Part A: Supporting Statement for Paperwork Reduction Act Submission

Study of Teacher Preparation Experiences and Early Teacher Effectiveness Phase II–Data Collection

July 18, 2014

Prepared for:
Melanie Ali
Institute of Education Sciences
555 New Jersey Ave, NW
Room 502b
Washington, DC 20208-5500

Submitted by:
Abt Associates Inc.
55 Wheeler Street
Cambridge, MA 02138

In Partnership with:

Belmont Research Associates
The Bench Group
Dillon-Goodson Research Associates
Education Analytics
Pemberton Research

Part A: Supporting Statement for Paperwork Reduction Act Submission

Table of Contents

Table of Exhibits

Appendices

A. Justification

Introduction

The U.S. Department of Education (ED) is conducting a study examining the relationship between teacher preparation experiences and early teacher effectiveness (The Study of Teacher Preparation Experiences and Early Teacher Effectiveness, formerly known as The Study of Promising Features of Teacher Preparation Programs). This Information Collection Request (ICR) is the second of two ICRs for the study. The first ICR (*Phase I—Recruitment*; OMB-PRA: 1850-0891) requested clearance for recruitment activities. This second ICR, *Phase II—Data Collection*, requests clearance for data collection activities (obtaining teacher contact information from districts, collecting data from teachers on preparation experiences via an online teacher survey, and obtaining student data from districts). The packages have been submitted separately because the process of recruiting districts had to begin before the teacher survey was developed and piloted.[1]

This study is sponsored by ED's Institute of Education Sciences (IES) and is being implemented by Abt Associates Inc. and its partners: Belmont Research Associates, The Bench Group, Dillon-Goodson Research Associates, Education Analytics (EA), and Pemberton Research (together, the "study team").

Overview of the Study

Title II, Part A of the Elementary and Secondary Education Act—the Improving Teacher Quality State Grants program—focuses on improving teacher quality and increasing the number of highly qualified teachers. This study, which is the first of its kind, is about how to improve teacher quality through more effective preparation of new teachers. It will look at how the nature and intensity of preparation experiences are related to teacher effectiveness in promoting student achievement. For the purposes of this study, "teacher effectiveness" is defined as teacher value-added (TVA). The study will measure TVA using analytic techniques that isolate teacher contributions to student test score gains. The study has three primary questions:

1) What are the relationships between teacher preparation experiences and teacher effectiveness in the first year of teaching, measured by teacher value-added? (RQ1)

2) What are the relationships between teacher preparation experiences and teacher effectiveness with English learners in the first year of teaching, measured by teacher value-added for English learners? (RQ2)

3) Do relationships between teachers' preparation experiences and teacher effectiveness in the first year of teaching differ depending on teachers' assessments of the usefulness of the preparation experiences? (RQ3)

Measuring Preparation Program Experiences

To answer the research questions, the study needs to identify preparation experiences that could reasonably be hypothesized to affect student test scores. The challenge is that little or no research has

[1] At the time of the first ICR, the study design called for an experiment; however, efforts to recruit teacher pairs which met the experiment's requirements were unsuccessful, and therefore a value-added approach is now the focus of the study.

examined relationships between preparation experiences and test scores. (The 2010 National Research Council report on teacher preparation highlighted this lack of research.) However, the recent Measures of Effective Teaching (MET) Study (Kane & Staiger, 2012) did identify a set of teacher practices related to test score gains. The study team used the MET study as a framework for identifying preparation experiences that would foster these practices. The survey asks teachers to rate the frequency of these experiences in their teacher preparation program.

Specifically, in the MET study, teachers were rated on their instructional practices using four different observation systems. This process resulted in each teacher being rated on practices in 32 domains of instruction. The study team distilled these 32 domains into 12 broad topic areas, which form the basis for the teacher survey questions about preparation experiences. This reduction was based primarily on data from the MET study on (a) the distribution of teacher ratings on each practice, (b) the relationship of teacher ratings on each practice to teacher value-added scores, and (c) correlations of the practices across the four measures. For each of the 12 topic areas, the study team developed a small set of specific instructional strategies from the coding manuals for the four MET measures. The strategies align with teaching behaviors that coders are trained to look for as evidence of a teacher's skill with the measure-specific practices. Finally, the study team conducted cognitive interviews with teachers to ensure that the instructional strategies were being interpreted as intended.

The same kind of evidence provided by the MET study to inform the identification of instructional practices for the general classroom does not exist for English learner-specific practices. Some evidence does exist, however, at the individual strategy level. The study used four research syntheses and the 2007 and 2014 IES Practice Guides on instruction for English learners to identify English learner-specific instructional strategies that are supported by evidence of improving English learner achievement (August & Shanahan, 2006; Baker et al., 2014; Calderón, Slavin & Sanchez, 2011; Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006; Genesee, Lindholm-Leary, Saunders & Christian, 2006; Gersten, Baker, Shanahan, Linan-Thompson, Collins & Scarcella, 2007). A set of English learner-specific strategies were selected by 1) eliminating redundancies across strategies that were identified in multiple reviews, 2) consulting content experts to guide and prioritize the selection of strategies, and 3) conducting cognitive interviews with teachers to ensure that the strategies were being interpreted as intended. The final set of English learner-specific instructional strategies were grouped together into one English learner-specific topic area on the survey.

Sample

An analysis of statistical power indicated that a sample of 6,450 teachers is needed to answer the primary research questions with reasonable precision. To meet this sample size, the study aims to include as many teachers as possible in recruited districts who teach reading/English language arts (ELA) and/or mathematics, the two subject areas consistently tested across states. Because analyses will be conducted separately by subject, focusing on elementary-grade teachers increases the likelihood that each teacher contributes to both the reading/ELA and mathematics sample, which reduces the cost of recruiting districts and teachers. Fourth grade is the lower bound for the teacher sample due to data availability—proficiency testing typically begins in third grade, which means fourth grade is the first grade for which two years of test scores needed to estimate TVA are available (i.e., test scores from grades 3 and 4). The study selected sixth grade as the upper bound because it is assumed to be the highest elementary grade.[2]

[2] In some school districts, some or all of the sixth grade classrooms will be in middle schools. However, data from the Common Core indicates that a substantial proportion of sixth grade classrooms will be in elementary

Limiting the study to first-year teachers would minimize issues with recalling preparation experiences. However, data on teachers entering the teaching force indicated that meeting the sample size requirement would require two years of data collection (i.e., recruiting teachers in 2014-15 and again in 2015-16). To limit data collection to one year, the study decided to include teachers in their second or third years of teaching for whom TVA in their first year of teaching could be estimated. The three-year upper bound is driven by concerns about recalling preparation program experiences for teachers who have been teaching for more than three years.

Teachers in their first through third years will be asked the same questions about their teacher preparation experiences. The primary analyses will examine the relationship of teachers' preparation experiences to effectiveness in their *first year*. For example, a teacher in her second year in 2014-2015 will be asked about her preparation experiences and those experiences will be related to her value-added in the *previous* year, which corresponds to her first year of teaching.[3]

The teacher sample will be recruited from a purposive sample of approximately 50 large and moderate-sized school districts across the country that can meet the student-teacher data linkage requirements for estimating TVA. Targeted districts include the 10 largest districts in the United States (based on the expected number of eligible teachers) and other moderate-to-large districts (or local education agencies). For the majority of these districts, the study team is currently conducting research projects, has existing partnerships, and/or has access to key decision makers. The targeted districts are in 25 states and 49% percent of them include 10 percent or more English learners. Districts will be asked to provide the study team with teacher contact information for administering the (online) teacher survey on preparation experiences in spring 2015. The study team will also request student data from districts to estimate teacher value-added. The data will include student scores on state assessments (reading/ELA and mathematics) and student demographic data.

Analysis and reporting

Hierarchical analysis (students nested within classes nested within teachers) will be used to estimate the relationship between preparation program experiences and student test scores, controlling for student, class, teacher, and school baseline covariates. Reading/ELA and Mathematics scores will be analyzed separately. For the subset of teachers with five or more English learners, the study will examine the relationship between general preparation experiences and the test scores of English learners as well as English learner-specific preparation experiences and the test scores of English learners. The study team will produce a study report that is expected to be available by fall 2017.

Phase II—Data collection request

This ICR seeks clearance to obtain the following data:

- From school districts:

schools.

[3] If second- or third-year teachers have not taught in an eligible grade or subject in their first year but have in their second and/or third year of teaching, they will still be surveyed and included in secondary analyses examining the relationship of teacher preparation experiences to teacher effectiveness in the second and/or third years of teaching. Only second- or third-year teachers who meet the eligibility criteria in their first year of teaching will be included in the sample used to answer the primary research questions.

  - Teacher contact information (emails, phone numbers) for first-, second-, and third-year teachers in the district in the 2014–15 school year;
  - Student data:
    - Reading/ELA and mathematics state assessment data for students in grades three to six for four years (2011–12 through 2014–15); [4]
    - Student demographic data (English learner status, Special Education status, free/reduced price lunch status, gender, race/ethnicity) for students in each year of test data.

- From teachers:
  - Data to verify sample eligibility, to determine teacher preparation pathway, to assess the frequency and usefulness of preparation experiences, and to measure background characteristics.

This ICR provides a detailed discussion of the study data collections and the analysis and reporting of the data, as well as an overview of the study, including its design and data collection procedures. Copies of the data collection instruments (the teacher survey and the District Administrative Records Collection Protocol) are included in the appendices.

A.1 Circumstances Making the Collection of Information Necessary

Data from this study will be used to identify promising preparation experiences. These results can inform efforts of stakeholders invested in teacher preparation, including national, state, and local policy makers; teacher preparation programs and certifying institutions; districts; and schools. Policy makers and administrators engaged in teacher preparation and certification can learn about which preparation experiences are related to teacher effectiveness. Districts and schools seeking objective information to guide teacher hiring and placement decisions also may find the results valuable.

The study will address the goals of the authorizing legislation, Title II, Part A of the Elementary and Secondary Education Act, section 2121-2123 as amended by No Child Left Behind (20 USC 6621-6623), one purpose of which is to "increase student academic achievement through strategies such as improving teacher…quality."[5] The study goes beyond current research that examines teacher preparation programs as a whole and focuses on relationships between preparation experiences and effectiveness. This research is particularly critical in the current educational context, in which more states are implementing teacher evaluation systems and preparation programs are more likely to be held accountable for the effectiveness of their graduates.

Evidence that teachers have a central role in improving educational outcomes has accumulated for decades (as noted by Hanushek & Rivkin, 2010 and by Price et al., 2013). Training effective teachers is therefore central to improving achievement, but there is scant evidence on how to train teachers effectively. In 2010, a committee of the National Research Council (NRC) reviewed research on teacher preparation and found that although an evidence base exists about the characteristics that are

[4] For districts that administer state tests in the fall of each year, data will be obtained for students in grades four to seven for SY2012–13 through SY2015–16.

[5] Other purposes of Title II are to improve principal quality, increase the number of highly qualified teachers, and hold local education agencies accountable for increases in student academic achievement. This study is most closely related to improving teacher quality through more effective teacher preparation and training.

valuable for teachers to have, the evidence did not support conclusions about how preparation programs can develop these characteristics (NRC 2010). There also is a particular concern about how to improve the training of novice teachers because they are more likely to teach disadvantaged students.[6]

The fact that effectiveness varies between teachers raises the possibility that teacher preparation programs could be part of the explanation. But studies of teacher preparation programs have found little variation between programs. Furthermore, Price et al. (2013) noted that studies of preparation programs overstate the heterogeneity of their effects. Some programs had large or small effects but the source was sampling error rather than true differences. After adjusting for sampling error, differences were in the range of 0.02 to 0.04 standard deviation units, about one or two percentile points on standardized tests.[7] Based on this research and simulations carried out by the study team, a reasonable conclusion is that program effects are small.

The approach these studies used to measure effects, which groups novice teachers by their program (typically a university or college), may explain why the effects are small. Implicitly, this approach assumes that preparation experiences vary only because of the program attended. All teacher candidates from one program are assumed to have the average experience of that program. If average experiences in different programs are similar, the result will be small test score differences, as research has found. But the approach conceals variation in experiences that may be occurring *within* programs, which observations and discussions with preparation program directors suggest is likely. What remains to be investigated is whether *specific* experiences within programs affect test scores. This information would be useful for program design and improvement above and beyond knowing that a particular teacher preparation program has teacher-candidates that are (slightly) more effective.

Effectiveness also could be explained by factors related to a teacher's selection of their preparation program. If this were true, and the study does not try to control for these selection factors, they could bias the relationships between preparation experiences and teacher value added.[8] The study team reviewed research on teacher value-added to assess the potential for selection effects and to identify additional data that the study could gather to control for selection factors. In particular, research on Teach for America (TFA) identified a small set of characteristics related to value-added—perseverance, leadership skills, and grade point average (GPA) in the last two years of college

[6] The National Center for Education Statistics reports that 17 percent of teachers had less than five years of experience in low-poverty schools compared to 21 percent of teachers in high-poverty schools (http://nces.ed.gov/pubs2009/2009324.pdf, Table 4). Clotfelter et al. (2006) show a similar differential exists in North Carolina schools and that the pattern has been stable over time and between elementary, middle, and high schools (http://sanford.duke.edu/research/papers/SAN06-08.pdf).

[7] Only one study (Koedel et al., 2012) adjusted for clustering of students within teachers and estimation error in program effect estimates. In this study, when using all teacher preparation programs in the sample (average number of teachers per program is 54), the standard deviation of program effects was 0.016 for mathematics and 0.019 for Communication Arts. When the analysis was limited to teacher preparation programs with at least 50 teachers (average number of teachers per program is 83), the variation of program effects effectively was zero.

[8] The study team also investigated whether teachers can select preparation experiences. If more or less effective teacher candidates select different experiences, a reverse correlation could arise in which effectiveness appears to predict experiences. Review of available research and discussions with program directors and experts on preparation programs indicated it was unlikely that teacher candidates could select their preparation experiences.

(Dobbie, 2011). The study will collect data on two of these three characteristics—perseverance and leadership skills—to help control for selection. Perseverance will be measured using the Grit Scale (Duckworth & Quinn, 2009). GPA in the last two years of college will not be included due to differences between the study sample and samples used in studies of TFA sample. A large proportion of teachers in the study sample will have been in teacher education programs in their final two years of undergraduate study. Based on reports from preparation program deans, these students will have more homogeneous grade point averages than the general student population from which TFA draws candidates. This homogeneity makes GPA a weak predictor of value-added for the study sample.

The study team is aware of no published studies that examine the types of training experiences teachers receive in their preparation programs and how variation in these experiences relates to differential teacher effectiveness. This study will be the first to provide ED with evidence about the relationships between preparation experiences and teacher effectiveness and will make a strong contribution toward building an evidence base on this critical issue.

A.1 Purposes and Use of the Information Collection

This ICR requests approval for *Phase II—Data Collection*. The study will administer an online teacher survey that will collect information on 1) eligibility; 2) preparation pathway; 3) teacher reports of the frequency and usefulness of preparation experiences; and 4) background characteristics.

The teacher survey is necessary because information on preparation experiences must be collected from teachers directly. Other data for the study—teacher contact information and state assessment data—will come from existing district data sources.

All data will be collected by the Abt study team. The schedule, purposes, and uses of the data being collected are summarized in Exhibit A-1. Together, the teacher survey and student data will be used to estimate the relationships between teachers' preparation experiences and teacher effectiveness.

Exhibit A-1. Data Collection Plan

Schedule	Data Collection	Respondent	Purpose	Use
Fall 2014	*Teacher contact information* • Emails and phone numbers for first-, second- and third-year teachers in the district in the 2014-15 school year	District Human Resources Office	• To send teachers the survey and contact them for non-response follow-up	• Contacting teachers
Spring 2015	*Teacher survey data* • Years of teaching, grades and subjects taught • Preparation pathway • Preparation experiences and usefulness • Background characteristics	Teachers	• To verify sample eligibility[9] • To gather data on teachers not available in existing data sources	• Eligibility • Descriptive • Covariates in analysis models • Independent variables in analysis models

[9] Since districts may only count years of teaching in their district when reporting teachers' years of experience, the survey will ask teachers to confirm that they are in the first, second or third year of their teaching career (not limited to their tenure in their current district).

Schedule	Data Collection	Respondent	Purpose	Use
Fall 2014 and Fall 2015	*Student data* • Reading/ELA and mathematics state test data for all students in grades 3–6 in 2014–15, 2013–14, 2012–13, and 2011–12 and teachers linked to these students • Demographic characteristics of students	District Testing & Accountability Office	• To estimate teacher value-added in teachers' first year of teaching[10]	• Dependent variables in analysis models • Covariates in the analysis models

Teacher contact information. The study team will request teacher email addresses and phone numbers for all teachers who have been in the district for one, two or three years in fall 2014.[11] (See Appendix 1 for the District Administrative Records Collection Protocol, which provides a complete list of the teacher data elements that the study team will request from districts.)

Teacher eligibility. The online survey will include questions to confirm that the teacher is eligible to be included in the sample to answer the primary or secondary research questions. These questions will ask teachers about their year of teaching (in any district) and whether they are or have been responsible for teaching reading/ELA and/or mathematics to at least one classroom of general education students in grades 4, 5 or 6 in any of the following school years: 2012-13, 2013-14 or 2014-15. Only teachers who meet the eligibility criteria for their first year of teaching will be included in the sample to answer the primary research questions. (See Appendix 2 for the teacher survey.)

Preparation Pathway. The online survey will be used to collect information about the type of preparation program a teacher was/is in (traditional, alternative, TFA), the name of their preparation program, where they are in terms of completing their program (e.g., completed, completed/will complete at end of first, second or third year of teaching), if they did student teaching, and the name(s) of the schools in which they did student teaching, if applicable.[12] These data will be used to characterize the sample. In addition, data on the timing of completion of program will be used in

[10] These data will also be used to estimate second- and third-year TVA for eligible teachers at no extra cost to the study or burden to districts since all four years of data are needed to estimate first-year TVA for first-, second- and third-year teachers.

[11] The study team will also use student-teacher linkages that are part of the student assessment data requested in fall 2014 to further narrow the sample to first-, second-, and third-year teachers who meet the study's grade and subject criteria. This set of teachers will be sent the survey and their eligibility will be confirmed via questions at the beginning of the survey.

[12] Teach for America and some alternative certification programs (e.g., teacher residency programs) continue to provide support to teachers in their second, and sometimes third year in the classroom.

sensitivity analyses.[13] Data on preparation program names will be used to examine the generalizability of the sample and to examine between- and within-program variation in preparation experiences.

Preparation experiences. The online survey will collect data on teachers' preparation experiences and their assessment of the usefulness of these experiences in their classroom instruction. Teacher preparation experiences are defined as the frequency of four types of preparation experiences related to 13 topic areas. For each topic area, teachers are asked about experiences related to a small set of specific instructional strategies. For each of the instructional strategies, respondents are asked to rate the frequency of four types of preparation experiences that teacher candidates might experience as part of their teacher preparation program:

- reading about, hearing about, or seeing a role play of this strategy, such as during coursework;
- observing a teacher using this strategy in a K-12 classroom (in videos or during fieldwork or student teaching);
- practicing this strategy in a K-12 classroom *prior* to becoming a full-time teacher; and
- receiving feedback on your use of this strategy, including what you did well or how you could improve, from program staff or a cooperating teacher.

Respondents are also asked to rate how useful their teacher preparation experiences have been for their classroom instruction. Data from the frequency items will be used to create measures of frequency of preparation experiences for addressing the first and second research questions. Data from the usefulness items will be tabulated and used to address the third research question.

Teacher background characteristics. The online survey will be used to collect information about demographic characteristics (age, gender, race/ethnicity) that will be used to characterize the study sample. The survey will also ask about two other teacher characteristics that prior research has shown to be related to teacher value-added—perseverance and leadership skills (see Section A.1). These data will be used to create covariates for use in analysis models. The final questions on the survey ask teachers for the information that is necessary to provide to ACT Inc., or the College Board to locate teachers' ACT or SAT scores and release them to Abt Associates. For the ACT, this includes full name at time of testing, current name (if different), state in which the test was taken, and date of birth. For the SAT, this information includes full name at time of testing, current name (if different), and social security number. Teachers are assured that their scores will be used for analysis purposes only and will never be identified or shared outside of the study researchers. These scores are a proxy for prior academic skill and will be used as another covariate in the models to control for selection bias.

Student data. The study team will request reading/ELA and mathematics achievement scores for all students in study districts in grades three through six for 2011–12 through 2014–15.[14] Student

[13] Second- and third-year teachers who have not completed their programs at the time of the survey but who will be included in the primary analyses because they have first-year TVA will be reporting on preparation experiences that are happening after the outcome. The study team expects this to be a small subset of second- and third-year teachers and will conduct sensitivity analyses to determine if excluding these teachers from the analyses affects the overall findings.

[14] For districts that administer state tests in the fall of each year, data will be obtained for students in grades four through seven for SY2012–13 through SY2015–16.

demographic data (English learner status, Special Education status, free/reduced price lunch status, gender, race/ethnicity and age) will also be requested. Data for all third through sixth grade students (including those whose teachers have more than three years of experience) are being requested in order to more precisely estimate the models. See Appendix 1 for the District Administrative Records Collection Protocol, which specifies the student data elements that the study team will request from districts.

A.2 Use of Information Technology and Burden Reduction

The data collection plan reflects sensitivity to issues of efficiency, accuracy, and respondent burden.

The study will use a combination of mechanical and electronic technology to collect data. For each data collection task, the study team has selected the form of technology that enables the collection of valid and reliable information in an efficient way while minimizing the burden on district staff and teacher respondents.

During *Phase II—Data Collection*, districts that agree to participate can provide teacher contact information and student data to the study team electronically. To minimize burden during data collection, a study team member will be assigned as the primary contact for each district and the email address and telephone number of this team member will be included on all data requests. The study team will accept the teacher contact and student data in whatever file format and structure is most convenient for the district and legible for the study team. Because both teacher and student files contain Personally Identifiable Information (PII), the study team will require that electronic files be securely transferred by authorized school district staff via Abt's secure data transfer portal that is automatically encrypted. The portal will also allow districts to upload files with minimal effort and time.

The teacher survey will be administered online, allowing teachers to complete the survey easily at a time and place most convenient for them. Additionally, online administration can reduce time and human error associated with manual data entry because the data will be entered directly by respondents and loaded automatically into an electronic data file. The Abt study team will provide a dedicated email address and toll-free number for teachers to call for assistance with all aspects of the survey.

A.3 Efforts to Identify Duplication

No other large-scale, multi-state study has explored the extent to which teachers' experiences in their preparation programs are associated with effectiveness. Two studies (Boyd et al., 2008; Seidel et al., 2012) link preparation experiences, as measured in surveys; however, both studies are on a smaller scale (in a single district (Boyd et al., 2008) or in an individual state (Seidel et al., 2012).

IES has completed several evaluations of alternative pathways to teaching (Constantine et al., 2009; Clark et al., 2013), including highly selective alternative routes such as Teach for America and the Teaching Fellows Program (Clark et al., 2013). These studies focused on differences between teachers who select into different preparation pathways, rather than individual experiences teachers had in their preparation programs. There have also been studies comparing teachers' preparation programs (e.g., Grossman et al., 2008; Boyd et al., 2006). These studies focused on teacher preparation program features, but did not examine variation of teachers' experiences within programs.

The focus of the proposed study on teachers' preparation experiences, therefore, will provide new information not currently available in the field of teacher preparation.

To the extent possible, the study team will use existing data rather than duplicate data collection efforts. For example, instead of testing students, the study will use data already collected by districts as outcomes. Teacher contact information is also already collected by districts; the study team's request will not require districts to collect any additional data. Data collected in the teacher survey are not available elsewhere.

A.4 Efforts to Minimize Burden on Small Businesses

The study is purposefully designed to focus on large rather than small districts in order to obtain the teacher sample needed in the most efficient manner and to minimize the number of small districts that are burdened by this work.

A.5 Consequences of Not Collecting the Information

The full data collection plan described in this supporting statement is necessary for conducting this study, which is consistent with the goals of Elementary and Secondary Education Act, Title II, Part A, to raise student achievement through the preparation, training, and recruitment of high-quality teachers. There is evidence from multiple research studies that a student's teacher predicts a substantial amount of variation in his or her academic achievement (Chetty et al., 2011, Kane & Staiger, 2008; Hanushek & Rivkin 2010[15]). And, while research projects such as the MET study have helped policy makers and practitioners better understand what good teachers look like in the classroom and which instructional strategies and behaviors are linked to teacher effectiveness, the field lacks evidence regarding how to *produce* good teachers. This study will be the first to explore this important issue.

The consequences of not collecting specific data in *Phase II—Data Collection* are described below.

- Should the study fail to collect **teacher contact information**, it will not be able to administer the teacher survey.
- Without collecting **teacher survey data,** the study team will not be able to measure the frequency and usefulness of teachers' preparation experiences, nor link them to any variation in teacher effectiveness.
- Without collecting the **student data**, the study team will be unable to generate teacher value-added scores and link teachers' preparation experiences to teacher effectiveness.

A.6 Special Circumstances Justifying Inconsistencies with Guidelines in 5 CFR 1320.6

There are no special circumstances concerning the collection of information in this study.

[15] See Price et al., 2013 for a review of this research literature on variation in teacher effectiveness.

A.7 Consultation Outside the Agency

A.7.1 *Federal Registrar* Announcement

The 60-day notice to solicit public comments was published in Volume 79, Number 96, page 28689 of the Federal Register on May 19, 2014. One comment was received but it was positive in nature and did not require a response.

A.7.2 Consultation Outside the Agency

The Abt study team assembled a Technical Working Group (TWG) (in consultation with ED) that is composed of consultants with various types of expertise in the areas relevant to this study. The TWG convened in fall 2013 and discussed the study design, recruitment, instrumentation, data collection, and analysis and reporting of study findings. Members of the TWG and their affiliation are listed in Exhibit A.2.

Exhibit A-2. TWG Members and Affiliations

Name	Affiliation
Cory Koedel	University of Missouri
David Francis	University of Houston
Mari Koerner	Arizona State University
Kata Mihaly	RAND Corporation
Steven Rivkin	University of Illinois at Chicago
Suzanne Wilson	University of Connecticut

A.7.3 Unresolved Issues

There are no unresolved issues.

A.8 Payments or Gifts to Respondents

During *Phase II—Data Collection*, the study team proposes to include a modest incentive of $30 to teachers for the completion of the teacher survey, which is estimated to take approximately 35 minutes to complete. All eligible teachers who complete the survey will be sent a $30 Amazon e-gift certificate (delivered to them electronically).

The study team has reviewed the research literature on the effectiveness of incentives in increasing response rates for surveys. In the Reading First Impact Study commissioned by ED, monetary incentives proved to have significant effects on response rates among teachers. A sub-study requested by OMB on the effect of incentives on survey response rates for teachers showed significant increases when an incentive of $15 or $30 was offered to teachers as opposed to no incentive (Gamse et al., 2008). In another study, Rodgers (2011) offered adult participants $20, $30, or $50 in one wave of a longitudinal study and found that offering the highest incentive of $50 showed the greatest improvement in response rates and also had a positive impact on response rates for the next four waves.

In 2005, the National Center for Education Evaluation submitted a paper to OMB outlining guidelines for incentives for NCEE Evaluation Studies. The teacher survey incentive proposed for the current study is consistent with those guidelines, which classified a survey that required 30 minutes of burden as “high burden” and recommended a $30 incentive.

A.9 Assurance of Confidentiality

The study team will conduct *Phase II–Data Collection* activities in accordance with all relevant regulations and requirements. These include the Education Sciences Institute Reform Act of 2002, Title I, Part E, Section 183, which requires "[all] collection, maintenance, use, and wide dissemination of data by the Institute … to conform with the requirements of section 552 of Title 5, United States Code, the confidentiality standards of subsections (c) of this section, and sections 444 and 445 of the General Education Provisions Act (20 U.S.C. 1232 g, 1232h)." These citations refer to the Privacy Act, the Family Education Rights and Privacy Act, and the Protection of Pupil Rights Amendment.

In addition, all data collected for the study (Phases I and II) shall remain confidential in accordance with Section 552a of Title 5, United States Code, the confidentiality standards subsection (c) and sections 444 and 445 of the General Educations Provision Act. Subsection (c) of Section 183, referenced above, requires the director of IES to "develop and enforce standards designed to protect the confidentiality of persons in the collection, reporting, and publication of data." The study will also adhere to requirements of subsection (d) of Section 183 prohibiting disclosure of individually identifiable information as well as making the publishing or inappropriate communication of individually identifiable information by employees or staff a felony.

The study team will help maintain the confidentiality of data by requesting that districts provide a crosswalk with teacher identification numbers rather than providing student data linked to teacher names.

In addition, the following verbatim language will appear on all letters, fact sheets, and other study materials:

> *Per the policies and procedures required by the Education Sciences Reform Act of 2002, Title I, Part E, Section 183, responses to this data collection will be used only for statistical purposes. The reports prepared for this study will summarize findings across the sample and will not associate responses with a specific program, district, or individual. Any willful disclosure of such information for nonstatistical purposes, except as required by law, is a class E felony.*

Data will be presented in aggregate statistical form only. All study staff involved in collecting, reviewing, or analyzing individual-level data will be knowledgeable about data security procedures and will sign nondisclosure agreements. Respondents will be assured that all information identifying them or their school or district will be kept private to the extent allowed by law. The confidentiality procedures adopted for this study during all rounds of recruitment, data collection, data processing, and analysis consist of the following:

- All paper files will be converted to an electronic format and the paper files will be shredded immediately after they have been converted.
- Electronic data files with sensitive data will be removed from computers and working servers in a manner that ensures that the information cannot be recovered.
- At the end of contract with ED, all identifiers will be destroyed.

A.10 Questions of a Sensitive Nature

There are no questions of a sensitive nature included in the information requested in *Phase II—Data Collection*.

A.11 Estimate of Response Burden

The total respondent burden for the data collection effort covered by this clearance request is 5,608 hours, for a total cost to respondents of $187,370. Exhibit A-3 presents time estimates of respondent burden for the data collection activities requested for approval in this submission. The burden estimates are based on the following assumptions:

- The study team will obtain teacher contact information and student data from 50 district-level personnel. The total number of hours for each district is 20 hours. This estimate is based on the study team's prior work with districts in obtaining similar data, as follows:
    - Extracting and submitting teacher contact information to the Abt study team will take approximately four hours per district.
    - Collecting and submitting student data will take 16 hours (8 hours two times during the study period).[16]
- The total cost to districts is $38,040, based on an hourly wage of $38.04 in May 2012 for database administrators.[17]
- The study team will send surveys to approximately 9,930 fourth- through sixth-grade teachers, of which 80 percent, or 7,944 teachers, are expected to complete the online survey. The online survey is estimated to take 35 minutes.[18]
- The total cost to teachers is $149,330, based on an hourly wage of $32.41 in May 2012 for elementary and middle school teachers.[19]

Exhibit A-3. Estimate of Respondent Burden

Respondent/ Data Collection Activity	# of Respondents	# of Respon	Hours per	Total Burden	Total Burden	Hourly Wage	Total Cost per Respon	Total Costs

[16] It would be preferable for districts to provide student data in fall 2014 and again in fall 2015; however, if districts are only able to provide student data once during the study period, the study team will work with districts to accommodate this schedule.

[17] Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Employment Statistics*, accessed online at http://www.bls.gov/oes/current/oes151141.htm (December 23, 2013).

[18] The total number of teachers estimated to complete the survey includes the 6,450 teachers with first-year TVA needed to meet sample size requirements for primary analyses plus an additional 1,494 second- and third-year teachers who will only have second- and/or third-year TVA and thus be excluded from primary analyses but included in secondary analyses (estimates based on data from the Schools and Staffing Survey, https://nces.ed.gov/surveys/sass/).

[19] Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Employment Statistics*, accessed online at http://www.bls.gov/oes/current/oes_nat.htm (December 23, 2013). Since some sixth-grade teachers may teach in middle schools, the burden estimate uses the hourly rate for elementary and middle school teachers, rather than the rate for elementary school teachers only.

		ses	Response	Hours per Respondent	Hours	per Respondent	dent	
District Staff								
Teacher Contact Information	50	1	4	4	200	$38.04	$152.16	$7,608
Student Data	50	1	16	16	800	$38.04	$608.64	$30,432
District Subtotal	*100*	*1*	*20*	*20*	*1,000*	*$76.08*	*$760.80*	*$38,040*
Teachers (Elementary and Middle)								
Teacher Survey	7,944	1	0.58	0.58	4,608	$32.41	$18.80	$149,330
Total (Districts & Teachers)	**8,044**				**5,608**			**$187,370**

A.12 Estimate of Total Capital and Startup Costs/Operation and Maintenance Costs to Respondents or Record-Keepers

There are no annualized capital, startup, or ongoing operation and maintenance costs involved in collecting data in *Phase II*.

A.13 Estimates of Costs to the Federal Government

The estimated cost to the federal government of the activities in both *Phase I-Recruitment* and *Phase II—Data Collection* is $10,454,503. These activities will be carried out over six years (fall 2011 to fall 2017). Thus, the *average* annual cost to the federal government is $1,742,417.

A.14 Changes in Burden

This is an addendum package to a clearance request that was approved in June 2012 to conduct study recruitment activities (see OMB Control Number 1850-0891). Completion of the data collection activities described in this addendum package requires additional burden totaling 5,608 hours.

A.15 Plans for Analysis, Publication and Schedule

A.15.1 Analysis Plans

This section presents the study's estimation approach for addressing the three primary research questions.

1. **What are the relationships between teacher preparation experiences and teacher effectiveness in the first year of teaching, measured by teacher value-added?**

To address the first research question, the study team will analyze relationships between teacher preparation program experiences (TPPEs) and teacher effectiveness, measured by value-added. TPPEs are defined as the frequency of specific types of preparation experiences related to the representative instructional strategies within each key instructional topic area. For each topic area, four analytic variables representing the four types of preparation experiences will be created: 1) opportunities to "read about, hear about, or see a role play of the strategies, such as during

coursework;" 2) opportunities to "observe a teacher using the strategies in a K-12 classroom (in videos or during fieldwork or student teaching);" 3) opportunities to "practice the strategies in a K-12 classroom *prior* to becoming a full-time teacher;" and 4) opportunities to "receive feedback from program staff or a cooperating teacher on your use of the strategies, including what you did well or how you could improve." These TPPE measures will be used as explanatory variables in the analytic models.

Overview of Approach

The relationship of TPPEs will be examined separately for reading/ELA and mathematics but the analysis approach will be identical. For both outcomes, the study team will combine data from different school districts and grade levels to estimate relationships between TPPEs and value-added. May et al. (2009) argue that the decision of whether to combine data across grades or states should be driven primarily by a study's research questions. Combining scores across grades or states is appropriate when the study's questions are about student performance on state tests. Combining scores is inappropriate when the study's questions are about attainment of specific skills. Because the current study's primary research questions fall into the first category—relating TPPEs to value-added, as measured by student performance on state tests—combining data across states and grades is appropriate.

The study will use a two-stage approach to analyze the relationship between preparation experiences and value-added. The first stage will produce TVA scores for each teacher. Each TVA score is a measure of the extent to which a teacher's students experienced achievement growth over a school year, adjusting for student characteristics. In this stage, achievement scores will first be standardized using state and year means and standard deviations. The study then will fit models using data from each unique combination of state, grade, and year, which will yield a set of state/grade/year-specific TVA scores. In participating districts, test scores of all students in relevant grades will be used in the stage 1 analysis. The use of all student data in the relevant grades will increase the precision of the estimated coefficients and resulting TVA score estimates.

In stage 2, the set of state/grade/year-specific TVA scores will be combined into a single data set. Using the combined data set, regression models will be estimated with the TVA score as the dependent variable and one or more TPPE measures as independent variables, as well as grade and state dummies, and other statistical controls. Only the TVA scores of teachers who have completed the survey on preparation experiences will be used in the second stage analysis. Each of the two stages is described in more detail below.

Analytic Model for Stage 1

The stage 1 analysis will produce a TVA estimate for each teacher. The approach will utilize a hierarchical linear modeling (HLM) framework, which is common in education research. The analytic model is a three-level hierarchical linear model with students (level-1) nested in classes (level-2), and classes nested in teachers (level-3).[20] The level-1 model has student prior year scores and other student characteristics as covariates. The level-2 model includes average student measures 'centered' at the teacher-level mean (the overall mean at the teacher level will be subtracted from the classroom mean). The level-3 model includes these covariates aggregated to the teacher-level. In the level-3

[20] The study assumes that more than half of fifth- and sixth-grade teachers will teach more than one class in a single school year.

model, the dependent variable is the conditional TVA and the level-3 residual is the part of the TVA that has not been explained by covariates. The level-3 residuals from the stage 1 analyses are used as the dependent variable in the stage 2 analysis. Note that the model does not include dummies ("fixed effects") for schools.[21]

Specifically, the level-1 model, or student-level model is:

$$Y_{ijk} = \alpha_{0jk} + \sum_{m=1}^{M} \alpha_{mjk}(X_{mijk} - \overline{X}_{m.jk}) + e_{ijk}$$

The level-2 model or class-level model is:

$$\alpha_{0jk} = \beta_{00k} + \sum_{m=1}^{M} \beta_{0mk}(\overline{X}_{m.jk} - \overline{X}_{m..k}) + u_{jk}$$

The level-3 model or teacher-level model is:

$$\beta_{00k} = \gamma_{000} + \sum_{m=1}^{M} \gamma_{00m}(\overline{X}_{m..k}) + \sum_{t=1}^{T} \psi_{00t}(Z_{tk}) + r_k$$

where, in the level-1 model,

Y_{ijk} is the spring reading/ELA or mathematics achievement test score from ith student (i in 1,2,...,n) in the jth class, (j in 1,2,..., J), nested in the kth teacher (k in 1,2,...,K teachers).

$(X_{mijk} - \overline{X}_{m.jk})$ is the mth of M student characteristics (e.g., prior year test score, gender, race/ethnicity age, English learner status, free/reduced price lunch status), centered at the class-level mean.

e_{ijk} is the student-level error, assumed distributed normal with mean zero and variance σ_e^2

α_{0jk} is the covariate adjusted mean of achievement scores for the jth class of the kth teacher.

[21] Inclusion of fixed effects for schools limits the analysis sample to teachers from schools where there is more than one eligible teacher in the school and where two or more teachers have different values on the TPPE measures of interest. This approach could result in a very small analysis sample that would be a poor representation of the full sample of teachers that will be surveyed.

In the level-2 model,

$(\overline{X}_{m.jk} - \overline{X}_{m..k})$ are the student characteristics aggregated up to class-level means, and centered around the teacher's mean.

u_{jk} is the class-level error, assumed distributed normal with mean zero and variance σ_u^2

In the level-3 model,

β_{00k} is the conditional *mean* of achievement scores for the k[th] teacher (i.e., the conditional TVA score)

γ_{000} is the conditional grand mean (predicted mean for teachers when all teacher-level covariates are zero).

$(\overline{X}_{m..k})$ are the means of student characteristics for the k[th] teacher, averaged over all of the classes taught by the k[th] teacher.

(Z_{tk}) are measures of school characteristics (e.g., percent LEP, percent FRPL), and dummies for district.

r_k is the teacher-level error, assumed distributed normal with mean zero and variance σ_k^2

As mentioned above, this model will be fit separately for each unique combination of state, grade, and school year. The fitted models will produce an estimated TVA "score" (the level-3 residual r_k) and its standard error.

Analytic Model for Stage 2

In stage 2, a data set will be created that merges teachers' TPPE measures and their TVA scores from stage 1. Regression models will be used to estimate the relationships between TPPEs and TVA scores. These models will use the TVA as the dependent variable and TPPE measures; measures of perseverance; leadership; and prior achievement (ACT/SAT scores), which are collected via the survey to account for selection; and indicator variables for state, grade, and teachers' years of experience at the time of the survey as explanatory variables.[22] The outcome variable in the model (TVA score) will be weighted inversely proportional to the square of the standard error of the TVA

[22] Some teachers may have multiple TVA records. If there are more than a negligible number of teachers with multiple records, the model will be fit as a two-level hierarchical model with repeated observations at level-1 and teachers at level-2.

score, which means teachers with more precisely estimated TVA scores will receive more weight in the estimation.

2. What are the relationships between teacher preparation experiences and effectiveness with English learners in the first year of teaching, measured by teacher value-added for English learners?

To address the second research question about relationships of TPPEs to value-added for English learners, the models described in the previous section will be fit to subsets of data consisting of student outcomes for English learners, and teachers of classes that have a minimum of five English learners. Specifically, in the stage 1 models that yield TVA estimates, all of the student-level (level-1) outcomes and covariate data will be specific to only English learner students. For the class-level (level-2) and teacher-level (level-3) covariates that represent aggregates of student characteristics, the data from all students (both English learners and non-English learners) will be aggregated such that these measures will represent the characteristics of all of the students in a teacher's class or classes. These models will yield state, grade, and year-specific TVA estimates that pertain to only English learners. Only teachers with at least five English learners and their English learner-specific TVA estimates will be used in the stage 2 models. Measures of TPPEs will include general measures (i.e., the same measures as will be assessed in the primary research question) and TPPEs that are English learner-specific.

3. Do relationships between teachers' preparation experiences and effectiveness in the first year of teaching differ depending on teacher assessments of the usefulness of the experiences?

In addition to rating the frequency of specific types of preparation experiences related to key instructional topic areas, teachers are also asked on the survey to rate the usefulness of their TPPEs.

To answer this research question, descriptive data will be tabulated from the usefulness items to describe teachers' perceptions of the usefulness of TPPEs for first, second and third year teachers separately, and for all teachers combined. The stage 2 analytic models described above will also be estimated, but in each model, the TPPE measure of interest will be replaced with a dichotomous variable for usefulness: 1 = useful (if TPPE is rated as useful or very useful), 0 = not useful (if TPPE rated as not useful or a little useful). The coefficient for the usefulness variable will represent the relationship between the perceived usefulness of a TPPE and TVA. That is, it will address the question of whether TPPEs that are rated as useful are more strongly associated with teacher effectiveness than TPPEs that are rated as not useful.

A secondary research question about the usefulness of TPPEs is whether teachers' assessment of the usefulness of their experiences changes over time. A descriptive analysis will compare the proportions of teachers who describe a TPPE as useful among the first-, second-, and third-year teacher samples.

A.15.2 Plans for Tabulation and Publication of Results

The Abt study team will disseminate the results of the Study of Teacher Preparation Experiences and Early Teacher Effectiveness in a single report expected to be released by fall 2017.

The report will include a description of the study design and findings from all three primary research questions. More specifically, the report will include:

- Descriptive analyses of the characteristics of the study sample;
- Analyses of the relationships between preparation experiences and teachers' effectiveness in their first year of teaching (RQ1);
- Analyses of the relationships between preparation experiences and teachers' effectiveness with English learners in their first year of teaching for the subset of teachers with five or more English learners in their first-year classroom (RQ2); and
- Analyses of whether the relationships between preparation experiences and teachers' effectiveness in their first year of teaching differ depending on teachers' evaluations of the usefulness of the preparation experiences in instruction (RQ3).

These analyses will be carried out using hierarchical linear modeling to take into account nesting (e.g., the nesting of students and classes within teachers) and will incorporate student covariates to maximize precision.

A.16 Approval to Not Display Expiration Date

No exemption is requested. The data collection instruments will display the expiration date.

A.17 Exceptions to Item 19 of OMB Form 83-1

The submission describing data collection requires no exemptions to the Certificate for Paperwork Reduction Act (5 CFR 1320.9).

References

August, D., & Shanahan, T. (2006). *Developing literacy in a second-language: Report of the National Literacy Panel on Language Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum Associates.

Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., Gersten, R., Haymond, K., Kieffer, M. J., Linan-Thompson, S., & Newman-Gonchar, R. (2014). *Teaching academic content and literacy to English learners in elementary and middle school* (NCEE 2014-4012). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: http://ies.ed.gov/ncee/wwc/publications_reviews.aspx.

Boyd, D. J., Grossman, P., Lankford, H., Loeb, S., Michelli, N. M., & Wyckoff, J. (2006). Complex by Design Investigating Pathways Into Teaching in New York City Schools. *Journal of teacher education*, *57*(2), 155-166.

Boyd, D.J., P.L. Grossman, H. Lankford, S. Loeb, and J. Wyckoff. (2009). Teacher preparation and teacher effectiveness. *Educational Evaluation and Policy Analysis*, 31, 416. DOI:10.3102/0162173709353129.

Calderón, M., Slavin, R., & Sánchez, M. (2011). Effective instruction for English learners. *The Future of Children*, *21*(1), 103-127.

Chetty, R., Friedman, J.N., & Rockoff, J.E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood* (No. w17699). Cambridge, MA: National Bureau of Economic Research.

Clark, M., Chiang, H., Silva, T., McConnell, S., Sonnenfeld, K., Erbe, A., & Puma, M. (2013). *The effectiveness of secondary mathematics teachers from Teach for America and the Teaching Fellows programs* (NCEE 2013-4015). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Clotfelter, C., Ladd, H. F., Vigdor, J., & Wheeler, J. (2006). High-poverty schools and the distribution of teachers and principals. *NCL Rev.*, *85*, 1345.

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). An Evaluation of Teachers Trained through Different Routes to Certification. Final Report. NCEE 2009-4043. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Dobbie, W. (2011). Teacher characteristics and student achievement: Evidence from Teach for America. *Mimeo, Harvard University*.

Duckworth, A. L. & Quinn, P. D. (2009). Development and Validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment, 91*(2), 166-174.

Francis, D. J., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). Practical guidelines for the education of English language learners: Research-based recommendations for instruction and academic interventions. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Gamse, B. C., Bloom, H. S., Kemple, J. J., & Jacob, R. T. (2008). Reading First Impact Study: Interim Report. NCEE 2008-4016. National Center for Education Evaluation and Regional Assistance.

Genesee, F., Lindholm-Leary, K., Saunders, W. M., & Christian, D. (Eds.) (2006). *Educating English language learners: A synthesis of research evidence.* Cambridge University Press.

Gersten, R., Baker, S.K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective literacy and English language instruction for English learners in the elementary grades*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education.

Grossman, P., Hammerness, K. M., McDonald, M., & Ronfeldt, M. (2008). Constructing Coherence Structural Predictors of Perceptions of Coherence in NYC Teacher Education Programs. *Journal of Teacher Education*, *59*(4), 273-287.

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *The American Economic Review*, 267-271.

Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation* (No. w14607). National Bureau of Economic Research.

Kane, T.J. & Staiger, D.O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. *MET Project Policy and Practice Brief prepared for Bill and Melinda Gates Foundation*.

Koedel, C., Parsons, E., Podgursky, M., & Ehlert, M. (2012). Teacher Preparation Programs and Teacher Quality: Are There Real Differences Across Programs?. *National Center for Analysis for Longitudinal Data in Education Research, Working Paper, 79*.

May, H., Perez-Johnson, I., Haimson, J., Sattar, S., & Gleason, P. (2009). *Using state tests in education experiments: A discussion of the issues* (NCEE 2009-013). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

National Research Council. (2010). Preparing teachers: Building evidence for sound policy. Washington, DC: National Academies Press (NRC).

Price, C., Goodson, B., Dynarski, M., Unlu, F., Caswell, L., & Shivji, A.. (2013). A review of the research on the relationship between teacher preparation programs and teacher effectiveness. Unpublished manuscript. Bethesda. MD: Abt Associates.

Rodgers, W. (2011). Effects of increasing the incentive size in a longitudinal study. *Journal of Official Statistics* 27(2), 279–299.

Seidel, K., Green, K., & Briggs, D. (2012). An exploration of novice teachers' core competencies: Impacts on student achievement and effectiveness of preparation. Institute of Education Sciences Award Number R305A120233, U.S. Department of Education. Project details online at www.portfolio.du.edu (Novice Teachers Project).