

Supporting Statement for Paperwork Reduction Act Submissions for the Study of Emerging Teacher Evaluation Systems in the United States

Part A. Justification

A.1 Circumstance Requiring Collection of Information

Most current teacher evaluation systems fail to differentiate effective from ineffective teachers or provide feedback and support in order to help teachers improve. While teachers' effectiveness in increasing student learning varies significantly, the majority of school districts do not evaluate teachers in a manner that distinguishes effective teachers from ineffective teachers or takes student achievement into account in the evaluation (Kane, 2009). A recent study of 12 districts in four states showed that, in districts with binary evaluation ratings (generally "satisfactory" or "unsatisfactory"), more than 99 percent of teachers received a satisfactory rating (Weisberg et al., 2009). Moreover, in districts with a broader range of ratings, 94 percent of teachers received one of the top two ratings and less than 1 percent received an unsatisfactory rating. In addition, three different studies of typical teacher evaluations used in districts found that these evaluations were not designed or used to provide feedback in order to help teachers to improve or to guide teacher professional development (Mathers, Oliva, & Laine, 2008).

Likewise, tenure and licensure are important milestones in educators' careers, but states and districts do not always take performance into account in making these decisions. Evidence of teacher effectiveness is a preponderant criterion in licensure advancement in only three states (National Council on Teacher Quality, 2011). A recent report noted that 46 states award teacher tenure "with little or no attention paid to how effective they are with students in their classrooms" (National Council on Teacher Quality, 2011).

To address the shortcomings of current teacher evaluation systems, the U.S. Department of Education (Department) has embarked on an ambitious agenda to support states and districts in implementing teacher evaluation systems that:

- Apply multiple measures of teacher effectiveness, including gains in student achievement and observation of classroom practice
- Distinguish between teachers of different levels of effectiveness
- Include a formative component that provides feedback to teachers in order to help them to improve
- Are aligned with a comprehensive human capital system

As part of this agenda, the Department's Office of Policy and Program Studies Service (PPSS) is supporting a "Study of Emerging Teacher Evaluation Systems in the United States,"

which will contribute to the Department's work by providing research-based information to aid state and local efforts to plan and implement comprehensive teacher evaluation systems. The study includes a review of the research on teacher evaluation practices, programs, and policies, and nine case studies. The case study sample will include five fully operational teacher evaluation systems and four systems in the early implementation phase.

The research team, made up of employees at Policy Studies Associates, its subcontractor, and a consultant, will develop and execute plans for the nine case studies, which will include (a) interviews with leaders and staff responsible for teacher evaluation systems and (b) review of key documents and artifacts from each of the study sites. The research questions and expected data sources for each question are displayed in Exhibit 1 on the next page.

Authorization to conduct this study is provided by the Elementary and Secondary Education Act (ESEA), as reauthorized by the No Child Left Behind Act (NCLB) (Public Law 107-110), Section 9601(a) and Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, 2010, Division D, Title III, (Public Law 111-117); Department of Defense and Full-Year Continuing Appropriations Act, 2011, (Public Law 112-10).

A.2 Indicate How, by Whom, and for What Purpose the Information Is to be Used

The study will produce two reports. One will synthesize the results of the literature review. The other will present findings and observations from the nine case studies. Both publications will be targeted at state and local practitioners and policymakers who are implementing new teacher evaluation systems. Both publications will be disseminated on the Web and will include artifacts from the teacher evaluation systems included in the case studies. The study team will also provide briefings for policy audiences. We anticipate that the study results will help states that seek increased flexibility under NCLB develop and implement new teacher evaluation systems. The study will also serve as a resource for other support that the Department will provide for the development and implementation of new teacher evaluation systems at the state and local levels.

A.3 Use of Information Technology to Reduce Burden

We will conduct extensive online searches of the websites of (a) state education agencies (including portions of the websites that may be dedicated to Race to the Top (RTT) activities, State Fiscal Stabilization Fund annual performance reports, and other relevant policy and program areas), (b) state teachers associations, (c) school districts, and (d) local teacher associations prior to each site visit. We anticipate that these searches will yield a substantial amount of documents pertaining to the history, structure, and operation of the teacher evaluation system in each of the sites. We also anticipate that additional program materials and artifacts will be identified as we work with sites to plan the data collection visits. Thorough analysis of these documents will allow the research team to use the extant information to develop at least

partial answers to the study's research questions and to refine and streamline interview protocols prior to the site visits so that the interviews focus only on issues not covered in the documents.

Exhibit 1: Data Collection Matrix

Research Questions	Data Sources					
	Document Review	Interviews				
		State Education Agency Staff	Other State-Level Stakeholders ¹	District Leaders/Staff ²	Other Local Stakeholders ³	Teachers (in focus groups) ⁴
1. Key System Components						
1. What standards, frameworks, and/or models of good teaching inform these systems?	X	X	X	X	X	X
2. What <i>models and data on changes in student learning</i> do emerging teacher evaluation systems use to rate teacher performance and how are these ratings developed? What provisions are made for evaluating teachers who teach subjects and/or grade levels for which there are no standardized student assessment data?	X	X	X	X	X	X
3. How do emerging teacher evaluation systems assess (a) instructional practice, (b) teacher planning and preparation, and (c) other dimensions of professional practice?	X	X	X	X	X	X
4. Who is responsible for the core evaluation tasks (e.g., observing practice, determining teacher performance ratings, determining ratings based on student learning gains, providing feedback to teachers, planning and facilitating follow-up support) and how are they prepared and held accountable for successful completion of these tasks?	X	X		X	X	
5. How and to what extent do system design and operation consider the unique circumstances of (a) novice teachers, (b) special education teachers, (c) teachers of English language learners, and (d) teachers whose regular assignments include ongoing collaboration and partnerships with other teachers (e.g., reading and math specialists, members of grade-level or subject-area teams)?	X	X	X	X	X	X
6. How and to what extent have state statutes, policy, guidance, and training and technical assistance influenced system design and operation?	X	X	X	X	X	

¹ Includes leaders/representatives of state teacher unions and other state-level, such as leader/members of state planning group

² Includes director of teacher evaluation system, director of assessment, director of professional development, representatives of planning group (including some teachers), principals and others responsible for evaluation data collection and feedback

³ Includes union leaders and others identified as involved in system planning

⁴ Includes teachers identified as being in one of the target groups for focus group interviews and who have been through the evaluation process. Does not include teachers included in interviews under #2 above

Exhibit 1: Data Collection Matrix (continued)

Research Questions	Data Sources					
	Document Review	Interviews				
		State Education Agency Staff	Other State-Level Stakeholders	District Leaders/Staff	Other Local Stakeholders	Teachers (in focus groups)
2. Use of System Outputs						
1. How do the developing systems use (or expect to use) the system outputs—teacher performance ratings—to (a) improve practice; (b) recognize and reward exceptional performance; (c) monitor and adjust the distribution of highly-effective teachers to ensure adequate staffing in low-performing schools; (d) alter the composition of the teacher workforce; (e) report to key stakeholders on teacher quality and equitable distribution; and (f) evaluate and improve system design and operations?	X	X	X	X	X	X
3. Implementation, Challenges, and Lessons Learned						
1. What were the key steps in system design and implementation? Who was involved? What role did they play? In retrospect, who was not involved but should have been involved?	X	X	X	X	X	
2. How and to what extent did implementation include efforts to raise and focus district-wide expectations for teacher performance and student learning?				X	X	X
3. How long did it take for the district to develop and implement the new teacher evaluation system?	X	X	X	X	X	
4. What resources are required for the annual administration of the new teacher evaluation system (e.g., contracts, staff time, etc.) compared with the previous system? (NOTE: Only asked in fully implementing districts)				X		
5. What challenges have been encountered in implementing and operating these systems and how have the challenges been overcome?	X			X	X	X
6. How have lessons from early implementation been translated into modifications to system design and operation?				X	X	
7. What evidence suggests that system implementation has (a) changed the discourse on professional practice and student learning; (b) re-defined expectations for principals and teachers as instructional leaders; and/or (c) re-focused professional development policies and practices?				X	X	X
8. What lessons from the experiences of the sites in the current study can inform next steps in developing these systems in other districts and states?				X	X	X

A.4 Avoidance of Duplication

We have determined that no previous federal studies have examined the design and implementation of local teacher evaluation systems that reflect the Department's vision and expectations for these systems. Further, this study does not duplicate two Institute for Educational Sciences (IES) efforts in the area of teacher evaluation:

1. ***Learning from Emerging Teacher Evaluation Practices to Advance Teacher Quality*** is a research effort by the Carnegie Foundation for the Advancement of Teaching that began in the fall 2010. The purpose of the project is to inform the development of rigorous teacher evaluation systems. The primary focus is on the technical concerns encountered by designers, development and accountability analysts who are building new tools and protocols for teacher assessment. This project is intended to (a) bring together multiple stakeholders to build a research agenda around teacher evaluation systems and protocols, and (b) quickly synthesize and disseminate findings, successes, and challenges related to teacher evaluation systems and protocols. During the three-year period, the Carnegie Foundation is functioning as an analytic hub and is providing a neutral convening context for this networked improvement community.

The Contracting Officer's Representative (COR) for the PPSS teacher evaluation study has met with IES staff overseeing the cooperative agreement for *Learning from Emerging Teacher Evaluation Practices to Advance Teacher Quality* and the PPSS COR has determined that the two research efforts are complementary and are not duplicative. In particular, the Carnegie effort will not include a literature review of the research on teacher evaluation systems and it will not include descriptive case studies of teacher evaluation systems. In addition, IES is ensuring that the research staff at Carnegie know about the PPSS teacher evaluation study and will share any recommended districts or community based organizations that may be good candidates for the PPSS study. Finally, one of the leaders of the Carnegie effort – Paul LeMahieu – serves on the Technical Working Group for the PPSS study.

2. ***Impact Evaluation of Teacher and Leader Evaluation Systems*** is being conducted through an IES contract that was awarded in September 2011. This evaluation will recruit districts to implement new teacher evaluation systems and will randomly assign schools in the districts to use the new teacher evaluation system. The IES study is designed to provide estimates of the impact of evaluation systems on teacher retention rates and student achievement whereas the PPSS study is designed to address qualitative questions about program implementation. The PPSS study is intended to provide practical, descriptive information for state and district practitioners by early 2013 whereas the data collection for the IES evaluation will not take place until the 2012-2013 school year. In addition, there will be no overlap in study sites because the IES study will recruit school districts that are not yet implementing teacher evaluation systems that reflect the Department's vision for such systems.

A.5 Methods to Minimize Burden on Small Entities

All entities participating in this data collection effort are state education agencies, school districts, or charter management organizations (CMOs). No small businesses will be involved in any way, and the state education agencies, school districts, and CMOs included in the study are large enough that they do not meet the definition of small entities.

A.6 Consequences of Not Collecting Information

Without the reports from this study, the Department will not have some key information that leaders and staff need in order to help guide and inform development of new teacher evaluation systems as outlined in requirements and guidance for RTT and TIF. These requirements and guidance call on states and districts to design and implement new teacher evaluation systems that rely on measures of student achievement gains and performance measures to rate teacher effectiveness. Through the RTT and TIF grant programs, the Department has emphasized policies related to evaluating educator effectiveness and providing educators useful and timely feedback needed to improve practice and, ultimately, student achievement. The RTT program, which awarded a total of \$4 billion in 2010, was, among other things, designed to spur states to improve teacher and principal effectiveness by designing and implementing rigorous, transparent and fair evaluation systems for teachers and principals that differentiate effectiveness using multiple rating categories that take into account data on student growth as a significant factor. Such evaluations are intended to be conducted annually and provide timely and constructive feedback to teachers and principals. Similarly, TIF awarded a total of \$448 million in new grants in 2010 to implement performance-based compensation systems which are based on effective performance evaluation systems.

In addition, the Department's proposal for reauthorization of ESEA includes formula grants to states and districts to improve the effectiveness of teachers and leaders, and ensure that students in high-need schools are being taught by effective teachers in schools led by effective principals. Under the proposal, districts would be required to implement evaluation systems that (a) meaningfully differentiate teachers and principals by effectiveness across at least three performance levels; (b) are consistent with their state's definitions of "effective" and "highly effective" teacher and principal; (c) provide meaningful feedback to teachers and principals to improve their practice and inform professional development; and (d) are developed in collaboration with teachers, principals, and other education stakeholders.

Thus far, research on how these systems are implemented or how they operate when they are fully implemented has been limited. While the literature review (which is not part of this clearance request) will provide a summary of existing research in the field, the case studies will go into greater depth by exploring how emerging teacher evaluation systems operate, the challenges they face, and the extent to which they are meeting expectations for providing meaningful assessments of teacher performance. Both the literature review and report on the case studies will inform new and ongoing district and state efforts to design and implement systems that meet the Department's expectations. The TIF program office, the Implementation and Support Unit (RTT), and Department leaders will share the literature review and case study report with grantees and other states and districts that are trying to implement new teacher

evaluation systems so that they can learn from past research and from on-going efforts in districts and in one CMO. The Department's leaders and program staff feel that the field needs to learn from the current efforts of districts and CMOs and that these case studies will provide this vital information in a timely fashion.

A.7 Explain Any Special Circumstances

None of the special circumstances apply to the planned case studies.

A.8 Consultation Outside the Agency

We consulted with a Technical Working Group (TWG) made up of experts on teacher evaluation to obtain their input on the design of the study. The interview protocols included in this clearance request reflect the early advice of the TWG, and the team will continue to consult these experts several times during the course of the study. The TWG members are listed below:

Joanna Cannon

Director of Research
Office of Accountability
Department of Education
New York City

Laura Goe

Associate Research Scientist
Educational Testing Service
Senior Researcher
National Comprehensive Center on Teacher
Quality

Paul LeMahieu

Senior Partner
Carnegie Foundation for the Advancement
of Teaching

Leigh McGuigan

Partner
The New Teacher Project

Tony Milanowski

Senior Research Analyst
Westat

Pamela Moran

Superintendent
Albemarle County (VA) Public Schools

Susan Sclafani

National Center on Education and the
Economy

Andy Sokatch

Vice President for Research
Teach for America

Rob Weil

Director
Field Programs, Educational Issues
American Federation of Teachers

In addition to consulting with the TWG, we have pre-tested each interview protocol with state and local leaders and staff who are familiar with emerging teacher evaluation systems. In addition, we have pre-tested the teacher focus group protocol with teachers who have participated in at least one annual cycle of the new teacher evaluation in their district. We discuss the results of the pre-test in more detail in Section B.4, below.

A.9 Payments of Gifts

We will invite each participating district to nominate a liaison to assist in scheduling interviews and identifying documents pertinent to the study that we may not have located in our online searches. Because this assistance is critical to the overall success of the study and will reduce the overall burden on district staff, we will provide honoraria of \$500 and \$250, respectively, to liaisons in districts with fully operational systems and in districts in the early phase of system implementation. As we have learned during the pre-testing of the interview protocols identifying potential respondents and scheduling interviews with them is time consuming and requires someone in the district to assist in the process. In the fully operational sites, the scheduling task is more complex because of the need to schedule seven teacher focus group interviews with teachers from different schools and with different kinds of assignments. We will also call on the site liaisons to assist in arranging locations for the interviews. We anticipate that most one-on-one interviews will take place in the respondents' offices and that the teacher focus groups will be scheduled for somewhat centralized locations that are reasonably convenient to the participants. We propose to pay the honoraria as a thank you to the site liaisons for what will be their invaluable assistance in completing our planned data collection activities.

No payments or gifts will be made to state education agency or district staff who participate in interviews for the study. However, because the teacher focus groups will last 75 minutes and some teachers may commute from one school to another school or to the district office to participate in the focus group discussion, we will provide each teacher with a \$10 gift certificate to Starbucks or Barnes and Noble as a token of appreciation.

A.10 Assurances of Confidentiality

We will make every effort to protect the privacy and confidentiality of all teachers, state agency staff, district staff, and other individuals who participate in the study. In particular, we will not identify any individuals by name in any reports or other communications about the study, such as briefings on the study findings for Department staff and others. In addition, we will not identify any schools.

The final project report will, as appropriate, look across sites and present aggregate findings or use case study data to provide examples of program implementation and challenges in a manner that does not associate responses with a specific site or individual. We do, however, anticipate identifying the participating districts in the study report because, among other things, we anticipate that the final report on the study will include brief profiles of each of the nine emerging teacher evaluation systems. Further, because the public nature of the teacher evaluation systems makes district leaders such as superintendents and teacher evaluation system

directors easily identifiable, they will consequently be identified by position/role but not by name in the report. We will invite respondents who may be identified by position/role in the project report to review the relevant text in draft prior to submitting the report to the Department in draft or in its final form and notify us of any errors of fact or about any questions or concerns that they may have about the text. Overall, except as explained here, the contractor will not provide information that associates responses or findings with an individual, district, or school to anyone outside the study team, except as may be required by law.

Prior to each individual interview and each focus group interview, we will explain the purpose of the study, the topics that we will cover in the interviews, and the confidentiality assurances discussed above. We will also indicate that participation in the study as well as responding to individual interview questions is voluntary and that respondents may decide not to participate or to end their participation at any time.

The study team has extensive experience in protecting the privacy and confidentiality of interview respondents. Safeguards to protect the privacy and confidentiality of all respondents that we will use in addition to the ones discussed above include the following:

- All individual contact information that may be available to the study team will be used for scheduling interviews only and will be destroyed as soon as the interviews and necessary follow-ups are completed.
- All audiotapes and notes from individual interviews and teacher focus groups as well as all documents that contain sensitive, personally identifiable information will be maintained in secure files accessible only by members of the study team.
- All audiotapes, interview notes, and sensitive documents will be destroyed upon submission of the final report on the case studies.
- Training for site visits will familiarize team members with the confidentiality provisions discussed above and their responsibilities for explaining those provisions to respondents and maintaining the necessary safeguards in storing and using interview data for analysis and reporting.

A.11 Justification of Sensitive Questions

The interview protocols do not include any questions of a sensitive nature.

A.12 Estimates of Respondent Hour Burden and Annualized Cost

Data collection for the five case studies of sites that are fully operational will include up to 264 individual interviews, including up to four telephone interviews with state-level respondents and up to 220 in-person interviews with district staff and local stakeholders.⁵ In

⁵ For purposes of this study, fully operational sites are sites that have completed at least one full annual evaluation cycle prior to the planned data collection. This means that these sites would have completed at least one annual

addition, we will conduct seven teacher focus groups, with up to seven teachers in each group. We estimate that the individual interviews will last 45 minutes and that the focus groups will last up to 75 minutes. Data collection for the four case studies of sites that are in the early stages of implementing new teacher evaluation systems will include up to 24 individual interviews, including up to four telephone interviews with state-level respondents and up to 20 in-person interviews with district staff and local stakeholders.⁶ (Following the Department’s specifications, data collection for the case studies of districts that are implementing new teacher evaluation systems will not include teacher focus groups.)

Our estimates of the number of interviews and the amount of time required to conduct them are displayed in Exhibit 2.

There are no direct monetary costs to respondents for this activity. At an estimated 467 hours and an average of \$30 per labor hour, the overall cost burden for information collected through the surveys will be \$14,010.

Exhibit 2: Number of Respondents and Labor Hours Expected for Each Participating Site

	Individual Interviews Per Site (45-minute interviews)	Focus Group Interviews Per Site (75-minute focus group interviews)	Interviews at the State Level Per Site (45-minute interviews)	Total Labor Hours
LEAs with Fully Operational Systems (5 LEAs)	Up to 220 respondents/60-minute interviews (2215 hours)	49 respondents (61 hours)	Up to 4 respondents (3 hours)	430395 hours
LEAs Implementing New Systems (4 LEAs)	Up to 20 respondents/45-minute interviews (15 hours)	0	Up to 4 respondents (3 hours)	72 hours
Total Hours for All LEAs	1980 respondents (17035 hours)	245 respondents (305 hours)	36 respondents (27 hours)	502467 hours

A.13 Estimates of Annual Cost Burden to Respondents

There is no total capital or start-up cost component to these data collection activities nor is there a total operation, maintenance, or purchase cost associated with the study.

evaluation cycle in the 2010-2011 school year or earlier.

⁶ For purposes of this study, sites that are in the early stages of implementation will be in their first year of implementation and/or will not have completed a full annual evaluation cycle at the time of the planned data collection.

A.14 Estimates of Annual Cost Burden to Federal Government

The estimated cost to the Federal government is \$639,037. This total includes costs already invoiced, plus budgeted future costs charged to the government by PSA for preparation of the literature review, study design, site selection, data collection (including travel for site visits), data analysis, and reporting.

A.15 Program Changes in Burden/Cost Estimates

This request is for a new information collection so no changes apply.

A.16 Plans/Schedules for Tabulation and Publication

This study will generate two products. The first product, a literature review, will summarize existing empirical research on emerging teacher evaluation systems in the United States. State education agencies and districts will be able to use this synthesis of information to help plan and improve their evaluation systems.

The second product, the project report, will present findings from case studies of five districts that have been operating new teacher evaluation systems for at least one full school year and four districts that are implementing new evaluation systems in 2011-2012. In addition to examining cross-cutting themes that emerge as we address the study questions, this report will profile each teacher evaluation system. Finally, the case study report will locate key findings and conclusions in the larger context of empirical research discussed in the literature review.

Pending the Department's review and approval, the final draft of the literature review will be completed in ~~April~~ ~~March~~ 2012. We will collect case study data ~~April~~ ~~February~~ 1, 2012, to ~~May~~ ~~June~~ 30, 2012. Pending the Department's review and approval, the final draft of the case study report will be completed by ~~February~~ ~~March~~ 9, 2013.

A.17 Expiration Date Omission Approval

Not applicable. All data collection instruments will include the OMB data control number and data collection expiration date.

A.18 Exceptions

Not applicable. No exceptions are requested.

References

- Kane, T. (2009, October 30). *Identifying effective teaching*. Presentation made by Thomas Kane on behalf of the Bill and Melinda Gates Foundation at the Center for American Progress.
- Mathers, C., Oliva, M., & Laine, S. (2008). *Improving instruction through effective teacher evaluation: Options for states and districts*. Washington, DC: National Comprehensive Center for Teacher Quality

National Council on Teacher Quality. (2011). *2010 state teacher policy yearbook: National summary*. Washington DC: National Council on Teacher Quality.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2010). *ESEA Blueprint for Reform*, Washington, DC: author.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York: The New Teacher Project. Retrieved February 23, 2010, from http://widgeteffect.org/downloads/TheWidgetEffect_execsummary.pdf.