

OMB Supporting Statement Part A: Alternative Student Growth Measures for Teacher Evaluation: Case Studies of Early Adopters

August 6, 2013

Table of Contents

Table of Contents	ii
SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT	1
PART A. JUSTIFICATION	1
1. Circumstances Necessitating the Collection of Information	1
a. Statement of Need to Study Alternative Measures of Student Growth.....	1
b. Research Questions.....	3
c. Study Overview.....	3
d. Recruitment of Districts.....	4
e. Data Collection Plan.....	5
2. Purposes and Uses of Data	9
3. Use of Technology to Reduce Burden	9
4. Efforts to Avoid Duplication	9
5. Methods to Minimize Burden on Small Entities	10
6. Consequences of Not Collecting Data	10
7. Special Circumstances	10
8. Federal Register Announcement and Consultation	11
a. Federal Register Announcement.....	11
b. Consultations Outside the Agency.....	11
9. Payments or Gifts	11
10. Assurances of Confidentiality	11
11. Justification for Sensitive Questions	13
12. Estimates of Burden Hours	13
13. Estimates of Cost Burden to Respondents	14
14. Annualized Cost to the Federal Government	14
15. Reasons for Program Changes or Adjustments	14
16. Plans for Tabulation and Publication of Results	14
a. Tabulation Plans.....	15
b. Publication Plans.....	17

17. Approval Not to Display the Expiration Date for OMB Approval	17
18. Exception to the Certification Statement	17
REFERENCES	18

SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT

This submission is a request for approval of data collection activities that will be used to support the Mid-Atlantic Regional Educational Laboratory (REL) Alternative Student Outcomes for Growth Measures Case Studies. The study is being funded by the Institute of Education Sciences (IES), U.S. Department of Education (ED), and is being implemented by ICF International and its subcontractor, Mathematica Policy Research.

This study aims to fill the gap in information available to districts and policymakers on measures of student growth that do not use state standardized tests via qualitative case studies of up to nine districts that are using alternative measures of student achievement growth in teacher performance ratings. The case studies will address what alternative outcome measures are used, how the alternative growth measures are implemented, challenges and obstacles in implementation, how the measures are being used. Where possible, the study team will examine the extent of differentiation produced by the measures—specifically, the distribution of teacher performance on the measures, as compared with the distribution of teacher performance on conventional value-added measures that are based on state assessments. The study team will conduct semi-structured interviews with district administrators leading teacher evaluation or effectiveness efforts, teacher representatives (such as union leaders), teachers (including both classroom teachers and instructional coaches), and principals. The data collected will be summarized and analyzed using a case study approach.

This submission requests approval to recruit districts for the study and conduct in-person and telephone interviews with staff in participating districts.

PART A. JUSTIFICATION

1. Circumstances Necessitating the Collection of Information

a. Statement of Need to Study Alternative Measures of Student Growth

The specific legislation authorizing this data collection is specified in Part D, Section 174 (20 U.S.C. 9564) of the Education Sciences and Reform Act (ESRA) of 2002. Part D, Section 174 provides for ED to enter into five-year contracts with entities to establish a networked system of ten regional laboratories that serve the needs of each region of the United States. The Regional Educational Laboratories (RELs) are to

carry out a range of activities to serve the needs of each region in the United States, including applied research, and development, dissemination, training, and technical assistance activities that focus on how to use data and analysis. The primary mission of the RELs is to help states and districts systematically use data and analysis to answer important issues of policy and practice with the goal of improving student outcomes.

In school districts that have sought to measure the effectiveness of teachers in raising student achievement, statistical methods that aim to measure growth or value-added are typically applied to students' scores on statewide standardized tests. These methods are broadly referred to as "growth models." Growth models have attracted considerable interest among educators and policymakers nationwide—and have been promoted by the US Department of Education via Race to the Top (RTT) and the Elementary and Secondary Education Act (ESEA) waiver process. Indeed, many states are beginning to require that some measure of student achievement growth be used as a component of the evaluation of teachers and/or principals. All five states in the REL Mid-Atlantic region have secured Race to the Top funding, which places a premium on teacher evaluation. Reforms to the teacher evaluation systems incorporate student performance, whether it be through a value-added model (e.g., DCPS's IMPACT system; Pennsylvania Value Added Assessment System – PVAAS) or a student growth model (e.g., DEDOE, MSDE, and NJDOE). But the utility of value-added models has been limited by the breadth and depth of the underlying student assessments. The reliance on statewide standardized tests for value-added models limits the grades and subjects that can be evaluated (often only grades 4-8, only in math and reading).

To permit evaluation of grades and subjects not tested through state exams, and to derive a more comprehensive picture of teacher effectiveness through multiple growth measures, some school districts have begun to use alternative student outcome measures in value-added models and other growth models. Alternative student achievement measures in use in growth models range from end-of-course curriculum-based assessments to commercially available standardized tests such as the Iowa Test of Basic Skills and the PSAT. Meanwhile, multiple districts have adopted student learning objectives as alternative measures of student achievement growth that are used in teacher evaluation but do not involve growth models. Student learning objectives differ from growth models in that they are specific growth targets for a teacher's own particular set of students, and the targets are typically set by individual teachers and approved by principals at the beginning of each school year.

This study, which was requested by the Pennsylvania Department of Education (a member of the Teacher Evaluation Research Alliance), will be of interest to officials and policymakers throughout the REL Mid-Atlantic region (and the country) who would like to learn more about the characteristics, uses, and implementation of alternative student growth measures. As noted above, large numbers of states and districts are exploring such measures, often as a result of changes in state-level evaluation systems inspired by Race to the Top grants or ESEA waivers. Within the region, Delaware is planning on extending evaluation of teachers in non-tested grades and subjects to include alternative measures and growth goals in the 2012-2013 school year. Pennsylvania, Maryland, and New Jersey are likewise working on changes to their teacher evaluation system with the aim of including some measure of student growth in the evaluation of teachers within and outside of tested grades and subjects. The Pittsburgh Public Schools have pioneered the development and use of value-added models that rely on

locally developed curriculum-based assessments for courses in secondary grades. The District of Columbia Public Schools have already started using student learning objectives (called Teacher-Assessed Student Achievement Data) and the district planning to add end-of-course exams to include high school teachers in their value-added models in subsequent years (District of Columbia Public Schools 2011).

Most of these measures are so new that the districts and states developing and using them have had to work thus far without much (if any) knowledge about challenges in the implementation of growth measures using alternate assessments. The findings of this study therefore have the potential to inform states and districts in deciding which measures are promising for use in evaluation systems, which are of doubtful value (for example, because they do not distinguish many teachers from each other, or because they show indications of rapid inflation from year to year), and the difficulty of effectively implementing the measures. Teacher evaluation systems will be works in progress for at least the next several years, and policymakers in the Mid-Atlantic region and beyond need information to inform their selection of measures, how the measures will be used, and how to minimize implementation costs and avoid obstacles.

b. Research Questions

The primary research questions to be addressed in the study are:

- What student outcome measures other than state standardized test scores are currently being used in growth measures to assess teacher performance?
- How have school districts implemented the data collection and analysis necessary for growth measures based on alternative student outcomes, and what obstacles have they encountered?
- How are the alternative measures being used for other purposes in addition to teacher evaluation? If they are used for other purposes (e.g., for planning instruction), are those purposes compatible with maintaining their validity and reliability for evaluation?
- How much weight does each alternative measure alternative receive in a teacher's overall evaluation, and how does the weight vary by grade and subject? How does the distribution of scores on the alternate measure compare to the distribution of scores on growth measures used with state assessments (e.g., value-added measures) or other measures of performance?
- What are the perceived benefits and drawbacks of using growth models (including VAM) based on each type of alternative outcome: student learning objectives, end-of-course curriculum-based assessments, and nationally normed assessments? What costs (notably in terms of time and effort) do they impose on teachers, principals, and districts? What are the estimated monetary costs associated with each of the following components: piloting, training, additional data collection, data system development, data analysis, and reporting

c. Study Overview

In this study, the study team will examine what alternative outcome measures are used, how the alternative growth measures are implemented, challenges and obstacles in implementation, how the

measures are being used, and, where possible, the distribution of teacher performance on the measures, as compared with the distribution of teacher performance on conventional value-added measures that are based on state assessments. The study will examine the use of three categories of alternative student growth measures: (1) end-of-course curriculum-based assessments to which statistical growth models are applied; (2) nationally-normed assessments such as the PSAT or ITBS, to which statistical growth models are applied; and (3) student learning objectives, which do not involve the application of statistical growth models, and instead are intended to account for growth implicitly, because they are selected separately for each teacher's students. Eight districts have been recruited to participate in the study; all three categories of alternative growth models are included among the eight districts.

Data will be collected using semi-structured interviews with district administrators leading teacher evaluation or effectiveness efforts, teacher representatives (such as union leaders), teachers (including both classroom teachers and instructional coaches), and principals. Prior to the interviews, the study team will review documents, available on each of the districts' websites, which address teacher evaluation and effectiveness efforts. The four types of respondents have been selected as the stakeholders who are most involved in the decision-making process and most impacted by implementation. District administrators and union representatives who played significant roles in developing the alternative growth measures will be best equipped to contribute information on the district's decisions related to choosing and implementing the specific measure, and to describe any outcomes or lessons learned during the process that can be applied by other districts as best practices. Teachers and principals—the respondents most directly affected by the implementation of these alternative growth measures—will provide perspectives on the costs and benefits of implementation that other districts can anticipate.

The data collected through interviews and document reviews will be summarized and analyzed using a qualitative approach. First, a single write-up synthesizing information collected in the interviews will be created for each district. Then, a coding system will be developed to characterize the variation across key variables or categories of information (e.g., how teacher performance ratings are used). After the data have been coded, the study team will use four main categories to examine and summarize the data: the type of alternative student outcome and growth model, the applications of the growth measure, the differentiation in teacher performance produced by the growth measure, and the implementation process, including perceived costs and benefits of the measures. The data will be grouped by type of alternative growth measure rather than by district, and the study team will focus on examining consistencies and variation both within and across the alternative growth measure categories.

d. Recruitment of Districts

The study team has recruited eight school districts to participate in the study. The study team began the recruitment effort by mailing districts an introductory package, which will include the following two documents:

- **Notification letter.** The one-page notification letter describes the importance of studying the implementation of alternative measures of student growth, provides an overview of the study design, summarizes the benefits of participating, and notes that a study team member will follow up by telephone to discuss the study in more detail(See Appendix A to Statement B).
- **Study summary.** The two-page summary describes the purpose of the study and the benefits of participation, identifies the study team, and provides contact information for the project director and the ED project officer. It also discusses the activities required of participating districts and schools (See Appendix B to Statement B).

Study team members followed up with phone calls involving one or more relevant district staff, describing the study in more detail, answering questions, and securing participation.

e. Data Collection Plan

The study data collection consists of on-site and in-person interviews with school and district staff who are knowledgeable about district implementation of alternative measures of student growth. During December 2013 – February 2014, the study team plans to conduct interviews with up to ten individual respondents in each district—specifically, at least one district administrator, two principals, three classroom teachers, and one union/association representative. In larger districts, the study team will conduct additional interviews. In districts with dedicated instructional leaders (e.g., instructional coaches and master teachers), one or two additional interviews will be conducted with instructional leaders to supplement interviews with classroom teachers and gain a broader perspective on the effects on teachers of implementing the measures. The number of instructional leaders interviewed will depend on the size of the district.

The selection of principal and teacher respondents will begin with a purposeful selection of the most credible and knowledgeable respondents. Because response bias is a concern and because the study team will not be able to select a representative sample of all subjects and grades in a district due to the small sample size of respondents, the team’s focus is on selecting teacher and principal respondents who will be most credible or accurate in their reports on implementation, application, and effectiveness of the alternative measures of student growth. To ensure a range of viewpoints, the study team will also solicit suggestions of teachers who actively struggled with implementation.

Table 1 describes the timeline for data collection.

Table 1. Data Collection Timeline

	2013												2014	
	1	2	3	4	5	6	7	8	9	10	11	12	1	2
Site selection and pilot interviews		●	●	●										

	2013												2014	
	1	2	3	4	5	6	7	8	9	10	11	12	1	2
Submission of initial OMB clearance package (60-day public comment period)								●	●					
Analysis of pilot interviews				●	●									
Submission of final OMB clearance package										●	●			
Data collection										●		●	●	●

The focus of each interview will vary by respondent type as shown in Table 2. Interviews with district administrators will primarily cover the type of student outcome and growth measure used, the implementation process, and the extent of differentiation in teacher performance produced by the growth measure. Interviews with principals, classroom teachers, and instructional leaders will focus on the applications of the growth measure, the perceived costs and benefits of the measure from the viewpoint of the respondent, and the respondent’s perception of the success of the district’s communication during implementation of the measure. Interviews with teachers’ union/association representatives will cover the development and implementation of the alternative measure—particularly related to the union’s involvement—and the respondent’s perception of the success of the district’s communication approach during the process.

Interview protocols have been devised for each type of interviewee: district administrators, principals, teachers and instructional leaders, and union representatives. Those protocols are available in Appendix D of Part B.

Table 2. Interview Topics by Respondent Type

	District Administrator	Principal	Teacher/ Instructional Leader	Union Representative
Type of student outcome and growth measure used	√	√	√	√
Design of growth measure	√			√
Piloting process	√			√
Training/PD during roll-out	√	√	√	√
Involvement of teacher or union in decision-making	√		√	√
Applications of growth measure	√	√	√	√
Feedback to teachers	√	√	√	√
Staffing decisions	√	√	√	√
Union role				
Differentiation produced by growth measure	√	√		√
Relative distribution	√	√		√
Reliability and validity	√	√		√
Implementation process	√	√	√	√
Assessment administration/ data collection	√	√	√	√
Data analysis	√			
Quality control processes	√	√		√
Perceived costs and benefits of the measure	√	√	√	√
Perceived success of the district's communication and teacher response	√	√	√	√
Changes in the district's approach	√	√	√	√
External funding	√			
Related district changes	√	√	√	√

The study team also plans to collect contextual information about the districts to describe the types of districts implementing alternative growth measures in each of the three categories. Specifically, the study team will collect information on district size, region, urbanicity, proportion of disadvantaged students, collective bargaining status, adoption of Common Core standards, and receipt of funding through RTT, School Improvement Grants (SIG), or the Teacher Incentive Fund (TIF) This information will be drawn primarily from the Common Core of Data; a full list of the specific variables and extant data sources the study team plans to use is included in Table 3.

Table 3. District Extant Data Collection Protocol

Variable	Type of Data	Categories	Data Source
Size	Total student enrollment	<ul style="list-style-type: none"> • Small (<10,000 students) • Medium (10,000-45,000 students) • Large (>45,000 students) 	Common Core of Data (CCD)
Region	Geographic categories	<ul style="list-style-type: none"> • North • South • West • Mid-West 	CCD
Urbanicity	NCES urban-centric locale code	<ul style="list-style-type: none"> • City • Suburb • Town • Rural 	CCD
Poverty	Proportion of students eligible for FRL	<ul style="list-style-type: none"> • High poverty (at least 65% of students eligible for FRL) • Low poverty (<65% of students eligible for FRL) 	CCD
Collective bargaining status	State collective bargaining requirement for school employees (binary)	<ul style="list-style-type: none"> • Collective bargaining required • Collective bargaining not required 	Department of Labor
Common Core State Standard adoption	Common Core State Standard adoption status (binary)	<ul style="list-style-type: none"> • In Common Core state • Not in Common Core state 	http://www.corestandards.org/in-the-states
RTT participation by year of alternative growth measure implementation	RTT participation status (binary)	<ul style="list-style-type: none"> • In RTT state • Not in RTT state 	US Department of Education (ED)
SIG funding by year of alternative growth measure implementation	SIG funding status (binary)	<ul style="list-style-type: none"> • Received SIG funding • Did not receive SIG funding 	State departments of education SIG records; district interviews
TIF funding by year of alternative growth measure implementation	TIF funding status (binary)	<ul style="list-style-type: none"> • Received TIF funding • Did not receive TIF funding 	Center for Educator Compensation Reform

2. Purposes and Uses of Data

Mathematica will collect and analyze the data for the REL Mid-Atlantic Regional Educational Laboratory (REL) Alternative Student Outcomes for Growth Measures Case Studies under contract number ED-IES-12-C-0006 with the Institute of Education Sciences, Department of Education.

The primary purpose of the study is to document the implementation of alternative measures of student growth. The study team will examine what alternative outcome measures are used, how the alternative growth measures are implemented, challenges and obstacles in implementation, how the measures are being used, and, where possible, the distribution of teacher performance on the measures, as compared with the distribution of teacher performance on conventional value-added measures that are based on state assessments. The findings of this study will fill a gap in information on the implementation of such alternative growth measures and inform states and districts in deciding which measures are promising for use in evaluation systems, which are of doubtful value, and the difficulty of effectively implementing the measures.

The data to be collected will be obtained from district and school staff interviews. After the data have been coded, the study team will use four main categories to examine and summarize the data: the type of alternative student outcome and growth model, the applications of the growth measure, the differentiation in teacher performance produced by the growth measure, and the implementation process, including perceived costs and benefits of the measures. The data will be grouped by type of alternative growth measure rather than by district, and the study team will focus on examining consistencies and variation both within and across the alternative growth measure categories.

3. Use of Technology to Reduce Burden

Data from the semi-structured interviews cannot be collected through such methods as web surveys or computer-assisted telephone interviews. The proposed interviews will be necessary to allow in-depth, conversational exchanges with respondents, and to obtain answers to both open-ended and detailed questions. Prior to conducting the interviews, the study team will examine district websites to collect preliminary details about (a) the implementation of alternative measures of student growth in the district and (b) source documents to corroborate information provided during interviews.

4. Efforts to Avoid Duplication

No comprehensive multi-district study has been conducted or is underway to address the research questions presented in this study. Few studies have evaluated the reliability and validity of value-added estimates that are based on alternative measures of student achievement, as demonstrated in Gill, Bruch, and Booker (2012). Johnson et al. (2012) is a notable exception. This study reports summary results for schoolwide value-added models in Pittsburgh that are based on end-of-course exams, PSATs, attendance, and the pass rate for core classes. The study finds that all of these measures can be used to statistically distinguish above-average and below-average school performance. Papay (2011) finds

correlations ranging from 0.15 to 0.58 across different model specifications between teacher value-added estimates based on the two alternative assessments and the state assessment.

There is also limited evidence on the reliability and validity of student learning objectives and their ability to distinguish among teachers. A recent study analyzed four years of data on student learning objectives in the context of a pilot program in Denver that tied additional teacher compensation to meeting goals set by student learning objectives (Community Training and Assistance Center, 2004). Results indicate a positive relationship between meeting student learning objectives and higher student achievement, providing some evidence of their predictive validity. The study also found that teachers improved the “quality” of their objectives over time, as they received more training and support on setting the targets. Ultimately, the pilot schools showed little variability in terms of teachers meeting and not meeting their student learning objectives. Goldhaber and Walch (2011) study the effect on student achievement of the compensation system that grew out of the pilot program analyzed in Community Training and Assistance Center (2004). They found that teacher effectiveness measures based on student learning objectives were highly correlated with teacher value-added based on standardized tests in math and reading.

To date, the type of implementation analysis proposed for this study has not been conducted across multiple districts and types of alternative student growth measures, and there is no alternative source for the information to be collected.

5. Methods to Minimize Burden on Small Entities

The study team has developed an efficient interview protocol that focuses on the data of most interest to districts and policymakers. Some of respondents, such as schools and small districts, will be small entities meaning they have a population of less than 50,000. The study team will also offer to speak with respondents via phone interviews rather than in-person visits when preferred by the district.

6. Consequences of Not Collecting Data

The data collection activities described in this submission are necessary for ED to document the implementation of alternative measures of student growth and assess their reliability and validity for evaluating teacher performance. This is a one-time data collection effort. The study represents a significant step in examining how to evaluate teachers using student growth measures in non-state-tested grades and subjects. Without the data collected in this study, ED will not be able to provide sufficient information to assist states and districts in determining how to implement a teacher evaluation system that incorporates alternative measures of student growth.

7. Special Circumstances

There are no special circumstances associated with this data collection.

8. Federal Register Announcement and Consultation

a. Federal Register Announcement

A 60-day notice to solicit public comments will be published in the *Federal Register* in early August. We will analyze the responses in early October and submit the final package for approval within two weeks of the receipt of public comments.

b. Consultations Outside the Agency

The study team has worked with IES to identify five members to serve as the Technical Working Group (TWG) for this study. These experts on teacher evaluation and student assessment have provided input on the study's design. For example, TWG members advised the study team to consider respondent bias, which led us to seek out some respondents who are unlikely to be advocates for the program. They also advised us to consider the relationship between the evaluation uses and instructional uses of student learning objectives, which the study team will do. The TWG includes the following individuals:

- Robert Boruch, University of Pennsylvania, 215-898-0409
- Laura Hamilton, RAND, 412-683-2300, x4403
- Christopher Hulleman, University of Virginia, 434-924-6998
- Andrew Porter, University of Pennsylvania, 215-898-7014
- Christopher Rhoads, University of Connecticut, 860-486-3321

9. Payments or Gifts

The study team will not give payments or gifts to districts for completing the interview or providing other study data. The study team does not believe it will be necessary to provide incentives for participation.

10. Assurances of Confidentiality

REL Mid-Atlantic (specifically, Mathematica Policy Research, a subcontractor to ICF International) follows the confidentiality and data protection requirements of IES (The Education Sciences Reform Act of 2002, Title I, Part E, Section 183). Mathematica will protect the confidentiality of all information collected for the study and will use it for research purposes only and will obtain security clearance through NCEE's security clearance office. No information that identifies any study participant will be released. All data will be kept in secured locations and identifiers will be destroyed as soon as they are no longer required. All members of the study team having access to the data will be trained and certified on the importance of data security. When reporting the results, the study team will present data in aggregate form only so that individuals and institutions will not be identified. The study team will also include the following statement in the requests to districts for participation in the study (Appendix A):

- Responses to the data collection activities will be used for research purposes only. The reports prepared for the study will summarize findings across the sample and will not associate responses with a specific district, school, or individual. The study team will not provide information that identifies you or your district to anyone outside the study team, except as required by law.
- The contractor follows the confidentiality and data protection requirements of the Institute of Education Sciences (The Education Sciences Reform Act of 2002, Title I, Part E, Section 183). The contractor will protect all information collected for the study and will use it for research purposes only. No information that identifies any study participant will be released. Information on respondents will be linked to their institution but not to any individually identifiable information. No individually identifiable information will be maintained by the study team. All institution-level identifiable information will be kept in secured locations and identifiers will be destroyed as soon as they are no longer required.

The following safeguards, which are routinely employed by Mathematica to carry out confidentiality assurances, will be applied consistently during the study:

- All employees sign a confidentiality pledge (Appendix B), which describes both the importance of and the employee's obligation to discretion.
- Personally identifiable information is maintained on separate forms and files, which are linked by sample identification number only.
- Access to hard copy documents is strictly limited. Documents are stored in locked files and cabinets, and discarded materials are shredded.
- The study team will submit a list of all people who have access to respondents and data to NCEE and will track all staff joining or leaving the study to ensure that signatures are obtained and clearances revoked.
- Access to computer data files is protected by secure user names and passwords, which are available to specific users only.
- Especially sensitive data is encrypted and stored on removable storage devices that are kept physically secure when not in use.

The plan for maintaining confidentiality includes ensuring that all personnel with access to individual identifiers sign confidentiality agreements and provide notarized nondisclosure affidavits. Also included in the plan is personnel training regarding (1) the meaning of confidentiality, particularly as it relates to handling requests for information and providing assurance to respondents about the protection of their responses; (2) controlled and protected access to computer files under the control of a single database manager; (3) built-in safeguards concerning status monitoring and receipt control systems; and (4) a secured and operator-controlled, in-house computing facility.

11. Justification for Sensitive Questions

There are no questions of a sensitive nature in the district interviews.

12. Estimates of Burden Hours

Table 3 shows the estimated burden hours for district staff who will participate in data collection. These estimates are based on the study team's experience collecting such data from district staff for similar studies. Please also refer to the IC burden table included as an attachment to this supporting statement.

Table 3. Estimated Response Time for Data Collection

Respondent/Data Request	Number of Targeted Respondents	Expected Response Rate (%)	Number of Respondents	Unit Response Time (Hours)	Total Response Time (Hours)	Total Cost (Based on hourly wage rate)
<i>District staff</i>						
Recruitment by phone	20	100	20	2	40	\$1944
Phone interview (up to 2 staff at each district)	18	100	18	1	18	\$875
<i>Principals</i>						
Recruitment by phone	27	100	27	1	27	\$1181
Phone interview (up to 3 principals in each district)	27	100	27	1	27	\$1181
<i>Teachers</i>						
Recruitment by phone	36	100	36	1	36	\$980
Phone interview (up to 4 teachers in each district)	36	100	36	1	36	\$980
<i>Instructional Leads/Coaches</i>						
Recruitment by phone	12	100	12	1	12	\$245
Phone interview (up to 4 teachers in each district)	12	100	12	1	12	\$245
<i>Teachers'</i>						
<i>Union/Association Representative</i>						
Recruitment by phone	9	100	9	1	9	\$245
Phone interview (up to 1 representative in each district)	9	100	9	1	9	\$245
Total	104		104		226	\$8,121

Note: Total cost estimates based on hourly wage rates for the Bureau of Labor Statistics' May 2012 National Industry-Specific Occupational Employment and Wage Estimates, NAICS 61100: Educational Services (http://www.bls.gov/oes/current/naics3_611000.htm). District staff burden estimates based on a mean hourly wage of \$48.59 for Human Resources Managers in Elementary and Secondary Schools (11-3121). Principal burden estimates based on a mean annual salary of \$90,980 and 2,080 annual hours for Educational Administrators in Elementary and Secondary Schools (11-9032). Teacher, instructional leads, and union/association representative burden estimates based on a mean annual

salary of \$56,620 and 2,080 annual hours for Preschool, Primary, Secondary, and Special Education Teachers (25-2000). Total costs rounded to the nearest dollar.

13. Estimates of Cost Burden to Respondents

There are no start-up costs or ongoing operation and maintenance costs for respondents.

14. Annualized Cost to the Federal Government

The estimated cost of the study to the federal government is \$345,339 over two years, for an average annual cost of \$172,670. The estimated total cost of the REL Mid-Atlantic Region five-year contract is \$32,353,087, of which this study accounts for \$345,339. Study costs were estimated using fully-burdened (“loaded”) hourly rates for staff, multiplied by the estimated number of hours of labor, and adding travel and other direct costs. Multipliers for indirect rates were applied to travel and other direct costs.

15. Reasons for Program Changes or Adjustments

This is a new data collection.

16. Plans for Tabulation and Publication of Results

Table 4. Timeline for Data Collection and Analysis

	2013										2014									
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8
Site selection and pilot interviews		●	●	●																
Submission of initial OMB clearance package (60-day public comment period)								●	●											
Analysis of pilot interviews				●	●															

	2013										2014									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Submission of final OMB clearance package											●	●								
Pilot report draft 1 (internal and IES review)						●	●													
Pilot report draft 2 (internal and IES review)							●	●												
Delivery of pilot results to requestors and report public release									●											
Data collection												●	●	●						
Data analysis													●	●	●					
Draft 1 of final report (internal and IES review)															●	●				
Draft 2 of final report (internal and IES review)																	●	●		
Delivery of results to requestors and final report public release																				●

a. Tabulation Plans

The study team will use the information gathered from the district and school staff interviews to examine and describe the implementation of alternative measures of student growth in three main categories: (1) end-of-course curriculum-based assessments to which statistical growth models are applied; (2) nationally-normed assessments such as the PSAT or ITBS, to which statistical growth models are applied; and (3) student learning objectives, which do not involve the application of statistical growth models, and instead are intended to account for growth implicitly, because they are selected separately for each teacher’s students.

To prepare the interview data for the analysis, the study team will develop a coding system organized into four broad categories of information: the type of alternative student outcome and growth model, the applications of the growth measure, the differentiation in teacher performance produced by the

growth measure, and the implementation process, including perceived costs and benefits of the measures. The study team will define codes for the variations that they expect to observe based on their initial literature review and refine the coding system as they collect data. For example, as part of their coding of applications of alternative growth measures, they will create categories to capture how teacher performance ratings based on these measures are used: for evaluating teachers, for targeting professional development, for performance-based compensation, and for assigning teachers to positions in the district.

The study team will use the four main categories in the coding system—the type of alternative student outcome and growth model, the applications of the growth measure, the effectiveness of the growth measure in differentiating teacher performance, and the implementation process (including the perceived costs and benefits of the measures)—to summarize key findings across all districts. The district will be the primary unit of analysis. The study team plans to address their research questions by summarizing and reporting on the following findings by category:

- **Type of alternative student outcome and growth model:** frequency of each outcome and growth model; range and mean years of implementation; range and frequency of subjects and grades affected; frequency of application to teachers of ELL and special education students
- **Applications of growth measure:** frequency of specific uses for alternative measures (evaluation, compensation, targeting of professional development, teacher assignment, tenure awards, etc.); range and mean frequency of feedback to teachers on alternative growth measure performance ratings
- **Differentiation of growth measure:** characteristics of reliability and validity analysis; range of reported reliability and validity for each measure; range and mean number of performance categories; range and mean percentage of teachers in highest and lowest performance categories; range of weights given to the alternative measures
- **Implementation process (including perceived costs and benefits):** mean length of piloting period; range and mean frequency of data collection; frequency of external vendor contracting; most frequently cited quality control measures; most cited obstacles to data collection and analysis; most frequently cited benefits (by respondent type); most frequently cited costs (by respondent type); range and mean monetary cost (by type of cost cited)

To facilitate cross-site analysis, the data will be grouped by type of alternative growth measure rather than by district. Across all districts, the study team will examine findings by key factors of interest to stakeholders: type of alternative growth measure and years of experience in implementation. The study team will summarize and report information in text and numeric tables to examine cross-site themes and the role of district contextual factors. While the study team will strive to provide quantifiable information where possible, much of their reporting will focus on themes and detailed descriptions rather than on summary statistics to help inform future program design. In particular, the study team will focus on providing specific details on the implementation process and perceived costs and benefits by type of alternative student outcome to help guide policymakers in decision-making and implementation.

The study team plans to examine variation not only across the measures but within the measures as well, when possible. In addition to looking at the commonalities of districts using the same alternative outcome, the study team will consider the range of experiences within a single category and even within a single district. Previous research suggests that districts vary in their approaches to implementing student learning objectives. In addition, there are likely differences in the extent to which assessments in a particular outcome category—for example, nationally normed assessments—are aligned with a single grade and subject. The PSAT, for example, is designed to reflect multiple years of learning that may not be attributable to a particular teacher.

b. Publication Plans

The study team will produce two “What’s Happening” reports during the course of the project. The first report, scheduled to be released in September 2013, will be based on document reviews and one pilot interview with a key staff person in each district, and will present preliminary findings. It will focus largely on the design and intended use of the alternative growth measures. The final report, scheduled to be released in August 2014, will focus on cross-measure analysis. The study team will present an overview of the types of alternative measures on which they collected data, followed by three sections on findings: (1) applications of the measures, (2) differentiation produced by the measures (including evidence of validity and reliability), and (3) implementation of the measures, including benefits and costs. Within each of the three sections the study team will compare findings across the types of alternative growth measures to allow readers to make comparisons based on the dimensions of interest. Both reports will be aimed at educators and policymakers and will use practitioner-friendly language to describe what the team has learned about the implementation of a variety of alternative measures of growth/value-added in the case study sites. The reports will be released by the REL and available on the REL web page. To avoid the risk of deductive disclosure of district identities, profiles of individual districts will not be included in the reports.

17. Approval Not to Display the Expiration Date for OMB Approval

The study will display the OMB expiration date.

18. Exception to the Certification Statement

No exceptions are being sought.

REFERENCES

- Community Training and Assistance Center (2004). *Catalyst for Change: Pay for performance in Denver*. Final Report. Retrieved on May 4th, 2012 from: <http://www.ctacusa.com/PDFs/Rpt-CatalystChangeFull-2004.pdf>
- District of Columbia Public Schools (2011). "IMPACT: The DCPS Effectiveness Assessment System for School-Based Personnel, 2011-2012," Washington, DC: District of Columbia Public Schools, 2011. Available at [<http://dcps.dc.gov/DCPS/Files/downloads/TEACHING%20&%20LEARNING/IMPACT/IMPACT%20Guidebooks%202010-2011/Impact%202011%20Group%202-Aug11.pdf>]. Accessed June 27, 2013.
- Gill, B., J. Bruch, and T.K. Booker. *Using alternative student growth measures for evaluating teacher performance: what the literature says*. (NCEE 2013-4006). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Mid-Atlantic Regional Educational Laboratory, 2013.
- Goe, L. and L. Holdheide. "Measuring Teachers' Contributions to Student Learning Growth for Nontested Grades and Subjects." National Comprehensive Center for Teacher Quality Research & Policy Brief, 2011.
- Goldhaber, D. and J. Walch. "Strategic Pay Reform: A Student Outcomes-Based Evaluation of Denver's ProComp Teacher Pay Initiative". Center for Education Data & Research Working Paper no. 2011-3.0, University of Washington, 2011.
- Goldstein, J. and P. Behuniak. "Growth Models in Action: Selected Case Studies." *Practical Research, Assessment, & Evaluation*, Vol. 10, No. 11, 2005, pp. 1-17.
- Johnson, M., S. Lipscomb, B. Gill, K. Booker, and J. Bruch. "Value-Added Models for the Pittsburgh Public Schools." Report to the Pittsburgh Public Schools. Cambridge, MA: Mathematica Policy Research, 2012.
- Kimball, Steven M., Matthew Clifford, Jenni Fetters, Jessica Arigoni, and Kim Bobola. "Evaluating Principals' Work: Design Considerations and Examples for an Evolving Field." Washington DC: US Department of Education, 2012.
- Marietta, G. "Multiple Measures of Teacher Effectiveness in Hillsborough County Public Schools: Implementing Value-added Measures." Retrieved on May 4th, 2012 from: <http://www.fadss.org/docs/content/eet/caseStudyHCPS/ImplementingValue-addedMeasuresHCPS.pdf>.
- Potamites, L., D. Chaplin, E. Isenberg, and K. Booker. "Measuring School Effectiveness in Memphis – Year 2." Report to New Leaders for New Schools. Washington, DC: Mathematica Policy Research, 2009.
- Zinth, J. D. (2010). "Teacher Evaluation: New Approaches for a New Decade." Teaching Quality-Evaluation and Effectiveness Issue Brief. Denver, CO: Education Commission of the States.