted. The benefits reported by respondents in the eight districts are itemized in table 4. (Appendix C provides additional information on benefits for student populations with special needs).

The most often reported benefit among respondents from districts that used student learning objectives was increased collaboration; alternative assessment-based value-added models were perceived as fairer than student learning objectives both in districts that used an alternative assessment-based value-added model and in districts that used student learning objectives. Respondents from all four districts that used student learning objectives reported that the measures increased collaboration. District administrators (in all four districts) and principals and teachers (in two of them) reported that the student learning objective process—including analyzing baseline data, setting targets, and selecting or designing assessments—increased collaboration, particularly among teachers. District administrators and principals noted that student learning objectives helped build community and accountability at the school level, galvanizing school staff around similar goals. In one district the student learning objective process requires teachers to work together in grade or subject teams to develop standards, rubrics or assessments for assessing standards, and targets for student learning objectives. To ensure consistency, teachers shared responsibility for scoring the assessments administered to each other's

Respondents from districts that used an alternative assessment-based value-added model indicated that results minimally influenced instructional practice

Table 4. Benefits of alternative student growth measures

District	Type of measure	Increased collaboration	Increased use of data	Increased focus on instructional practice	Teacher empowerment
A	Student learning objective	✓	✓	✓	✓
B	Student learning objective	✓			
C	Student learning objective	✓	✓	✓	✓
D	Student learning objective	✓	✓	✓	✓
E	Alternative assessment-based value-added model		✓	✓	✓
F	Alternative assessment-based value-added model		✓		
G	Alternative assessment-based value-added model		✓	✓	✓
H	Alternative assessment-based value-added model		✓		
Total		4	7	5	5

Source: District site visits and interviews.

students. Principals and teachers noted that the collaborative process increased teachers' opportunities to share with colleagues and fostered discussions of teaching practice. One district respondent pointed out that student learning objectives also fostered collaboration between teachers and principals by providing an avenue for teachers to ask their principals for help when needed.

In all four districts that used an alternative assessment-based value-added model, district administrators noted that a key benefit was the perceived increase in fairness resulting from a wider scope of teachers being evaluated using a value-added model. In two of these districts, which had also recently begun implementing student learning objectives, principals, teachers, and teachers' union respondents viewed alternative assessment-based value-added models as more rigorous and fairer than student learning objectives. In another district a teachers' union respondent noted that the divide between teachers evaluated using value-added models—regardless of whether they were based on state assessments or alternative assessments—and teachers evaluated using only non-value-added models was more likely to induce teacher resentment and backlash about other teachers being held to less rigorous evaluation standards.

Improved instruction, increased use of and conversation about data, or both were primary benefits of implementation, regardless of type of measure. Respondents in seven districts reported an increase in both their use of and conversations about data. District-generated reports or external vendor-generated reports often compared teacher performance on the alternative student growth measures at the classroom and school levels. For district administrators and principals, examining these data resulted in increased conversation about student achievement and school expectations for student achievement. Data from the student assessments administered to measure progress on student learning objectives were particularly useful because teachers could see student responses to specific test items, which are not provided in state assessment results.

In all four districts that used an alternative assessment-based value-added model, district administrators noted that a key benefit was the perceived increase in fairness resulting from a wider scope of teachers being evaluated using a value-added model

Across three districts that used student learning objectives, respondents of each stakeholder type in two districts and a district administrator in the third district noted that student learning objectives encourage teachers to target their instruction around goals, particularly when held accountable by evaluation scores. The student learning objective process also helped principals better understand how to help individual teachers. Using and monitoring data during the academic year enabled teachers to respond to academic needs in their classrooms and to tailor their instruction accordingly.

District administrators, principals, and teachers' union respondents noted that analyzing alternative student growth measures enabled the identification of specific weaknesses in schools and classrooms. A principal might use data to find ways to support teachers or to identify goals that could feed into a school improvement plan that addressed those weaknesses. However, stakeholders in districts that used an alternative assessment-based value-added model could use the results only to identify weak schools and teachers and, to some extent, how school and teacher effectiveness varied across student subgroups. The value-added model results alone do not provide any information about why or how certain schools or teachers are less effective.

Alternative student growth measures could create a sense of validation and empowerment among teachers. The ability of alternative student growth measures, especially

student learning objectives, to empower teachers was reported as a benefit by five districts. Respondents—including a teachers' union representative—from a district that used an alternative assessment-based value-added model reported that the results can be rewarding for teachers. High levels of progress validate teachers' efforts; these efforts can then be celebrated. In three districts respondents reported that student learning objectives gave teachers a stake in how they were evaluated, which in turn could foster a strong sense of buy-in among the teachers. One teacher appreciated the opportunity to weigh in on the most appropriate assessments for measuring student growth based on his experience in the classroom. The implicit confidence in teacher-developed assessments via the student learning objective process enhanced teacher engagement in it.

Costs and challenges of alternative student growth measures and how they were addressed

This section examines the costs, drawbacks, and challenges faced by the sample districts during the implementation of alternative student growth measures. Additional factors affecting the implementation of alternative student growth measures in the sample districts are presented in appendix D.

Implementing alternative student growth measures incurred substantial financial costs.

All districts incurred financial costs as a result of implementing the measures, though some of the costs were subsidized by outside funders. Costs included hiring and training staff as well as purchasing the assessments and infrastructure needed to implement the alternative student growth measures. Respondents found it difficult to place a dollar figure on the costs. One large district that served more than 50,000 students in more than 100 schools reported that the system and staff time for implementing student learning objectives cost \$400,000–\$500,000 a year (or approximately \$1,000–\$1,500 per teacher). The respondent also noted that the amount does not include the stipends teachers receive as a result of the district's compensation system.

A second district that used an alternative assessment-based value-added model and that served fewer than 8,000 students in fewer than 10 schools reported that the commercial end-of-course high school assessments administered by the district cost approximately \$18 per student. The respondent also estimated that just one of the three alternative assessments administered by the district costs more than \$10,000 per year to score and report the results. The district is able to afford these assessments because of support from its Teacher Incentive Fund grant, but the respondent noted that when the grant ends, the district would need to use its general fund to pay for the administration and scoring, which might not be sustainable.

The most frequently cited challenges in districts that used student learning objectives related to ensuring a high level of rigor across teachers and securing sufficient teacher time.

Three of the four districts that used student learning objectives noted that ensuring that the measure was consistently implemented with a high level of rigor was a substantial implementation challenge. All types of stakeholders expressed concern about the potential for some teachers to “game the system” by setting easily attainable goals. In two districts, district administrators, principals, and teachers noted that teachers lacked the expertise to develop or select rigorous assessments to measure student growth. For district administrators and principals, inconsistency in the rigor of goals set by teachers was a concern because it would render student learning objectives invalid as a means of assessing teacher

Three of the four districts that used student learning objectives noted that ensuring that the measure was consistently implemented with a high level of rigor was a substantial implementation challenge

contributions to student growth. For teachers and teachers' union representatives, perceptions that some teachers were deliberately setting low growth targets reinforced the perception that student learning objectives were not a fair measure of performance and should be viewed by teacher evaluators as only a compliance exercise.

District administrators, principals, teachers, and teachers' union representatives in the same three districts also cited finding time for teachers to develop and monitor student learning objectives as a major obstacle. Although student learning objectives required substantial time from district administrators and principals for implementation as well, all the responding principals considered the teacher time requirement to be a more substantial challenge than time requirements for other staff. In two districts none of the principal respondents perceived the time required of them to approve goals and assess progress to be an unusual burden outside of other responsibilities.

Although the majority of teacher respondents could not put an hour count on the student learning objective process, noting that it took "a lot" of hours, three of the eight teacher respondents across two districts estimated that it took three to eight hours to develop each student learning objective. Assessing the students took approximately five hours per student learning objective. Principals tasked with meeting with each staff member to review midyear student learning objective performance, as was the case in one district, noted that each meeting took at least an hour.

Although the process differed in districts that were using an alternative assessment-based value-added model, which does not require substantial teacher and principal time to implement, district administrators in districts that used both types of alternative student growth measures reported that adopting the measures required additional district staff time. In each district using alternative assessment-based value-added models, external vendors were responsible for calculating value-added model estimates. However, to support the implementation of alternative student growth measures, districts had to establish infrastructure, develop quality control practices, evaluate the quality of assessment, ensure rigor, and support staff in using the measures.

In three districts, district administrators also noted a lack of data capacity as a hindrance to implementation. Two of the three districts lacked an electronic system for teachers to submit goals and upload documentation of attainment. They instead relied on a tedious paper-based submission process. In the third district the lack of data infrastructure for teachers to continuously access student data and monitor progress toward goals made it difficult to encourage teachers to use student learning objectives as a tool for ongoing instructional improvement rather than as a compliance exercise.

To overcome these obstacles, districts piloted or implemented a variety of solutions. To ensure a consistently high level of rigor, one district incorporated an assessment of the degree of difficulty of student learning objective goals, guided by a standard rubric, into the principal approval process. Two districts focused on increasing teachers' assessment capacity. One district implemented a centralized solution, expanding the pool of assessments available to teachers. The other district used coaches and a collaborative process involving subject-area and grade-level teams to help teachers gain assessment expertise by learning from their peers. This process also facilitated the development of vertically aligned standards and tools to assess those standards. Sharing the scoring of outcome measures among

Although using an alternative assessment-based value-added model does not require substantial teacher and principal time to implement, district administrators in districts that used the measures reported that adopting them required additional district staff time

team members similarly aimed to promote consistency and peer accountability. None of the districts had any empirical evidence that these efforts had succeeded in making the student learning objectives consistent and rigorous.

Districts found it more difficult to find time for teachers to implement student learning objectives. Two districts began to include more team-based student learning objective goals to relieve the burden on teachers. Principals in one of those districts were also trying to integrate student learning objective–related assessment development, goal-setting, and monitoring into existing teacher team meetings that had already been set aside for data analysis. The district also provided substitute teachers to allow teachers to meet with student learning objective coaches during the school day.

The challenges that most often impeded implementation of alternative assessment–based value-added models were related to quality and administration. The most frequently cited implementation challenge for alternative assessment–based value-added models was issues related to the assessments themselves. Respondents from all four districts that used an alternative assessment–based value-added model consistently cited assessment quality and alignment with the curriculum, assessment administration, or both as ongoing struggles. Across three of the four districts, all respondents described the process of administering the alternative assessments as especially problematic.

Respondents from all four districts that used an alternative assessment–based value-added model consistently cited assessment quality and alignment with the curriculum, assessment administration, or both as ongoing struggles

Respondents voiced concern that the additional assessments imposed costs in terms of lost instruction time and decreased morale. In one district, implementing new assessments specifically for use in a value-added model was considered disruptive to classroom instruction. One teacher and teachers' union representative estimated that administering the end-of-course assessments (including makeup assessments) used in the district's value-added model disrupted a teacher's regular classroom schedule for a month, even though teachers themselves did not administer the assessments. Limited facilities resulted in a lengthy process of alternating through test administrations, and students had to be provided sufficient opportunities to take the assessments, particularly those required at the high school level. Teacher respondents cited a sense of assessment fatigue. District administrators in two of the three districts that noted assessment administration challenges also explained that finding sufficient space and computers for students to complete the homegrown end-of-course assessments was an annual problem.

In three of the four districts that used an alternative assessment–based value-added model, teachers perceived some of the outcome measures to be misaligned with the curriculum. Two of the districts used homegrown end-of-course assessments; the third used commercial assessments. Principals, teachers, and teachers' union respondents in one of the districts also noted that some homegrown assessments were plagued by obvious mistakes, which rendered them unusable for including in a value-added model.

Districts struggled to find solutions to address the assessment-related challenges of implementing an alternative assessment–based value-added model. In one district the homegrown assessments were continually reviewed and revised with input from teachers to improve validity and alignment with the curriculum. This process required considerable resources, however. To increase the usability of value-added model results, two districts continually revised the written feedback provided to teachers on their results, based on teacher input. Nonetheless, respondents in both districts expressed frustration that the

delay in receiving value-added model results, despite every effort at efficient assessment scoring, decreases their usefulness.

Complexity of the measure and test security presented additional implementation challenges for alternative assessment–based value-added models. In three districts that used an alternative assessment–based value-added model, various respondents cited the complexity of the measure as a key implementation challenge. Respondents believed that the lack of transparency of the measure increased teacher resistance. (This issue is a challenge related to implementing any value-added model, whether based on state assessments or alternative assessments.) Teachers’ union representatives and district administrators across two of the three districts also struggled to present value-added model results as actionable data for teachers to use to improve instruction. Teacher respondents typically received scores in the fall following spring test administration and were uncertain how to use the previous school year’s data during the current school year.

Other obstacles cited by district administrators included a lack of data capacity, especially for tracking and verifying links between teachers and their students. Administrators in districts where teachers administered the alternative assessments also noted the challenge of ensuring test security. When high stakes such as teacher evaluation ratings are involved, teachers can face competing motivations. Respondents indicated that the high-stakes nature of the measure could motivate teachers to seek out training and resources to improve their instruction. Nonetheless, as with other high-stakes tests, the stakes could also induce less desirable means of raising test scores. Moreover, procedures for test security are typically less rigorous for alternative assessments than for state standardized assessments, potentially increasing the opportunity and temptation to inflate scores artificially.

Districts developed a number of strategies to limit teachers’ temptation or ability to manipulate results. One district had assistant principals submit assessment results rather than teachers, and it can develop assessments “on-the-fly” when case test security concerns arise. Another district tracked the number of assessments provided to schools and then scored the assessments at the central office. Procedures such as these may promote the rigor and integrity of the assessments and goals, but they cannot eliminate the inherent conflicts of interest created by using the measures for high-stakes purposes.

Districts that used an alternative student growth measure struggled to communicate effectively with stakeholders. Regardless of the type of alternative student growth measure implemented, districts shared a common ongoing challenge: communicating effectively with stakeholders about the measure and its implementation. In seven of the eight districts, communication issues were cited by district administrators; principals, teachers, and teachers’ union representatives; or both groups. Respondents in districts that used student learning objectives complained specifically about a lack of communication from the district. Most teachers and principals described communication from the district as infrequent and limited to emails or materials posted on websites. In one district the perceived lack of communication was considered especially problematic because of impending changes to the measure expected to be implemented in the 2014/15 school year. Principals and teachers who were not participating in a pilot of the changes were perceived to be largely unaware of the new guidelines. Principals, teachers, and teachers’ union respondents also expressed frustration with the district’s unilateral approach to the incorporation of student learning

Other obstacles cited by district administrators in districts that used alternative assessment–based value-added models included a lack of data capacity and the challenge of ensuring test security

objectives into the teacher evaluation system and to subsequent changes in how student learning objectives were used in the system.

Districts that used alternative assessment-based measures were particularly challenged in their outreach to stakeholders as a result of a reported lack of teacher engagement with, and sometimes even resistance to, the measure. Across all four districts, district administrators and teachers' union representatives described an ongoing skepticism and general resistance to using value-added models of any kind to evaluate teachers. In two of the districts, teachers' union representatives reported that many teachers did not pay attention to their alternative assessment-based value-added model scores. In these districts the teachers' union played a substantial role in raising teacher awareness of the growth measure. One district worked closely with the teachers' union and relied on teachers' union representatives to remind teachers to verify their student rosters in advance of value-added model estimation and to check their individual value-added model results. In the other district teachers' union representatives prompted teachers to be aware of their value-added model results (which were not yet tied to teacher evaluation) as evidence of performance should any human resource issues arise. For example, a teacher at risk of nonretention based on classroom observation scores might make a case for retention (with professional growth support) by citing above-average value-added model results, among other evidence.

More research is needed on the statistical properties of alternative student growth measures and on innovative solutions to implementation costs and challenges

In two districts, one that used student learning objectives and one that used an alternative assessment-based value-added model, district administrators or teachers' union representatives reported a high level of reliance on principals to provide instructions and communicate with teachers about the alternative student growth measures. This approach was perceived as a problem. A teachers' union representative in the district that used an alternative assessment-based value-added model noted that the quality of communication and resulting teacher understanding and use of the growth measures varied widely by school. The respondent noted that teacher acceptance of the measure was directly related to how well the principal framed it as a tool to help improve instruction.

Implications and limitations of the study

This study provides information for districts to consider in deciding whether to use one of two types of alternative student growth measures to evaluate teachers. However, more research is needed on the statistical properties of the measures and on innovative solutions to implementation costs and challenges. Moreover, the study is limited by the small number of districts and respondents in the sample. The lack of a large, representative sample precludes nationwide generalizations. Nonetheless, the implementation experiences of the sample districts may be applicable for individual states and districts.

These early-adopting districts experienced several benefits from alternative student growth measures. With both types of alternative growth measure, districts reported increased data use among educators. Districts using student learning objectives also saw increased collaboration and teacher empowerment. Districts using alternative assessment-based value-added models had the benefit of an objective measure of student achievement growth that could be fairly compared across teachers.

Nonetheless, several findings emerged regarding limitations of the alternative student growth measures. First, value-added models are not useful for informing instruction

without additional information on teaching practice. Value-added models are “black box” measures: they provide information on how teachers compare with each other in promoting student achievement growth. But on their own, the measures do not explain why one teacher is more effective than another and therefore provide no information about how teachers might improve. They are also not available in time to be used to inform decisions about the current school year. Value-added models may be useful for formal evaluation and human resource decisions that ultimately improve instructional quality, but they do not provide data that can immediately inform instructional improvement.

Second, using student learning objectives for teacher evaluation creates an inherent conflict of interest. Using any measure of student achievement growth for high stakes may encourage “teaching to the test,” and student learning objectives exacerbate this problem by asking teachers to set their own targets. In essence, student learning objectives ask teachers to grade themselves. School districts may adopt procedures to counter the incentive to lower standards, but as long as student learning objectives are developed by teachers and used for evaluation, the conflict of interest remains. Principals and teachers in the case study districts worried that colleagues might deliberately set low growth targets to ensure a good evaluation result; this worry contributed to the perception that student learning objectives were not fair measures of performance. In the absence of high stakes, student learning objectives can be useful for instructional planning and improvement. Applying high stakes in the form of formal evaluations may undermine the value of student learning objectives for instructional planning.

Third, student learning objectives are extremely labor intensive. A comprehensive student learning objective process requires each teacher to gather substantial baseline data on every student and requires principals to carefully review the student learning objectives of all their teachers. Also, principals and teachers alike must be able to predict with reasonable reliability and validity the likely end-of-year achievement of their students. District and state officials interested in using student learning objectives for evaluative purposes might consider whether frontline educators have the time and the skills to create student learning objectives that permit reliable, valid, and fair comparisons across teachers.

District and state officials interested in using student learning objectives for evaluative purposes might consider whether frontline educators have the time and the skills to create student learning objectives that permit reliable, valid, and fair comparisons across teachers

Appendix A. Study sample selection, data collection, and analysis

This appendix provides details on the study sample selection, data collection, and analysis.

District sample selection

The study team used a recent literature review (Gill et al., 2013) and the National Council on Teacher Quality (n.d.) online database to identify an initial pool of 13 districts that use alternative student growth measures to assess teacher effectiveness. The sample was not limited to districts in the Mid-Atlantic Region because so few of them have experience with alternative student growth measures and because the issues are not unique to the region. District websites were examined to identify the measures in use, the first year of implementation, and the purpose of the measure. To ensure the inclusion of at least one district using each measure, potential sample districts were grouped by the type of measure used and then ranked within each type based on the year the measure was first implemented; districts with the longest period of implementation within each type of measure were prioritized. Rather than being selected as a representative sample, districts were selected for length of implementation because they were more likely to have progressed beyond a pilot stage and would have better insight into medium- and long-term benefits and challenges. Four of the districts that used a value-added model and four of the districts that used student learning objectives had at least two years of implementation experience as of the 2012/13 school year, and they agreed to participate. The districts in the sample are located in eight states.

Respondent sample selection

One administrator from each district was interviewed during an initial pilot period in the 2012/13 school year. After obtaining Office of Management and Budget clearance, the study team recruited the eight districts in the pilot sample to participate in a second round of interviews with multiple types of respondents in the 2013/14 school year. The study team completed applications to conduct research in each of the eight districts and received approval for five of the eight districts. The study timeline did not permit recruitment and research application in any additional districts to replace the three districts that did not grant approval.

Four of the eight districts provided permission for the study team to interview school principals and teachers. In these four districts the study team worked with district administrators (and sometimes with teachers' union representatives) to identify a purposive sample of school principals with varied success in implementation and varied perspectives on the alternative student growth measures. To identify teachers, the study team asked for suggestions from school leaders—and in some cases, district administrators and teachers' union representatives—of teachers with both positive and negative experiences with the alternative student growth measures. The study team then targeted one district administrator and one teachers' union representative to interview in each of the districts and recruited from among suggested principals and teachers with the aim of interviewing two to three principals and three to four teachers in each district. In each of the four districts, one district administrator, between one and four principals, and between one and five teachers were interviewed. Teachers' union representatives were interviewed in three of the four districts.

A fifth district provided access to district administrators who were involved in implementing the alternative student growth measures. In this district the study team interviewed seven district administrators, including curriculum leads who work closely with district teachers. The study team also interviewed a teachers' union representative.

In the three remaining districts the study team was unable to interview principals and teachers, so it targeted representatives from teachers' unions instead. The team interviewed teachers' union representatives in one of the three districts. In the other two districts the study team used only the interviews completed with district administrators during the pilot period to address the research questions.

In total, the study team interviewed 18 district administrators, 11 principals, 14 teachers, and 6 teachers' union representatives across the eight districts (table A1). Although the sample was limited, interviews with multiple stakeholders in each district, supplemental document reviews, and cross-site analyses allowed for triangulation to enhance the validity and reliability of the findings (Merriam, 2001; Wiersma, 2000).

Data collection and analysis

The interviews were conducted during district site visits and by phone (see appendixes E–H for the interview protocols). Each interview lasted 45–60 minutes. Interviews were recorded, with the permission of the respondents, and transcribed. The interviewers also reviewed documents collected on district websites and from district administrators, such as evaluation reports, teacher handbooks, and lists of frequently asked questions, and included this information with citations in the interview transcripts. The interview transcripts were then coded in Atlas.ti, using a coding scheme that had been finalized after two rounds of piloting and subsequent revisions. Codes were grouped into families that corresponded to the primary research questions. All coders coded three of the same interviews at the beginning of the coding process to ensure a sufficient level of consistency; a senior study team member also checked coding at several points throughout the process.

Table A1. Stakeholder respondent samples, by district

District	Measure	District administrators (n = 18)	Principals (n = 11)	Teachers (n = 14)	Teachers' union representatives (n = 6)
A	Student learning objectives	1			
B	Student learning objectives	1			
C	Student learning objectives	2	3	5	
D	Student learning objectives	2	4	4	1
E	Alternative assessment-based value-added model	7			1
F	Alternative assessment-based value-added model	2	1	1	1
G	Alternative assessment-based value-added model	1			2
H	Alternative assessment-based value-added model	2	3	4	1

Note: Because many of the districts requested anonymity, this report does not identify the districts.

Source: Authors' compilation.

The study team then used the coded data to triangulate findings within districts and to identify and examine variations and commonalities across districts and types of measures along the dimensions of implementation raised by the research questions: uses of measures, differentiation in teacher performance, benefits, costs or drawbacks, implementation challenges, and contextual factors.

Appendix B. How the sample districts incorporate alternative student growth measures into teacher evaluation systems

This appendix describes the sample district approaches to integrating alternative student growth measures into their teacher evaluation systems as well as how the districts combined the measures with other measures of teacher performance to produce overall evaluation ratings.

Districts that included alternative student growth measures in their fully implemented teacher evaluation systems had at least four to five years of experience and had integrated the measures into district policies and practices

During the 2013/14 school year three districts had implemented teacher evaluation systems that included student learning objectives or an alternative outcome-based value-added model. Among these districts, one included student learning objectives in evaluations for all teachers, and two included an alternative assessment-based value-added model in evaluations for subsets of teachers. One of the latter districts was also implementing student learning objectives in its teacher evaluation system to comply with state requirements and to include student growth in evaluation ratings for an even broader population of teachers. All three districts had at least four years of implementation experience as of 2013/14.

These districts refined their approaches to using alternative student growth measures for teacher evaluation over time. One district readjusted the weight of the measure in teacher evaluation in response to state legislation requiring that 50 percent of a teacher's evaluation rating be based on student growth. Another district initially used student learning objectives in place of value-added scores for teachers in grades and subjects not tested by the state. Based on feedback from teachers who were evaluated using student learning objectives, the district reduced the weight of student learning objectives for teachers without other measures of student growth and extended it as a requirement for all teachers in the district, regardless of grade and subject taught (reducing the weight of individual state assessment-based scores for teachers who received them). The district reported that adding student learning objectives for teachers who also received value-added model scores was perceived as a positive change that made the evaluation system more equitable for different types of teachers. District administrators indicated that extending student learning objectives to all teachers gave all teachers in the district a greater sense of agency in their evaluation ratings. However, district administrators also expressed some concern regarding the rigor of student learning objectives and quality control measures when compared to state assessment-based value-added models.

Two of the three districts also refined their general approaches to implementing alternative student growth measures. The district that used student learning objectives increased the number and range of off-the-shelf assessments available to teachers for use in student learning objectives. By increasing assessment capacity at the district level and reducing the use of assessments developed by individual teachers, district administrators hoped that student learning objectives would be more consistent and rigorous across teachers and schools, since teachers could rely on assessments available districtwide across a broader scope of grades and subjects. One of the other two districts also refined the processes associated with assessment administration for its alternative assessment-based value-added model. The district introduced additional test security procedures to ensure that end-of-course

assessments were administered consistently across the district. The district also added a roster confirmation process to ensure that students were properly linked to their teachers in the value-added model.

Though administrators from these three districts continued to refine the implementation of alternative student growth measures, district and teachers' union respondents in one district (where principals and teachers were not interviewed) and district administrator, principal, and teacher respondents in the other two districts perceived the measures as completely integrated into the districts' overall approaches. Principals and teachers reported that because the measures became part of the overarching culture and theory of action in these districts, principals and teachers required minimal ongoing training or professional development related to the measures. However, new teachers received information regarding the measures during their orientation.

Districts that were concerned about stakeholder resistance to the use of alternative student growth measures in teacher evaluation used a piloting phase to allay concerns

Four districts (two that used student learning objectives and two that used an alternative assessment-based value-added model) were piloting the alternative student growth measures for use in teacher evaluation (as opposed to piloting the measures in general). The districts were piloting the measures in one of two ways: implementing an alternative assessment-based value-added model in all schools for informational purposes but not attaching stakes to performance or implementing student learning objectives in a subset of schools with some stakes attached, such as additional compensation. This section describes the two piloting approaches and the implications for future use of the measures in teacher evaluation systems.

Two districts were piloting an alternative assessment-based value-added model in all schools. Two districts that used an alternative assessment-based value-added model opted to pilot the measure in all their schools before using them in teacher evaluation ratings. In these districts teachers received their scores on the measure and might have participated in conversations with their school leaders regarding performance on the measures during the pilot phase. Administrators in both districts reported that providing results to teachers was intended to help them understand the measure before stakes were attached to it. District administrators indicated that the pilot period was intended to reduce teacher anxiety about using the measures for evaluation. These district respondents hoped that providing information on student growth would encourage the use of data to drive instruction and help teachers improve their instructional approaches.

One of the districts that was applying the piloting approach used the measure for rewards but not for negative consequences during the pilot period. (The district rewarded teachers who were identified as high performing with additional compensation or career opportunities.)

Two districts were piloting student learning objectives in a subset of schools. The remaining two districts, which were both implementing student learning objectives, were piloting the measure in a subset of schools or with a subset of teachers, with the intention of implementing student learning objectives districtwide following the pilot. These districts were already using the measure in a performance-based compensation system, but the

districts were planning to extend use of the measures to teacher evaluation as one among multiple measures of teacher performance. Although these districts used similar piloting approaches, one district's approach involved modifying an existing student learning objective measure. In this district student learning objectives had been used districtwide for nonevaluative purposes. The district was working to refine the measure and its use prior to bringing its implementation to scale. The other district first introduced student learning objectives for use in a performance-based compensation system in a subset of high-need schools. The district was piloting a teacher evaluation system using multiple measures of performance, including student learning objectives, in the same subset of schools. The district intended to implement the evaluation system districtwide following the pilot period.

Districts were in the process of finalizing how to incorporate measures in teacher evaluation systems. Regardless of the approach to piloting alternative student growth measures, the four districts were negotiating how the measures would be integrated into the evaluation systems with relevant stakeholders, including teachers' unions. Two districts engaged the teachers' union, as well as teachers and principals, in the design of their alternative student growth measures. Although these respondents did not have reservations about the design of the measures, teachers' union representatives reported concerns about the use and weights of the measures for evaluative purposes. In spring 2014, district administrators and teachers' union representatives in both districts were working toward a resolution.

Districts in the piloting stage were also addressing state legislation that specified the elements that must be included in teacher evaluation systems. In three of these districts, states enacted legislation requiring 50 percent of a teacher's evaluation rating to consist of measures of student growth. These districts—two used an alternative assessment-based value-added model and one used student learning objectives—were working with stakeholders to determine a fair mix of student growth measures to meet the 50 percent requirement. In one of the two districts that used an alternative assessment-based value-added model, a teachers' union representative expressed concern that districts would place too much weight on value-added results relative to other student outcome measures in order to comply with state law. In contrast, the teachers' union in the other district pushed for value-added results to be added to evaluations immediately. In this state, legislation requires the use of student growth measures in teacher evaluation systems but does not specify how districts should meet this mandate. The teachers' union supported taking advantage of the lack of specificity by including value-added results in teacher evaluation immediately to ensure that teachers would be evaluated using established, district-developed measures of student growth rather than using state-imposed measures of student growth.

In the district that used student learning objectives, a teachers' union respondent expressed concern about the weight of student learning objectives in teacher evaluation because the teachers' union considered the measure to be a tool for collaboration rather than evaluation. The respondent asserted that differences in the intensity of guidance and support provided to teachers by their school leaders could result in large differences in student learning objective rigor, making the measure inappropriate for use in teacher evaluations. The district was sensitive to this concern, and district administrators reported that they were developing quality control measures to ensure consistent implementation and rigor across schools before using student learning objectives in teacher evaluations.

Teacher evaluation systems in development relied on multiple measures of teacher performance, which varied both within and across districts

The seven districts currently using or planning to use alternative student growth measures for teacher evaluation were combining them with other measures of teacher performance to create overall evaluation ratings. The other measures varied by district and included state assessment–based value-added models (at the individual or school level) or other statistical measures of student growth, classroom observations, student surveys, and measures of professionalism. These additional measures varied not only across the sample districts but also within districts based on grade and subject taught. Among the three districts that had already incorporated alternative student growth measures in their evaluation systems, one district evaluated all teachers using student learning objectives, and another planned to use student learning objectives to evaluate all teachers, although the weight of the measure will vary depending on the availability of a value-added model for the teacher. In the third district all teachers except kindergarten teachers were evaluated using a value-added model, but the type of assessment (state assessment or one of several alternative assessments) included in the value-added model varied by grade and subject. A summary of the measures used in teacher evaluation systems across the sample districts appears in table B1.

Across the three districts that had already integrated alternative student growth measures in their evaluation systems, the weights vary by measure type. One district included student learning objectives as 15 percent of all teachers’ evaluation ratings, and the weights of other measures varied by grade and subject taught. Another district included alternative assessment–based value-added model results in evaluation ratings along with a state assessment–based value-added model and added student learning objectives to evaluation ratings (box B1). In this district the weight of value-added results depended on grade and subject taught but never exceeded 26 percent. Student growth, measured through value-added and student learning objectives, accounted for 50 percent of teachers’ evaluation ratings in total. The third district that integrated alternative student growth measures into

Table B1. Measures used or proposed for use in teacher evaluation systems, as of 2013/14 school year

District	Student growth measures				Professional practice measures		
	Individual state assessment–based value-added model	Individual alternative assessment–based value-added model	Individual student learning objective	Team or school measures of growth	Observation	Student surveys	Other measures of professional practice
B			✓	✓	✓	✓	✓
C	✓		✓		✓		✓
D	✓		✓	✓	✓	✓	✓
E	✓	✓			✓		
F	✓	✓	✓		✓		
G	✓	✓			✓		
H	✓	✓	✓		✓	✓	✓

Note: Excludes one sample district that is no longer implementing student learning objectives.

Source: District site visits and stakeholder interviews and publicly available documentation on district websites.

Box B1. Adding student learning objectives alongside alternative assessment-based value-added measures

Two sample districts that were already implementing alternative outcome-based value-added measures were also introducing student learning objectives. In both districts state legislation mandated the use of student growth in evaluation ratings for all teachers. In the alternative assessment-based value-added model, one district developed its own assessments, and the other district purchased nationally normed assessments for use in grades and subjects not tested by state assessments. The alternative assessments enabled the districts to produce value-added results for a broader population of teachers than just those in state-tested grades and subjects. However, the alternative assessments (both commercially purchased and home-grown) still applied only to a subset of the teachers who were not covered by state assessments. Consequently, these districts adopted student learning objectives to measure growth for all teachers. Although all teachers completed the student learning objective process in these districts regardless of grade or subject taught, student learning objective inclusion in evaluation ratings for teachers who received value-added results was still undecided. For these districts student learning objectives provided an option for districts to comply with state law while encouraging data use among all teachers.

teacher evaluation ratings applied the same weight to value-added results regardless of the assessment used in the model or the grade and subject taught. In this district value-added results constituted 40 percent of a teacher's evaluation rating. Districts that were still pilot testing in the 2013/14 school year had not finalized the weights of each measure in their teacher evaluation systems.

Appendix C. Benefits of using alternative student growth measures for student populations with special needs

This appendix describes contextual considerations around the benefits of using alternative student growth measures that are specific to teachers of student populations with special needs.

The sample districts applied alternative student growth measures to teachers of all students, including students in special education, English learner students, and talented and gifted students. However, assessing these student populations, particularly if the district used student learning objectives, required additional considerations. Across the four districts that used student learning objectives, two special education teachers and one teacher of English learner students were interviewed. One was an inclusion classroom teacher, one taught only English learner students using a “pull-out” model, and the third taught with both the inclusion and pull-out model (some students stayed in the same classroom; others were in the classroom for only half a day).

All three teachers collaborated with the grade-level student learning objective teams to ensure that they were using similar grade standards and assessments. However, the growth targets varied and considered the special learning needs of the student populations. One teacher with an inclusion classroom noted this as a benefit. The teacher found it more difficult to see growth from students in special education on state assessments and appreciated seeing growth and progress more clearly through the student learning objective process.

Inclusion teachers also appreciated the opportunity to collaborate with other teachers on student learning objective standards and assessments. But the teacher in the pull-out model experienced no benefits related to collaboration as a result of the student learning objective process and struggled to implement student learning objectives in isolation.

Appendix D. Contextual factors influencing implementation: funding, policies and politics, and teachers' unions

This appendix describes the external factors that the eight sample districts reported played a role in the implementation of alternative student growth measures.

External funding

All eight districts received financial support from external funders to measure student growth and used the funding to design and implement the alternative student growth measures. Five of the districts received funding from the U.S. Department of Education's Teacher Incentive Fund to develop and implement performance-based teacher and principal compensation systems in high-need schools. Additional sources of funding included the federal Race to the Top initiative and local or national foundations (including the Bill & Melinda Gates Foundation). Administrators from each district reported that they would not have been able to implement the alternative student growth measures without the additional funding; at least three districts expressed concerns about the sustainability of the measures in the district without continued external support.

Policies and politics

Respondents in four districts reported that recent state policy changes related to teacher evaluation systems had resulted in some adjustments to their implementation of alternative student growth measures. Two of the districts—both of which used an alternative assessment-based value-added model—were piloting student learning objectives and planned to implement student learning objectives alongside an alternative assessment-based value-added model to comply with state requirements. Both districts expressed frustration with unclear or frequently changing state guidelines, which made implementation more challenging. One other district was struggling with the state's directive to use an alternative assessment-based value-added model to evaluate teachers when the assessment administered by the district had not initially been intended for use in teacher evaluation.

In another district an administrator attributed the cessation of student learning objectives to changes in district administration. The administrator noted that the philosophy of the recently appointed superintendent did not align with the use of student learning objectives. In addition, the district anticipated new guidelines from the state on using student growth to evaluate teachers and preferred to drop other measures of growth until receiving this guidance.

Relatedly, policies and school culture around data use also emerged as a key factor influencing implementation of alternative student growth measures in the sample districts (box D1).

Teachers' unions

In six of the eight districts, district administrators, teachers, or teachers' union representatives reported that the teachers' union had been involved or highly involved in designing the alternative student growth measure and continued to play an active role in monitoring and adapting the measure for teacher evaluation. Four districts used an end-of-course

Box D1. Culture of data use and implementation of student learning objectives

Among the sample districts, school culture concerning data emerged as a key factor in facilitating or impeding implementation of alternative student growth measures with the intent to ultimately improve instruction and increase student achievement. District administrators, teachers, and principals from all four districts that used student learning objectives indicated that school leaders determined whether student learning objectives informed instructional practice and served as more than compliance exercises. Among schools with an established culture of data use, implementation of student learning objectives was considered “business as usual.” In these schools, school leaders established the structures necessary for teachers to effectively work with student data to inform their practices. Because practices such as shared planning time already existed in these schools, implementation of student learning objectives could occur within the existing school schedule.

According to district administrators, principals, teachers, and teachers’ union respondents, schools without a strong culture of data use tended to struggle more with implementing student learning objectives, which prevented student learning objectives from influencing instructional practice. Within these schools, policies were not in place to foster collaboration and data use, and the implementation of student learning objectives consequently focused on compliance rather than instructional improvement. In these schools, teachers tended to examine student data only at the beginning and end of the year because they lacked the time or support to collaboratively look at data to improve their instruction.

Even in schools with an established culture of data use, principals and teachers reported that time presented a serious barrier to implementation of student learning objectives. Although these schools may use common planning periods to discuss student data and instructional improvement, student learning objectives themselves may require teachers to spend considerable additional time administering and scoring assessments, monitoring progress, and analyzing data.

Ultimately, school leaders established the tone for creating a collaborative culture around student data use. In schools with a strong culture of data use, school leaders worked with teachers to establish student learning objective goals and provide teachers with ongoing feedback regarding goal progress. Yet providing this kind of feedback is both labor and time intensive. Districts must also consider how to ensure that structures are in place to foster data use among teachers who do not have school-level peers. This might mean release time to work with subject matter peers from other buildings, as suggested by one district respondent. Policies such as integrating performance on alternative student growth measures into school leader evaluation ratings can give school leaders an incentive to engage, and policies such as common planning time can relieve some, but not all, of the student learning objective time burden.

assessment-based value-added model, and two districts used student learning objectives. In one district that used an end-of-course assessment-based value-added model and that was located in a right-to-work state (that is, a state in which teachers are not required to join a union and there is no collective bargaining), the teachers’ union representative met regularly with district administrators to resolve issues related to measure implementation and assisted in communicating with teachers. In the other two districts that used an alternative assessment-based value-added model, the teachers’ union representatives reported being less supportive of using the measures in teacher evaluation. Both representatives noted that the debate about using value-added models in teacher evaluation had become

much more contentious across the country, with shifts in support by national teachers' unions. The backlash was perceived to be making implementation more difficult than when the measures were initially introduced in the districts.

Teachers' unions in districts that used an alternative assessment-based value-added model were more likely than those in districts that used student learning objectives to take a stance on issues related to the design of the model or the use of the measure when negotiating with the district. For example, in one district the teachers' union insisted on a minimum number of years of teacher data in the value-added model when the district proposed to use fewer years of teacher data to maximize the number of teachers for whom value-added model scores could be estimated. In another district the teachers' union agreed to the use of an alternative assessment-based value-added model in teacher evaluation scores—a state requirement—but stipulated in the contract that the value-added model not be used for hiring or firing decisions. In a third district the teachers' union successfully pushed a requirement that principals not have access to teacher value-added model results until after evaluations of instructional and professional practice are completed. (In this district the individual teacher value-added model results from the previous school year are released in the fall; principals do not have access to the scores until after evaluations are completed in the spring, although they do have access to the school-level value-added model results.)

Among the two districts that reported more limited involvement of the teachers' union, both used student learning objectives, and one is in a right-to-work state, where teachers do not collectively bargain with the district.

Appendix E. District administrator interview protocol

A. Introduction and consent for audio recording

The purpose of this section is to introduce the respondent to the study and request consent for audio recording.

1. Thank you for taking the time for the interview today. My name is [name] and I work for Mathematica Policy Research, which is conducting this study for the Mid-Atlantic Regional Educational Laboratory (REL). As I mentioned [in my email/when we spoke on the phone], we are conducting case studies of districts implementing measures of teacher performance that include student growth but that measure growth on outcomes other than state assessments. These alternative outcome measures might include nationally normed assessments, curriculum-based assessments developed by the district, or “student learning objectives” (SLOs) developed by individual teachers for their own students. Student growth might be measured through a simple subtraction of end-of-year achievement from beginning-of-year achievement, through a student growth percentile method, through a value-added statistical model, or implicitly in SLOs that are based on an understanding of baseline achievement levels in a class at the beginning of the school year. We plan to examine which alternative outcome measures are used, how the growth measures are implemented, the challenges and obstacles in implementation, how the measures are being used, and the effectiveness of these measures in differentiating teacher performance. We expect our findings to benefit states and districts considering the use of such alternative growth measures. The purpose of this interview is to gather information on the implementation of [insert specific growth measures used in the district] in the district from your perspective as a district administrator. We will be interviewing teachers, principals, and union representatives in the district to gather information on their perspectives as well. We will not identify you by name in the study reports.

Before beginning, I would like to get your permission to record the interview. This will ensure that we accurately capture information about district policies and will prevent me from having to ask questions and take notes at the same time. The audio recordings from all individuals interviewed within your district will be compiled into a single district summary that is for internal analysis purposes only and will not be published or shared with anyone outside the study team. You will not be identified in the district summary. This summary will then be aggregated with other district summaries and analyzed for trends and variations across multiple districts. Do I have your permission to record the interview?

B. Type of student outcome and growth measure used

1. Please describe the student outcome measure(s)—other than state assessments—used in growth measures of teacher performance implemented by the district.
 - a. Prompt: Does the district use any student assessments developed by the district or schools in the district?
 - b. Prompt: Does the district use SLOs for purposes of measuring student growth associated with individual teachers?
 - c. Prompt: Does the district use any end-of-year, nationally normed, or summative assessments purchased from testing companies (not including interim or benchmark assessments)?

2. In what subjects and grades are these measures implemented? Do these measures apply to teachers of English language learner students and special education students or students with IEPs? For how many school years has each measure been in place?
3. (If the district uses SLOs) Please briefly describe any formal guidelines the district provides to teachers or principals regarding the development of SLOs. Do teachers receive any training or professional development on developing SLOs? How about principals? Is there a districtwide template used by teachers? What is the process for approving SLOs submitted by teachers? Are teachers required or permitted to submit more than one SLO for a class? Does this vary by grade and subject area? Does the district have any way to determine whether the SLOs are of comparable difficulty for different teachers?
4. I would like to learn a bit more about how the alternative measure(s) used in the district was designed. What was the process for deciding on the specific measure currently implemented in the district? Did the district use examples from other districts or states? Was the student outcome measure administered in the district prior to its use in growth or value-added teacher performance measures? If not, was there a piloting period for the alternative assessment? How was it selected or designed?
5. Does the district use a value-added (VA) model for assessing teacher performance (aside from school accountability)? If not, how does the district measure growth (e.g., percentile method, simple subtraction)? How long has the growth/VA model been in place? Has the district made any changes to this approach to measuring growth during that period?
6. (If the district uses a widely recognized statistical model for measuring growth/value-added) What is the name of the growth/VA model and/or who designed it? Is the district using the same statistical model for alternative measures as it uses for state assessments?
7. Prior to the implementation of this alternative growth model, how did the district measure teacher performance (for example, measures of professional practice via classroom observations)? Is the district continuing to use these measures in addition to the alternative measure(s) you've described?

C. Implementation of data collection and analysis

1. I would like to learn more about the administration of these alternative student assessments in the district. (For each assessment) How is the assessment administered (for example, timing, frequency, who administers)? What is the process for collecting assessment results?
2. What is the procedure for analyzing assessment results and assigning teacher performance ratings? Who conducts the analysis? Is there a department in the district specifically responsible for the analysis? If not, does the district use an external vendor for this task?

3. Please describe any quality control measures you have in place for the collection and analysis of the teacher performance data (for example, use of external vendors for test administration/data collection, evaluation of model or VA estimate bias by external consultant, auditing of teacher SLO documentation).
4. What kinds of obstacles, if any, have you encountered in collecting and analyzing this data? For example, have you found you have sufficient staff and resources? (If applicable) Has the district encountered any difficulties collecting documentation of SLOs from teachers?
5. How has the district addressed these issues or challenges? From your perspective, has the district made any significant improvements to its approach?

D. Differentiation produced by the alternative growth or value-added measure

1. Please describe the scale for rating teacher performance on the alternative measures. Is it binary (reached target or not)? Categorical (for example, beginning, developing, proficient, advanced)? Continuous (for example, percentile, NCE, z-score)? (If continuous) What type of rating/estimate is used? (If categorical) What are the categories used?
2. Please describe the distribution of teacher performance across these ratings for each year the alternative measure has been in place. What proportion of teachers is in each category? (If respondent is unsure across all categories) Approximately what proportion of teachers is in the top and bottom categories? Are there any reports or documents you can provide that might offer some additional detail? If not, is there another staff member who might be able to provide more information?
3. Please describe the scale for rating teacher performance on any other measures currently or previously used in the district (for example, measures of professional practice or state test-based VAM estimate—use response to B.5. to frame question appropriately). Has the district made any effort to compare teacher performance across the different measures currently or previously used in the district? If so, how does the distribution of teacher performance on these current measures compare to the distribution of teacher performance on the alternative measures that you previously described? How does the distribution of teacher performance under your district's old system compare to the distribution of teacher performance on the alternative measures?
4. Do you have any evidence on the proportion of teachers whose scores on the alternative measure are statistically distinguishable from average?
5. Do you have any information about the reliability or validity of the results? In other words, has the district examined teacher scores across multiple years or measures? Do individual teachers tend to earn similar scores from one year to the next? Do individual teachers tend to earn similar scores across different measures (within the same year)?

E. Use of alternative growth or value-added measure

1. For what purposes is the district using each alternative student outcome measure used in assessing teacher performance?
 - Is it used for evaluating teachers? In what grades and subject areas?
 - Is it used as part of a performance-based compensation system? In what grades and subject areas?
 - Is it used as a means of targeting professional development to specific teachers?
 - Is the measure used to make decisions about teacher assignments?
 - Are there any other ways in which the district is using these measures?
 - If there are others in the district who can provide additional information on this issue, can you give us their contact information?
2. What type of feedback is provided to teachers regarding their performance on each measure? When is this feedback provided? Who provides the feedback? Is it provided in writing? In person?
3. If used in teacher evaluation, is the measure combined with any other measures of teacher performance (for example, value-added or other growth measures based on state assessments or classroom observation measures)? How much weight does each performance measure receive in a teacher's overall evaluation rating? Does the use of the measure in teacher evaluation vary by grade and/or subject area?
4. Does a teacher's performance rating on the alternative measure affect whether he or she is awarded tenure? Does it affect whether a teacher is targeted for additional monitoring and support, such as professional development, coaching, or additional interim evaluations?

F. Benefits and drawbacks

1. Could you summarize some of the main benefits, from your perspective, of implementing this alternative VAM or growth measure of teacher performance in the district? For example, have you been able to measure the performance of a larger percentage of teachers in the district? How have teachers in the district responded to the use of alternative student outcomes to measure their performance?
2. What do you perceive as the primary costs of implementing the measure? Are there specific monetary costs to the district (for example, administering additional tests and collecting data)? Has the district encountered any obstacles implementing the technical aspects of collecting and analyzing the teacher performance data? Has the measure allowed for a sufficient level of comparability across teachers? If SLOs are used in the district, has the district been able to ensure that all teachers are setting sufficiently high standards for student growth?
3. Could you describe the district's approach to communicating with the various stakeholders as the alternative measure was rolled out? How do you perceive the district's communication of the use of the alternative measure in evaluations to principals, teachers, and the teachers' union/association?

G. District context

1. Have there been any changes in the district administration or policies that may have affected the implementation of these alternative measures of teacher performance (for example, change in superintendent or implementation of a performance-based pay system)?
2. Please describe the teachers' union/association's involvement in the development of the VAM or growth measure. To what extent did the district engage the union/association in development and implementation? Did the district have the support of the teachers' union/association for the implementation of the VAM or growth measure of teacher performance using an alternative student outcome? Why or why not? (If applicable) How did the union/association respond to the incorporation of the VAM or growth measure into the teacher evaluation system?
3. Has the district received funding or resources from external sources (such as federal government, foundations) to develop or implement this measure?
4. Is there anything else that you think might be important to know about the development and implementation of this alternative growth or value-added measure in the district? Are any other key stakeholders in the district with whom you think it would be important for us to speak?

Appendix F. Principal interview protocol

A. Introduction and consent for audio recording

The purpose of this section is to introduce the respondent to the study and request consent for audio recording.

1. Thank you for taking the time for the interview today. My name is [name] and I work for Mathematica Policy Research, which is conducting this study for the Mid-Atlantic Regional Educational Laboratory (REL). The Mid-Atlantic REL's mission is to "build research alliances with education practitioners and policymakers to help states and districts use data to answer important issues of policy and practice." The results of this study will inform publicly available reports aimed at states and districts considering implementing measures of teacher performance that include student growth. As I mentioned [in my email/when we spoke on the phone], we are conducting case studies of districts implementing measures of teacher performance that include student growth, but that measure growth on outcomes other than state assessments. These alternative outcome measures might include nationally normed assessments, end-of-course curriculum-based assessments, or "student learning objectives" (SLOs) developed by individual teachers for their own students. Student growth might be measured through a simple subtraction of end-of-year achievement from beginning-of-year achievement, through a student growth percentile method, through a value-added model, or implicitly in SLOs that are set based on an understanding of baseline achievement levels in a class at the beginning of the school year. We plan to examine what alternative outcome measures are used, how the growth measures are implemented, challenges and obstacles in implementation, how the measures are being used, and the effectiveness of these measures in differentiating teacher performance. We expect our findings to benefit states and districts considering the use of such alternative growth measures. The purpose of this interview is to gather information on the implementation of [insert specific growth measures used in the district] in the district from your perspective as a principal. We will be interviewing teachers, principals, and union representatives in the district to gather information on their perspectives as well. We will not identify you by name in the study reports. We also will not identify the districts included in the study in subsequent reports.

Before beginning, I would like to get your permission to record the interview. This will ensure that we accurately capture information about district policies and will prevent me from having to ask questions and take notes at the same time. The audio recordings from all individuals interviewed within your district will be compiled into a single district summary that is for internal analysis purposes only and will not be published or shared with anyone outside the study team. You will not be identified in the district summary. This summary will then be aggregated with other district summaries and analyzed for trends and variations across multiple districts. Do I have your permission to record the interview?

B. Type of student outcome and growth measure used

1. Are teachers in your school measured based on [identify the alternative outcome growth measure(s) discussed in district interview]?

2. (If the district uses SLOs) As the principal, what role do you play in the development of SLOs? Is there a districtwide template used by teachers? What is the process for approving SLOs submitted by teachers? Are teachers required or permitted to submit more than one SLO for a class? Does this vary by grade and subject area? From your perspective, are SLOs of comparable difficulty across all teachers in your school? Do SLOs in your school involve classwide targets or separate targets for each student? How much time do you devote to the development of SLOs?
3. Can you describe the process for rolling out the alternative measure(s) in the district? What role did you play in the decisionmaking process? Who else was involved in the decisionmaking process? Did you receive any training or professional development related to the implementation of the measure?

C. Implementation of data collection and analysis

1. I would like to learn more about the administration of [fill in the assessment name or SLOs] in the district. (For each assessment) How is the assessment administered (for example, timing, frequency, who administers)? What is the process for collecting assessment results?
2. Are you aware of any quality control measures in place in the district for the collection and analysis of the teacher performance data collected? (If district uses SLOs) For example, what documentation of student performance do your teachers need to submit for your SLOs, and are teachers subject to any audits, such as auditing SLOs to ensure rigor? (If district uses a VAM/growth model) For example, do teachers administer or score the student assessments, or is administration and/or scoring conducted by independent proctors or an external firm?
3. What kinds of obstacles, if any, have you encountered in preparing for and administering the non-state student assessments or SLOs? How many extra hours, if any, have you had to devote to these tasks, per teacher? Does the amount of time devoted to these tasks vary by teacher?
4. Have there been any changes to the district's approach to using these student growth measures? Have these changes resulted in notable improvements?

D. Differentiation produced by the alternative growth or value-added measure

1. What proportion of teachers in your school is in each performance category according to the alternate growth measure? (If respondent is unsure across all categories) Approximately what proportion of teachers is in the top and bottom categories?
2. Have you found the performances of individual teachers in your school to be consistent across years and measures? Does the measure accurately reflect teachers' contributions to student learning? Is the measure capturing the appropriate student learning standards?

E. Use of alternative growth or value-added measure

1. For what purposes is your school using each alternative student outcome measure used in assessing teacher performance? Is it used for evaluating teachers? In what grades and subject areas? Do teachers receive targeted professional development based on their performance? Does the measure affect teacher assignments? Are there any other ways in which you are using these measures (for example, tenure)?
2. Is the performance of teachers within your school used for evaluating your performance as a principal? Have you taken any schoolwide or teacher-level actions as a principal based on this performance information?
3. Do teachers receive any feedback regarding their performance on each measure? What type of feedback, and when is it provided?

F. Benefits and drawbacks

1. Could you summarize some of the main benefits, from your perspective, of implementing this alternative VAM or growth measure of teacher performance in your school? For example, have you found that receiving this kind of information about teacher performance in your school has had any benefits for your effectiveness as a principal?
2. What do you perceive as the primary drawbacks or costs (for example, time) of implementing the measure? Do you think the measure allows for a sufficient level of comparability across teachers on performance? If SLOs are used, do you think all teachers are setting sufficiently high standards for student growth?
3. How do you perceive the district's communication of the use of the alternative measure in evaluations to teachers and principals?

G. District context

1. Have there been any changes in the district administration or policies that may have affected the implementation of these alternative measures of teacher performance (for example, change in superintendent or implementation of a performance-based pay system)?
2. Is there anything else that you think might be important to know about the development and implementation of this alternative growth or value-added measure in the district? Are any other key stakeholders in the district with whom you think it would be important for us to speak?

Appendix G. Teacher interview protocol

A. Introduction and consent for audio recording

The purpose of this section is to introduce the respondent to the study and request consent for audio recording.

1. Thank you for taking the time for the interview today. My name is [name] and I work for Mathematica Policy Research, which is conducting this study for the Mid-Atlantic Regional Educational Laboratory (REL). The Mid-Atlantic REL's mission is to "build research alliances with education practitioners and policymakers to help states and districts use data to answer important issues of policy and practice." The results of this study will inform publicly available reports aimed at states and districts considering implementing measures of teacher performance that include student growth. As I mentioned [in my email/when we spoke on the phone], we are conducting case studies of districts implementing measures of teacher performance that include student growth, but that measure growth on outcomes other than state assessments. These alternative outcome measures might include nationally normed assessments, end-of-course curriculum-based assessments, or "student learning objectives" (SLOs) developed by individual teachers for their own students. Student growth might be measured through a simple subtraction of end-of-year achievement from beginning-of-year achievement, through a student growth percentile method, through a value-added model, or implicitly in SLOs that are set based on an understanding of baseline achievement levels in a class at the beginning of the school year. We plan to examine what alternative outcome measures are used, how the growth measures are implemented, challenges and obstacles in implementation, how the measures are being used, and the effectiveness of these measures in differentiating teacher performance. We expect our findings to benefit states and districts considering the use of such alternative growth measures. The purpose of this interview is to gather information on the implementation of [insert specific growth measures used in the district] in the district from your perspective as a teacher. We will be interviewing teachers, principals, and union representatives in the district to gather information on their perspectives as well. We will not identify you by name in the study reports. We also will not identify the districts included in the study in subsequent reports.

Before beginning, I would like to get your permission to record the interview. This will ensure that we accurately capture information about district policies and will prevent me from having to ask questions and take notes at the same time. The audio recordings from all individuals interviewed within your district will be compiled into a single district summary that is for internal analysis purposes only and will not be published or shared with anyone outside the study team. You will not be identified in the district summary. This summary will then be aggregated with other district summaries and analyzed for trends and variations across multiple districts. Do I have your permission to record the interview?

B. Type of student outcome and growth measure used

1. What measures of student growth are used in assessing your performance?
 - a. Prompt: Does the district use any end-of-course curriculum-based assessments?

- b. Prompt: Does the district use SLOs for purposes of measuring student growth associated with individual teachers?
 - c. Prompt: Does the district use any end-of-year, nationally normed or summative assessments purchased from testing companies (not including interim or benchmark assessments)?
2. (If the district uses SLOs) Please briefly describe any formal guidelines the district provides to you regarding the development of SLOs. Is there a districtwide template used by teachers? What is the process for having your SLOs approved? Are you required or permitted to submit more than one SLO for a class? Do your SLOs involve classwide targets or separate targets for each individual student? From your perspective, are SLOs of comparable difficulty across all teachers in the district?
 3. (If the district uses alternate assessments to measure growth) Does the district use a growth or value-added model, applied to a measure of student achievement other than the state assessment, to measure your performance? If yes, what student assessment is used?
 4. What role did teachers play in the alternative measure(s) design process? Can you describe the process for rolling out the alternative measure(s) of student growth (including SLOs) in the district?
 5. Did you receive any training or professional development related to the implementation of the measure?

C. Implementation of data collection and analysis

1. I would like to learn more about the administration of these student assessments in the district. (For each assessment) How is the assessment administered (for example, timing, frequency, who administers)? What is the process for collecting assessment results?
2. What kinds of obstacles, if any, have you encountered in preparing for and administering student assessments or submitting student assessment results? How many extra hours, if any, have you had to devote to these tasks?
3. Have you made any changes to your approach to teaching or to the content of your lessons since the implementation of this alternative assessment or SLO?
4. Have there been any changes to the district's approach to using alternative student growth measures? Have these changes resulted in notable improvements?

D. Use of alternative growth or value-added measure

1. For what purposes is the district using each alternative student outcome measure used in assessing teacher performance? Is it used for evaluating you? Is it used as part of a performance-based compensation system? Do teachers receive targeted professional development based on their performance? Does the measure affect teacher assignments? Are there any other ways in which the district is using these measures?

2. Do you receive any feedback regarding your performance on each measure? When and how is this feedback provided? Who provides this feedback to teachers? Is feedback provided in person? In writing? Is there a way for you to respond to the feedback?
3. Does your performance rating on the alternative measure affect whether you are awarded tenure?

E. Benefits and drawbacks

1. Could you summarize some of the main benefits, from your perspective, of implementing this alternative VAM or growth measure of teacher performance in the district? For example, have you found that receiving this kind of information about your performance has had any benefits for your effectiveness in the classroom? How has receiving this kind of performance information affected your teaching, including how you feel about teaching and the way you set goals and relate to students?
2. What do you perceive as the primary drawbacks or costs of implementing the measure? Do you think the measure allows for a sufficient level of comparability across teachers on performance? Do you think the measure accurately reflects your contribution to student learning? If SLOs are used in the district, do you think all teachers are setting sufficiently high standards for student growth?
3. How do you perceive the district's communication of the use of the alternative measure in evaluations to teachers? Do you think teachers in the district understand the measure and how they can change its level? Do you think teachers in the district have a positive or negative view of the use of alternative student outcomes to measure their performance?

F. District context

1. Have there been any changes in the district administration or policies that may have affected the implementation of these alternative measures of teacher performance (for example, change in superintendent or implementation of a performance-based pay system)?
2. Did the district have the support of the teachers' union/association for the implementation of the VAM or growth measure of teacher performance using an alternative student outcome? Why or why not?
3. Is there anything else that you think might be important to know about the development and implementation of this alternative growth or value-added measure in the district? Are any other key stakeholders in the district with whom you think it would be important for us to speak?

Appendix H. Teachers' union representative interview protocol

A. Introduction and consent for audio recording

The purpose of this section is to introduce the respondent to the study and request consent for audio recording.

1. Thank you for taking the time for the interview today. My name is [name] and I work for Mathematica Policy Research, which is conducting this study for the Mid-Atlantic Regional Educational Laboratory (REL). The Mid-Atlantic REL's mission is to "build research alliances with education practitioners and policymakers to help states and districts use data to answer important issues of policy and practice." The results of this study will inform publicly available reports aimed at states and districts considering implementing measure of teacher performance that include student growth. As I mentioned [in my email/when we spoke on the phone], we are conducting case studies of districts implementing measures of teacher performance that include student growth, but that measure growth on outcomes other than state assessments. These alternative outcome measures might include nationally normed assessments, end-of-course curriculum-based assessments, or "student learning objectives" (SLOs) developed by individual teachers for their own students. Student growth might be measured through a simple subtraction of end-of-year achievement from beginning-of-year achievement, through a student growth percentile method, through a value-added model, or implicitly in SLOs that are set based on an understanding of baseline achievement levels in a class at the beginning of the school year. We plan to examine what alternative outcome measures are used, how the growth measures are implemented, challenges and obstacles in implementation, how the measures are being used, and the effectiveness of these measures in differentiating teacher performance. We expect our findings to benefit states and districts considering the use of such alternative growth measures. The purpose of this interview is to gather information on the implementation of [insert specific growth measures used in the district] in the district from your perspective as a teachers' union representative. We will be interviewing teachers, principals, and union representatives in the district to gather information on their perspectives as well. We will not identify you by name in the study reports. We also will not identify the districts included in the study in subsequent reports.

Before beginning, I would like to get your permission to record the interview. This will ensure that we accurately capture information about district policies and will prevent me from having to ask questions and take notes at the same time. The audio recordings from all individuals interviewed within your district will be compiled into a single district summary that is for internal analysis purposes only and will not be published or shared with anyone outside the study team. You will not be identified in the district summary. This summary will then be aggregated with other district summaries and analyzed for trends and variations across multiple districts. Do I have your permission to record the interview?

B. Type of student outcome and growth measure used

1. Are teachers in this district measured based on [identify the alternate outcome growth measure(s) discussed in district interview]?

2. What role did the teachers' union/association play in the design of the alternative measure(s)? Can you describe the process for rolling out the alternative measure(s) in the district (including SLOs)? What role did the teachers' union/association play in the implementation? Did the union raise any concerns regarding the alternate measure(s)? If so, how did the district respond to and address those concerns? Were teachers provided with any training or professional development related to the implementation of the measure?
3. Does the district use a VA model? If not, how does the district measure growth (for example, percentile method, simple subtraction)? How long has the growth/VA model been in place? Has the district made any changes to this approach to measuring growth during that period?

C. Implementation of data collection and analysis

1. I would like to learn more about the administration of these alternative student assessments in the district. (For each assessment) Are you familiar with how the assessment is administered (for example, timing, frequency, who administers)? What is the process for collecting assessment results?
2. Are you aware of any quality control measures in place in the district for the collection and analysis of the teacher performance data? (If district uses SLOs) For example, what documentation of student performance do teachers need to submit for SLOs? Is this documentation monitored regularly? Is this documentation subject to any audits? (If district uses a VAM/growth model) For example, do teachers administer or score the student assessments, or is administration and/or scoring conducted by independent proctors or an external firm?
3. What kinds of obstacles, if any, have teachers encountered in preparing for and administering the student assessment or submitting student assessment results? How much extra time, if any, have teachers had to devote to these tasks? What is your perspective on how to address these obstacles? Have you found the district to be responsive to these concerns?
4. Have there been any changes to the district's approach to using alternative student growth measures? Have these changes resulted in notable improvements?

D. Use of alternative growth or value-added measure

1. For what purposes is the district using each alternative student outcome measure used in assessing teacher performance? Is it used for evaluating teachers? In what grades and subject areas? Is it used as part of a performance-based compensation system? In what grades and subject areas? Do teachers receive targeted professional development based on their performance? Does the measure affect teacher assignments? Are there any other ways in which the district is using these measures?
2. What role did the teacher's union/association play in determining how the alternative student outcome measure would be used? Is the measure used in just cause terminations or tenure denial? If so, does the teacher's union/association support the measure's use in these decisions?

3. Do teachers receive any feedback regarding their performance on each measure? What type of feedback? When and how is this feedback provided? Who provides this feedback to teachers?
4. If used in teacher evaluation, is the measure combined with any other measures of teacher performance (for example, value-added or other growth measures based on state assessments or classroom observation measures)? How much weight does each performance measure receive in a teacher's overall evaluation rating? Does the use of the measure in teacher evaluation vary by grade and/or subject area?
5. Does a teacher's performance rating on the alternative measure affect whether he/she is awarded tenure?

E. Benefits and drawbacks

1. Could you summarize some of the main benefits, from your perspective, of implementing this alternative VAM or growth measure of teacher performance in the district? For example, has the district been able to measure the performance of a larger percentage of teachers in the district? How have teachers in the district responded to the use of alternative student outcomes to measure their performance? Do you perceive this alternative measure as an effective way to improve teacher quality in the district?
2. What do you perceive as the primary drawbacks or costs of implementing the measure? Do you think the measure allows for a sufficient level of comparability across teachers on performance? If SLOs are used in the district, do you think all teachers are setting sufficiently high standards for student growth?
3. How do you perceive the district's communication of the use of the alternative measure in evaluations to teachers?

F. District context

1. Have there been any changes in the district administration or policies that may have affected the implementation of these alternative measures of teacher performance (for example, change in superintendent or implementation of a performance-based pay system)?
2. Did the district have the support of the teachers' union/association for the implementation of the VAM or growth measure of teacher performance using an alternative student outcome? Why or why not?
3. Has the district received funding or resources from external sources (for example, federal government, foundations) to develop or implement this measure?
4. Is there anything else that you think might be important to know about the development and implementation of this alternative growth or value-added measure in the district? Are any other key stakeholders in the district with whom you think it would be important for us to speak?

References

- Gill, B., Bruch, J., & Booker, K. (2013). *Using alternative student growth measures for evaluating teacher performance: What the literature says* (REL 2013–002). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. <http://eric.ed.gov/?id=ED544205>
- Gill, B., English, B., Furgeson, J., & McCullough, M. (2014). *Alternative student growth measures for teacher evaluation: Profiles of early-adopting districts* (REL 2014–016). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. <http://eric.ed.gov/?id=ED544797>
- Merriam, S. B. (2001). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.
- National Council on Teacher Quality. (n.d.). NCTQ teacher contract database. Retrieved January 14, 2013, from <http://www.nctq.org/districtPolicy/contractDatabaseLanding.do>.
- Whitehurst, G. J. (2013, April 24). Teacher value-added: Do we want a ten percent solution? *The Brown Center Chalkboard blog*. Washington, DC: Brookings Institution. Retrieved June 3, 2013, from www.brookings.edu/blogs/brown-center-chalkboard/posts/2013/04/24-merit-pay-whitehurst.
- Wiersma, W. (2000). *Research methods in education: An introduction*. Boston, MA: Allyn and Bacon.

The Regional Educational Laboratory Program produces 7 types of reports



Making Connections

Studies of correlational relationships



Making an Impact

Studies of cause and effect



What's Happening

Descriptions of policies, programs, implementation status, or data trends



What's Known

Summaries of previous research



Stated Briefly

Summaries of research findings for specific audiences



Applied Research Methods

Research methods for educational settings



Tools

Help for planning, gathering, analyzing, or reporting data or research