

# **AFT Educational Foundation Educator Evaluation for Excellence in Teaching and Learning (E3TL) Consortium Development Grant**

## Table of Contents

Competitive Priorities	1
Selection Criteria A: Need for the project and quality of the project design	2
Selection Criteria B: Strength of Research, Significance of Effect, and Magnitude of Effect	15
Selection Criteria C: Experience of the Eligible Applicant	19
Selection Criteria D: Quality of Project Evaluation	22
Selection Criteria E: Strategy and Capacity to Bring to Scale	29
Selection Criteria F: Sustainability	32
Selection Criteria G: Quality of the Management Plan and Personnel	34
Bibliography	

***Competitive Priorities: AFT Educational Foundation***  
**Educator Evaluation for Excellence in Teaching and Learning (E3TL) Consortium**  
**Development Grant**

The United States student population continues to grow in its diversity and complexity—and so must our public education system. There are more than 6.5 million students with disabilities (SWDs) in our public schools today. Many of these students are being served in general education settings. Limited English proficient (LEP) students currently make up 10 percent of public school enrollment and are one of our lowest-achieving groups of students. Teachers need support and guidance on how best to meet the needs of these students.

The AFTEF proposal, while addressed directly to the support of effective teachers, also supports the department’s efforts regarding the needs of LEP students limited English proficient and SWDs (competitive priority 7). The proposal seeks to develop and incorporate the use of professional teaching standards for general education teachers of LEP students and SWDs to judge teacher effectiveness and assist teachers in improving their practice. Including professional teaching standards for LEP students and SWDs in an overall teacher evaluation system will identify effective practices for working with diverse students in general education settings and assist teachers in successfully educating students with varied learning and linguistic needs.

Two working groups will be formed, one focusing on LEP students and the other on SWDs to address the inclusion of professional teaching standards for LEP students and SWDs who are served in the general education curriculum and taught by general education teachers. Experts from the field as well as practitioners from the participating New York and Rhode Island districts will serve on these working groups to develop both standards and performance rubrics. Teachers will also receive professional development around the newly created standards as they are incorporated into the overall teacher evaluation system.

## **The Educator Evaluation for Excellence in Teaching and Learning (E3TL) Consortium**

**Selection Criteria A. Need for the Project and Quality of the Project Design (up to 25 points)** In determining the need for the project and quality of the design of the proposed project, the Secretary considers the following factors:

In an effort to encourage teacher-developed, sustainable, innovative and collaborative reform efforts, the American Federation of Teachers Educational Foundation (AFTEF) created and launched the American Federation of Teachers (AFT) Innovation Fund in early 2009. One of the priority areas of the Fund was to address innovations in teacher evaluation that included standards for effective teaching, new criteria for tenure, and multiple measures of effectiveness including student learning outcomes. Two state federations—the New York State United Teachers (NYSUT) and the Rhode Island Federation of Teachers (RIFT)—applied for and were awarded a grant to collaboratively design “state-of the art” performance-based teacher evaluation systems in 12 urban and suburban New York and Rhode Island schools districts. They proposed to create systems that would be aligned closely to their respective states’ professional teaching standards and include multiple sources of evidence of teacher effectiveness, including evidence of student learning. These systems are based on the AFT continuous improvement framework for teacher evaluation (Appendix H). That framework is designed to measure teachers on the practices that, over time, produce desirable student outcomes and that provide teachers the opportunity to hone effective practices.

Since October, 2009, labor-management “Innovation Design Teams” from all 12 districts have been meeting monthly to examine their current evaluation standards, determine state expectations, identify gaps in the current systems and agree to a format and process for developing a new system. Each local union president and superintendent, along with union and administrative representatives, has participated fully in these meetings. The Design Teams’

work has been guided by experts such as Laura Goe from the Educational Testing Service, Charlotte Danielson, Principal of the Danielson Group and Susan Moore Johnson from Harvard University. Design Teams have also received technical assistance from the AFT. These teams will complete the design phase of their work in July 2010 and pilot their new systems during the fall of 2010-2011.

Ten of the original 12 districts across New York and Rhode Island, NYSUT, RIFT and AFTEF have formed the Educator Evaluation for Excellence in Teaching and Learning (E3TL) Consortium. This Consortium has developed an i3 project that addresses Absolute Priority 1: Innovations that Support Effective Teachers and Principals, and Competitive Priority 7: Innovations to Address the Unique Learning needs of Students with Disabilities and Limited English Proficient Students. The purpose of the consortium is to assist districts with implementation of the performance-based teacher evaluation systems designed during their original AFT Innovation grant.

**(1) The extent to which the proposed project represents an exceptional approach to the priorities the eligible applicant is seeking to meet (i.e., addresses a largely unmet need, particularly for high-need students, and is a practice, strategy or program that has not already been widely adopted).**

*Addressing Students' Unmet Needs*—Since the publication of the Coleman Report in 1966, research has increasingly focused on the relationship between teacher effects and student achievement. While the Coleman Report highlighted the influence of out-of-school factors on student achievement, research today identifies teachers as the most important in-school factor impacting growth in student achievement (Hanushek, 1992; Kane et al. 2006; Clotfelter, Ladd, & Vigdor, 2006). This is particularly true for disadvantaged students (Nye, Konstantopoulous & Hedges, 2004).

Likewise there is a growing body of econometric research focused on the influence of “effective” teachers on student achievement (Aaronson, Borrow and Sander, 2003; Gordon, Kane and Staiger, 2006; Kane, Rockoff and Staiger, 2006; Rivkin, Hanushek and Kain, 2005). So far, these effects are significant, ranging from 0.1-0.25 student-level standard deviations (Kane, Taylor, Tyler, Wooten, 2010).

Yet, while we know that “effective” teachers matter to student achievement, current methods for identifying and developing these teachers in practice are lacking. Most state mandated evaluation policies and evaluation systems in school districts do not identify effective teachers nor do they help teachers to develop and improve (Brandt et al., 2007; Ellet & Garland, 1987; Loup et al., 1996; Weisberg et al., 2009). These findings, coupled with current state and federal legislative initiatives (e.g., Race to the Top, School Improvement Grants), have compelled districts and states to revisit their teacher evaluation policies to create performance-based teacher evaluation systems that can be used to improve practice and address a largely unmet need in improving student achievement, particularly for high-needs students.

***An Innovative Approach***— First, this project is innovative in its approach to identifying, developing and improving teaching quality. Current teacher evaluation systems are inadequate at identifying, developing and improving teaching quality; these systems usually rely on a subjective principal checklist to rate teachers as “satisfactory” and “unsatisfactory.” E3TL districts will implement comprehensive, performance-based teacher evaluation systems designed not only to sort but also to improve instruction in 10 districts across Rhode Island and New York. Research shows that districts that adopt rigorous performance-based teacher evaluation systems can be used reliably to make consequential decisions, such as employment status, career-ladder status and compensation decisions (Odden, 2004) and are correlated to positive student

achievement gains (Archibald, 2007; Gallagher, 2004; Kimball et al., 2004; Harris & Sass, 2007; Holtzapple, 2003; Milanowski, 2004a and 2004b; Kane et. al, 2010; White, 2004).

Performance-based teacher evaluation systems are designed to assess the quality of teacher performance on one or more important aspects of teaching (NCCTQ, 2008, Odden 2004). Such systems: (a) are based on professional teaching standards of practice that define what teachers should know and be able to do; (b) measure teachers' knowledge and skills as they are used in practice; (c) measure multiple constructs of teaching practice (e.g., instruction, planning, assessment); (d) incorporate multiple sources of evidence (e.g., teacher developed portfolios, structured classroom observation protocols, lesson plans, student learning, videos of instructional practice and teacher work samples); (e) include assessment methods that are scored by trained and calibrated assessors who use rubrics—written scales that define levels of quality performance based on standards of practice—to make judgments of performance quality (NCCTQ, 2009, p.5); and (f) align professional development resources around the evaluation.

Second, this project is innovative because it attempts to address issues of scale and capacity. As a cross-state, multi-district consortium, the resources of districts, local and state unions can be pooled and used to leverage capacity needs. Finally, this proposal is innovative is that the consortium acts as a professional learning network where challenges and opportunities can be collaboratively addressed and lessons learned can be shared.

**(2) The extent to which the proposed project has a clear set of goals and an explicit strategy, with the goals, objectives, and outcomes to be achieved by the proposed project clearly specified and measurable and linked to the priorities the eligible applicant is seeking to meet.**

*Goals*—The overall goal of this project is to implement effective performance-based teacher evaluation systems with fidelity that (a) articulate an agreed upon vision of effective teaching, (b) accurately identify teachers on a continuum of performance and (c) provide

accurate data that can be used to help teachers improve and develop to increase student achievement. (See Table 1 for a more detailed presentation of the goals, objectives and outcomes of this proposed project.).

**Table 1**  
**Educator Evaluation for Excellence in Teaching and Learning (E3TL) Consortium**

<b>Hypothesis: Implementing a rigorous performance-based teacher evaluation system with fidelity will strengthen teaching and improve student outcomes.</b>		
<b>Goals</b>	<b>Objectives</b>	<b>Outcomes</b>
1. Educate stakeholders about the standards of practice, purposes and goals of the evaluation system and the evaluation process.	<ul style="list-style-type: none"> <li>- Create/refine training and accompanying materials to help stakeholders learn the purposes of the evaluation system and the evaluation process (e.g., evaluation scheduled, how evidence is collected and assessed, etc.).</li> <li>- Create/refine and provide training and accompanying materials to help stakeholders learn the standards—including exemplars of teaching performance as measured along a continuum of practice (e.g. from excellent to meets the standard to needs improvement to unsatisfactory)</li> </ul>	- Positive changes in stakeholder attitudes regarding the purposes and potential uses of teacher evaluation and buy-in from stakeholders.
2. Train teacher evaluators to accurately assess teaching performance and assist teachers in improving their practice.	<ul style="list-style-type: none"> <li>- Create/refine training (stimulus) materials for a certification process for teacher evaluator candidates.</li> <li>- Conduct training for teacher evaluator candidates.</li> <li>- Create/refine professional development for evaluators on how to interpret teacher evaluation data to assist teachers in improving their practice.</li> </ul>	- Increased accuracy in identifying effective practices and teachers.
3. Implement the new evaluation system with fidelity in 10	-Conduct paired observations with expert consultants and evaluator.	- An increase in the percentage of teachers meeting the standards over

	-Provide on-site training to address evaluator fidelity to observation model protocols; assessment of evaluator observation artifacts to (a) identify areas of necessary support / growth; (b) assess the degree that evidence is aligned with standards and to determine initial degrees of accuracy in assessing teacher practice; (c) establish evaluator inter-rater reliability	time. - An increase in student achievement and a closing of achievement gaps between student groups.
4. Develop standards of effective practice in the teaching of SWDs and LEP students and incorporate these standards into a performance-based teacher evaluation system.	<ul style="list-style-type: none"> <li>- Form a committee comprised of national experts, teachers and administrators from participating districts to create a set of standards and performance rubrics designed to define and measure effective teaching practices for general education teachers who teach both SWDs and LEP students in inclusionary settings.</li> <li>- Develop materials and training for stakeholders about the standards of practice for SWDs and LEP students.</li> <li>- Develop materials for certifying evaluators using standards for SWDs and LEP students.</li> </ul>	<ul style="list-style-type: none"> <li>- A set of standards, with performance rubrics for assessing teacher practice in the instruction of SWDs and LEP students in inclusionary settings.</li> <li>- Additional training materials for stakeholder education</li> <li>- Additional materials for certifying evaluators based on the standards.</li> </ul>

**Theory of Action**—The theory of action undergirding this project is that implementing performance-based teacher evaluation systems will strengthen teaching and increase student learning. Such systems are based on professional teaching standards that identify effective practices that lead to desired student outcomes. These systems, when implemented with fidelity, provide valuable information on a teacher’s strengths and weaknesses thus allowing for targeted professional development to develop and improve teachers.

*The Project Design – Strategy for Meeting the Goals*—This project will assist 10 districts (see Table 2) in implementing performance-based teacher evaluation systems that provide training and professional development based on “best practices” that improve teacher knowledge, classroom instruction, and student achievement and produce sustained change over time. E3TL districts represent a diverse cross-section of American public education. Urban and suburban districts will participate in the project (see Table 3 for key district demographics).

**Table 2**

<b>Educator Evaluation for Excellence in Teaching and Learning (E3TL) Consortium</b>	
<b>New York Districts</b>	<b>Rhode Island Districts</b>
City School District of Albany	Cranston Public Schools Network
Hempstead Public Schools	Pawtucket School Department
Marlboro Township Public Schools	Providence School District
Plattsburgh City School District	West Warwick Public Schools
North Syracuse Central School District	Woonsocket Education Department

**Table 3**

<b>Demographic Variable</b>	<b>Number/Percent Across All E3TL Districts</b>
Total Number of Schools	147
Total Number of Teacher FTEs	6,880
Average percentage of FRPL students served	51%
Average percentage of SWDs students served	19%

The project will begin with a “pilot” phase in which the evaluator for this project, the American Institutes for Research (AIR), will conduct a formative evaluation to gather and provide data for full implementation of the project. Approximately 20 schools across 10 districts in Rhode Island and New York will participate in the pilot phase. The project has three major components: (1) developing materials; (2) delivering training; and (3) implementing the system. These components are discussed below.

## **1. Developing materials**

This grant will fund the development and revision of training materials (documents, training manuals, videos of teaching practice, web-based materials) based on Charlotte Danielson's Framework for Teaching (FfT) (See Appendix H for the FfT and a sample rubric). These materials will be revised and adapted to ensure alignment with both the teaching standards in Rhode Island and New York as well as the key design features that the districts have adopted (e.g., the evaluation cycle, schedule, etc.). Expert consultants from the Danielson Group will work with the AFTEF and the participating districts to revise the training materials needed for district- and building-level teams and evaluator candidates.

Materials will be developed and/or revised (based on existing materials from the Danielson Group) in two areas: (a) stakeholder education; and (b) observation skills and coaching.

*Stakeholder education materials*--To ensure trust in the system, both teachers and administrators evaluating practice must have the same understanding of the professional teaching standards and the evaluation process, and must consistently reach similar conclusions when weighing the evidence to determine levels of performance.

*Observation skills and coaching materials*—The system is dependent on common procedures and shared understandings among the evaluators. Materials developed and revised for observation skills and coaching will address how to ensure accurate measurement of teacher practice and will include topics such as the measurement of teacher practice using criteria, effective observation techniques, the evidence to judgment cycle, and a process for providing feedback to improve practice with observation narratives and impact statements.

## **2. Delivering training**

Training will be delivered for two distinct purposes: (a) stakeholder education; and (b) observation skills and coaching. Expert consultants from the Danielson Group will work with the participating districts to deliver the training.

*Stakeholder education training*—Stakeholder awareness training will include in-depth instruction on each state’s adapted version of the Framework for Teaching and its use as a tool for teacher observation and evaluation. District design teams along with building design teams will participate in the stakeholder education training. During the pilot phase, the Danielson Group will deliver two state-level trainings (4 total) using a “Train the Trainer” model. Each training session will be designed for no more than 45 participants (a total of 180 participants). Building level teams will then conduct training at their school sites to all teachers with the support of web-based applications and video materials. Initial training will take 24 hours (3 days) with another 8 hours of follow-up, onsite support and training by the expert consultants.

*Observation skills and coach training*—The observation skills and coach training is designed to ensure accurate measurement of teacher practice using the criteria. Training of observers and evaluators typically includes in-depth training on the criteria and building understanding of how professional practice is represented in each of the standards of practice. Training on the standards of practice is followed by specific training on how to use the criteria to assess teacher practice. Observation and evaluation training includes an overview of effective observation techniques, the evidence to judgment cycle, the role of evidence in observing and assessing practice, distinguishing between evidence and opinion, influences of bias, aligning evidence with the standards of practice, making professional judgments about teacher practice,

and providing feedback to improve practice through the composition of effective observation narratives and impact statements.

Typically, training professional observers consists of a six day training schedule and one day of scoring (56 hours total). Table 3 is a sample schedule of topics covered.

**Table 3—Sample Training Schedule**

Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
Build recognition of teacher practice at all levels of performance for the standards of practice and training on the observation techniques, protocols, and timeline of the model.		Observing professional practice; gathering evidence of observations, aligning evidence, interpreting evidence and judging teacher performance; A certification assessment at the conclusion of the training.			

Follow up training and support for those who do not attain the desired level of proficiency to be certified will be provided. Approximately 60 teachers and administrators from the participating districts will be initially “certified.”

### **3. Implementing the system**

All districts, with assistance from the AFTEF, will implement a performance-based teacher evaluation system that has the core features addressed in Section A(1). Some variation may occur as districts in the consortium have developed their own instruments and procedures to assess teachers against their professional teaching standards (which have been adapted from the FFT). Variation will be studied regarding the procedures for collecting and using multiple sources of evidence and data for teacher evaluations, plans for ongoing monitoring and evaluation of the performance based teacher evaluation systems; evaluation schedules and the like.

Implementation activities that will be funded by this grant include:

*Paired observations*—In order to build a high degree of inter-rater reliability, paired observations completed by consultants and observers / evaluators will be conducted. In addition

to addressing inter-rater reliability, the objectives of the paired observations will be to further district evaluators' recognition of teacher practice at all levels of performance, and to correct any misunderstanding in the implementation of the standards of practice.

*On-site training* that will address:

- Evaluator fidelity to observation model protocols
- Assessment of evaluator observation artifacts to for the following purposes: (a) identify areas of necessary support / growth; (b) assess the degree that evidence is aligned with standards and to determine initial degrees of accuracy in assessing teacher practice; (c) establish evaluator inter-rater reliability

### **Six Criteria for High Quality Professional Development**

Both the materials developed and revised as well as the training will follow “best practices” in professional development. In addition, the teacher evaluation process implemented will focus on the work of teachers with their students, a required component of high-quality professional development. When designing professional development, the following core and structural features have been found to be important to its success (Garet et al., 2001).

#### Core Features

*Focus on Content*—Professional development that focuses on what students are expected to learn and how students learn the subject matter improves student achievement (Cohen & Hill, 2001; Garet et al., 2001; Kennedy, 1998; Carpenter, Fennema, et al., 1989, McCutchen et al. 2002).

- The materials that will be developed and revised based on the FfT will directly connect with what students are expected to learn and how students learn the subject matter. For example,

many of the domains and components focus specifically on what and how students learn (e.g., Component 3c—engaging students in learning).

*Opportunities for active learning*—Teachers benefit from opportunities to observe and be observed by expert teachers; to integrate learning into classroom practice; opportunities to review student work with others; and to reflect, discuss, and write about their learning (Garet et al., 2001; Lieberman, 1996; Loucks-Horsley et al., 1998).

- The materials will include opportunities for teachers to observe practice in person and through the use of technology. A video collection of teachers engaging in a range of practices will be utilized. The implementation of the teacher evaluation system will also provide opportunities for teachers to naturally integrate learning into their classroom.

*Coherence*—Professional development is more effective when the activities and goals involved are aligned with other initiatives designed to change instruction; when they are consistent with teachers’ personal goals for their development; and when they afford opportunities for teachers to communicate with others involved in similar professional development activities (Cohen & Hill, 1998; Garet et al., 2001; Grant, Peterson & Shojgreen-Downer, 1996; Lieberman & McLaughlin, 1992).

- During the design phase, each of the participating districts went through an alignment process to ensure that the teacher evaluation systems they were developing were aligned with other curricular and instructional initiatives. Further, the project will fund a professional learning network for educators across districts and states so that educators will have both a school-based professional learning community as well as a broader learning community.

## Structural Features

*Form*—Research suggests that professional development activities incorporated in teachers' daily school work, such as coaching, mentoring, and in-school discussion groups, provide more opportunities for active learning and encourage greater coherence of activities with teachers' and schools' larger goals and teachers' communications with others is most effective (Garet et al., 2001; Hargreaves & Fullan, 1992; Little, 1993; and Stiles, Loucks-Horsley, & Hewson, 1996.)

- This project will specifically include a formal coaching model in which evaluators (administrators and expert teachers) will assist teachers with interpreting the evaluation results and planning for improvement in practice.

*Duration of the activity*—Duration refers both to the time span of the effort and the number of hours committed to the effort. Duration appears to be supported by the form of the activity. In turn, both span and number of hours of professional development are associated with opportunities for active learning (Garet et al., 2001; Cohen & Hill, 2001; and O'Connor, 1999).

- Participants of the training will receive initially over 32 hours of professional development and evaluators would receive over 80 hours over the course of the school year. In this Train the Trainer model, all stakeholders will receive equivalent hours of training from building-level teams.

*Collective participation*—Including teachers from the same school, same department within the school or, ideally, the same grade level in the school fosters opportunities for collegial development that improves professional development in the short-term and helps sustain it over the long-term (Ball, 1996; Knapp, 1997; Talbert & McLaughlin, 1993; Elmore, 2002).

- Building level teams will be created to receive and deliver the training, and common planning time around the teaching standards, assessment of the standards and professional

development based on the evaluation data will be devoted for the purpose of fostering collegial learning.

**B. Strength of Research, Significance of Effect, and Magnitude of Effect (up to 10 points).**

**The Secretary considers the strength of the existing research evidence, including reported practice, theoretical considerations, and the significance and magnitude of any effects reported in prior research, on whether the proposed project will improve student achievement or student growth, close achievement gaps, decrease dropout rates, increase high school graduation rates, or increase college enrollment and completion rates. Eligible applicants may also demonstrate success through an intermediate variable that is strongly correlated with improving these outcomes, such as teacher or principal effectiveness. In determining the strength of the existing research evidence, the Secretary considers the following factors:**

**(1) The extent to which the eligible applicant demonstrates that there are research-based findings or reasonable hypotheses that support the proposed project, including related research in education and other sectors.**

*The proposed project is based on the hypothesis that rigorous, comprehensive performance-based teacher evaluation systems will strengthen teaching and improve student outcomes.* Educational research and similar innovative programs yielded promising results that support this hypothesis.

*Performance-based evaluation systems can help teachers improve their practice*

(Darling-Hammond & Snyder, 2000; Kimball, 2002; NCCTQ, 2008; NCTAF, 1996; Pecheone, Pigg, Chung & Souviney, 2005; Sartain, Stoeliga & Brown, 2009; Sato, Wei & Darling-Hammond, 2008;).

Historically, teacher evaluation has been criticized for its inability to provide teachers with valuable information to assist teachers to improve their practice (McLaughlin, 1990; Seafross & Enz, 1996). In comparison, performance-based teacher assessment systems evaluate actual teaching practice and provide specific information aimed at improving practice. Teachers and principals alike who experience performance-based evaluation generally report that the

assessment accurately measures teacher performance and provides information that can be used to improve teaching (Sartain, Stoelinga & Brown, 2009).

*Performance-based evaluation systems are correlated with student achievement gains* (Archibald, 2005; Gallagher, 2004; Kimball et al., 2004; Harris & Sass, 2007; Holtzapple, 2003; Milanowski (2004a and 2004b), Kane et. al, 2010; White, 2004). Research confirms that robust performance-based teacher evaluation systems are associated with student learning gains in ways that other teacher evaluation methods (e.g., principal ratings alone) are not (Nelson, 2009; Harris & Sass, 2007; Jacob, Lefgren, Harris & Sass, 2007; Little, Goe & Bell, 2009). Four studies by the highly respected Consortium for Policy Research in Education-University of Wisconsin (CPRE) positively correlate performance-based teacher evaluation scores with student achievement growth (Gallagher, 2004; Kimball et al., 2004; Milanowski, 2004; White, 2004). Furthermore, systems based on the Danielson FFT are found to be positively correlated with student achievement growth (Holtzapple, 2003; Milanowski (2004a and 2004b), Kane et. al, 2010).

*Fidelity of implementation of a performance-based evaluation system is crucial to its ability to improve teacher practice and influence student achievement* (King et al., 2010; NCCTQ, 2009; Sartain, Stoelinga & Brown, 2009). In particular, research has identified the following lessons learned about the implementation of a performance-based teacher evaluation system:

- Providing high-quality training for stakeholders around the standards of practice, purpose and goals of the evaluation system as well as the evaluation process is critical to the success of a performance-based teacher evaluation system (Mathers, Olivia, & Laine, 2008).

- Systematically training classroom observers and evaluators (e.g., intra- and inter-rater reliability) helps ensure their ability to accurately assess teacher performance (Joint Committee, 1995; Little, Goe & Bell, 2009; Mujis, 2006; NCCTQ, 2008; Shannon, 1991; Shavelson, Webb & Burstein, 1986). Conversely, lack of training leads to overall ineffective teacher evaluations (Loup et al., 1996, Wise et al., 1984).
- Systematic communication about the evaluation should take place with teachers prior to and after the evaluation process (Joint Committee, 1995; Mathers, Oliva & Laine, 2008; Wise et al., 1984). This is particularly important for promoting professional growth. There is also some evidence that training or professional development for evaluators around leading conversations about teacher observations is beneficial (Sartain, Stoelinga & Brown, 2009).

**(2) The extent to which the proposed project has been attempted previously, albeit on a limited scale or in a limited setting, with promising results that suggest that more formal and systematic study is warranted.**

*This project is based on similar programs that have yielded promising results suggesting that more formal and systematic study is warranted.*

*Cincinnati Public Schools (CPS) Teacher Evaluation System (TES) is a robust performance-based evaluation of teachers.* TES, developed with the help of AFT, uses a modified version of Charlotte Danielson’s well-accepted and research-based framework (Danielson, 2007); the same framework undergirds AFT’s *A Continuous Improvement Model For Teacher Development and Evaluation*. TES assesses teachers’ skills and responsibilities in four domains: Planning and Preparing for Student Learning, Creating an Environment for Student Learning, Teaching for Student Learning, and Professionalism. Sixteen standards on high-quality standards for teaching are then described across those domains and form the foundation for evaluation and professional development; each standard is described by a rubric. Evaluators participate in intensive professional development on using the system. Each year, all

CPS teachers are evaluated to some extent every year; about fifteen percent of teachers each year receive a more comprehensive evaluation. Evaluations are both high stakes and also are used formatively to guide teacher growth.

*Montgomery County Public School's (MCPS) Professional Growth System (PGS) uses multiple measures to evaluate teachers and guide professional development.* PGS includes a multi-stakeholder Implementation Team, a comprehensive professional development program, a Peer Assistance and Review (PAR) program, a standards-based teacher evaluation system and a career ladder. The system includes in-depth classroom observations and a significant emphasis on professional growth. In the PAR program, jointly run by the teachers' union (MCEA) and the administrators' association (MCAASP), teachers are evaluated and then recommended for continued employment, continued assistance, non-renewal or dismissal. Teachers are also provided with intensive assistance as needed.

**(3) The extent to which the eligible applicant demonstrates that, if funded, the proposed project likely will have a positive impact, as measured by the importance or magnitude of the effect, on improving student achievement or student growth, closing achievement gaps, decreasing dropout rates, increasing high school graduation rates, or increasing college enrollment and completion rates.**

*The project will have a positive impact on improving student achievement for all students, including those most at risk of failure and those subgroups who traditionally perform at lower levels on tests.* Kane et al.'s 2010 study of TES found that CPS teachers' TES performance ratings were found to be positively correlated with value-added estimates of student learning in math and reading over the three years of the evaluation; e.g., the TES system appears to be based on a valid assumption that the performance-based teacher evaluation system is correlated with an increase in student learning (Milanowski et al, 2001).

CPRE's evaluation of TGS provides evidence of positive outcomes of teacher evaluation based on multiple measures and tied to professional growth (Koppich, 2009). The rate of teacher

resignation and contract nonrenewal for teachers who have not been rated effective is higher than average districts' numbers. In addition, every student subgroup in this multicultural, multilingual, socioeconomically-diverse school system has made significant gains on the state test. The number of MCPS students scoring proficient or advanced on the Maryland State Assessment in elementary and middle school math and reading has risen over the years TGS has been in place. Also, the gap has decreased significantly between white and African-American third graders: on the 2003 Maryland state reading exam was 35 points; by 2008, it was reduced to 19 points. The gap for Hispanic students narrowed from 43 to 17 points over those years. Mathematics trends show similar progress (Childress, Doyle, and Thomas, 2009).

**Selection Criteria C: Experience of the Eligible Applicant (up to 25 points)**

**In determining the experience of the eligible applicant, the Secretary considers the following factors: (1)The past performance of the eligible applicant in implementing projects of the size and scope proposed by the eligible applicant.**

*AFTEF is eligible for an i3 grant.* This i3 grant is a partnership between AFTEF and 10 districts. AFTEF is the AFT's nonprofit 501(c)(3) foundation and conducts, sponsors and disseminates research in education and related fields with the main objective of improving and restructuring the education system. All E3TL districts meet the federal requirement for classification as LEAs.

*E3TL is strongly positioned to receive the 20 percent match required for this grant.* The AFTEF currently has a commitment for more than \$500,000 in private philanthropy dollars to support its current work in the area of focus on this grant proposal. AFTEF is part of the new i3 Foundation Registry where i3 grantees can register their grant to receive a portion of the \$500 million in funds committed for matching funds. The AFTEF, based on communications with current funders, has every expectation that those organizations will provide at least a \$1 million

(20 percent) match to federal funds when AFTEF is awarded the \$5 million it seeks through this application.

*AFTEF, the nonprofit official partner in this proposal, has an outstanding history of implementing high-quality projects similar in size and scope to the proposed E3TL.* For example, the AFT has worked since 1981 with the Toledo Board of Education and the Toledo Federation of Teachers to implement a highly successful program of teacher professional development and evaluation. The Toledo Plan includes components for new and veteran teachers. Trained teacher-evaluators, along with management, conduct evaluations based on performance standards set by the Toledo Public Schools, with help from the AFT. Many educators and researchers consider The Toledo Plan to provide important clues for the next – generation model of teacher evaluation.

The AFT Educational Research and Dissemination (ER&D) Program is a union-sponsored, research-based professional development program designed to help local unions build the capacity to deliver high-quality professional development services in collaboration with their school districts. ER&D delivers scientifically based research in a focused, sustained framework that promotes the application of research-validated concepts and strategies. Created in 1981, ER&D was recognized by the American Educational Research Association as an exemplary program that "bridges the gap between research and practice." Beginning with a single course delivered to teachers in three pilot sites, the program has now includes 12 strands available to all union constituents—teachers, paraprofessionals, and school-related personnel—in more than 200 locals across the country. The ER&D Program meets the criteria for "high quality professional development" as defined in the No Child Left Behind Act of 2001.

**(1) The extent to which an eligible applicant provides information and data demonstrating that (b) In the case of an eligible applicant that includes a nonprofit organization, the**

**nonprofit organization has significantly improved student achievement, attainment, or retention through its record of work with an LEA or schools.**

*AFTEF's work has led to improved teacher professional development, evaluation, and practice, as well as increased student achievement, attainment and retention.* For example, AFTEF's work with the Chicago Public Schools (CPS) and the Chicago Teachers' Union (CTU) on the Fresh Start Program demonstrates the AFTEF's longitudinal success in implementing projects focused on at-risk student populations that result in positive teacher and student outcomes. This project also provides further evidence of AFTEF's experience in implementing projects of the size and scope of the E3TL as well as its capacity to forge a strong LEA-level labor/management collaborative, such as the one proposed in this i3 grant.

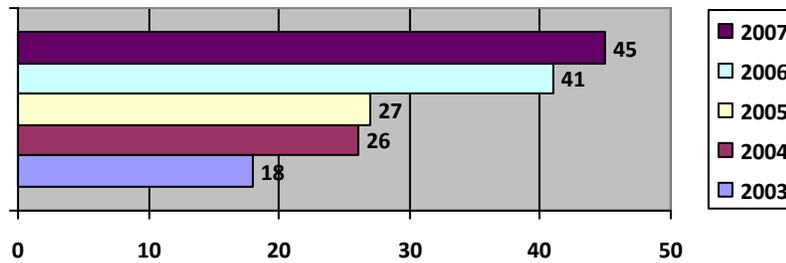
AFTEF began working with CPS and CTU in 2003, beginning with an effort for teacher-led reform in ten of the district's schools that were in danger of being reconstituted. The Partnership Initiative gave the schools an opportunity to implement research-based reform models and funded positions to support reform and provide professional development. Schools made major changes in the way they shared leadership and defined the roles of teachers, principals, union delegates, area instructional officers, and area instructional coaches.

External evaluation found that over a two-year period some schools demonstrated instructional changes and gains in achievement (Wright, Ross, & Luttrell, 2005). Based on the success of the two-year Partnership Initiative, AFTEF and CPS-CTU began implementing Fresh Start in high-poverty schools with histories of low student achievement. Fresh Start includes lessons learned from the Partnership Initiative as well as the implementation of a peer review system. The model provides strong, research-based on-going support to "graduate" schools from the program while simultaneously adding new schools.

A 2009 evaluation shows that the average percent of Fresh Start students meeting or exceeding reading and mathematics standards on the Illinois Standards Achievement Test

(ISAT) increased significantly from 2003-2007, by 27 percentage points in reading and by 23 percentage points in mathematics (see the graphs below) (Wright, Okunbor and Ross, 2009).

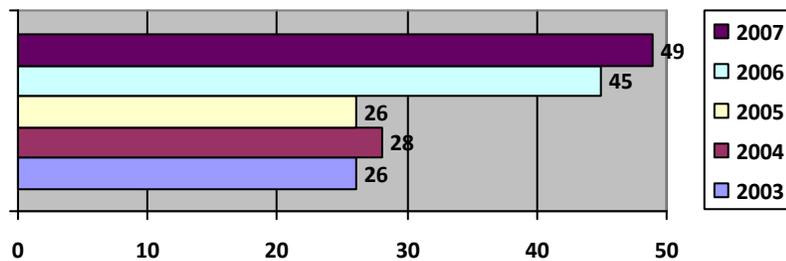
**Percent of Fresh Start Students Meeting/Exceeding ISAT Reading Standards, 2003-2007**



The percentage of Fresh Start schools making

2007 ISAT reading gains was the same as those of other students in the district, significant because Fresh Start schools represent the lowest performing CPS schools and face significant challenges in the areas of student mobility and teacher turnover.

**Percent of Fresh Start Students Meeting/Exceeding ISAT Mathematics Standards, 2003-2007**



**D. Quality of the Evaluation (15 points)** In determining the quality of the evaluation, the Secretary considers the following factors.

- (1) The extent to which the methods of evaluation are appropriate to the size and scope of the proposed project.
- (2) The extent to which the methods of evaluation will provide high-quality implementation data and performance feedback, and permit periodic assessment of progress toward achieving intended outcomes.

**(3) The extent to which the evaluation will provide sufficient information about the key elements and approach of the project to facilitate further development, replication, or testing in other settings.**

**(4) The extent to which the proposed project plan includes sufficient resources to carry out the project evaluation effectively.**

*Research Objectives and Sample*—The AFT Educational Foundation is requesting an i3 grant to support the implementation of performance-based teacher evaluation systems based on the FFT, in nearly 150 schools and for almost 7,000 teachers across two AFT affiliates (10 districts) in New York and Rhode Island. As noted earlier, through this effort, AFT seeks to build additional supports around the implementation of the framework, which currently is being piloted, used, or adapted by a number of districts (Danielson, 2009; Jerald, 2009; Weisberg et al., 2009).

To assess the outcomes of this new system, we propose an independent implementation evaluation designed and conducted by the American Institutes for Research (AIR) to collect data on implementation and outcomes. Seven research questions underlie the evaluation:

1. To what extent are the teacher evaluator and stakeholder training being implemented with fidelity to the FFT, and does the training reflect features of high-quality professional development across all districts?
2. To what extent is the new teacher evaluation system being implemented with fidelity across all study districts, reflecting standards set for selecting evidence of practice (i.e., teaching) and procedures set for collecting evidence?
3. To what extent do districts implement all features of the evaluation system (e.g., system is rigorous, transparent, and fair and uses multiple rating categories and multiple measures of effectiveness)?
4. What changes in attitudes and beliefs regarding the purposes and uses of teacher evaluation are reported by all stakeholder groups?

5. Do teacher evaluators demonstrate increased accuracy in identifying effective practice and effective teachers (e.g., more precision in using the evaluation framework and in conducting observations and conferences)?
6. Do teachers in participating districts improve their practices during the implementation of the evaluation system (e.g., increased percentage of teachers receiving higher ratings by teacher evaluators)?
7. Does student achievement in participating districts improve (e.g., change in the level of student achievement as measured by state assessments in ELA and mathematics)?

The implementation, evaluation, data collection, data analysis, and reporting activities are organized and timed to reflect the phased-in plan of the development grant. In Year 1, AIR will conduct a formative evaluation to gather data on the pilot implementation of the assessment system in up to three sample schools (e.g., one elementary, one middle, and one high school) from each participating district (we estimate that up to 10 districts will agree to participate). In each sample school, up to 20 teachers will be invited to participate in the evaluation activities throughout the implementation. The teachers participating will reflect core academic subjects that are tested by the state and additional core subjects at the high school level. The study team estimates that up to 720 teachers may be invited to participate in the data collection in the pilot year, with the goal of retaining at least 70% of teachers over the implementation years of the study. Experience in other longitudinal studies has shown that up to 40% of teachers may change teaching assignments or move out of teaching or to another school or district each year. Principal mobility is also likely to be high over the 5-year grant period. The study team will follow the surviving teachers in the analyses. However, if, as observers of evaluation systems have hypothesized, the quality of the evaluation system improves teacher effectiveness and leads to

retention of effective teachers, the mobility rate may be reduced, owing to the participation of the entire district, the buy-in by all teachers, and the support for the implementation by district and school administrators.

The formative evaluation will provide targeted information about the components and processes of the assessment system and will be a pilot for data-gathering instruments. Baseline data on student, school, and district characteristics will also be collected in Year 1. In Years 2–4, AIR will conduct a full-scale evaluation of the implementation of the evaluation system, collecting survey and interview data from this same group of participants and data on teacher and student outcomes. AIR will analyze data and report findings to AFTEF and ED. At the end of Year 5, AIR will produce a report on the findings from the implementation evaluation.

***Data Collection Methods***—The evaluation team will use a mixed method approach to collecting data. The team will develop an observation/fidelity protocol to document features of the professional development provided to teacher evaluators and building-level teams during the pilot. Interview protocols to use with evaluators, building teams, principals, affiliate representatives, and program developers and survey instruments to be administered to teachers will be designed and piloted in Year 1. AIR’s interview protocols and surveys will be informed by instruments previously used in large-scale research studies on program implementation, professional development, and school reform.

These evaluation tools will be employed in the following data collection activities each year of the study, including the pilot year, but not Year 5, which is the reporting year.

Observation of Teacher Evaluator and Stakeholder Training. AIR will observe the weeklong training of teacher evaluators and stakeholders in the FFT rubrics during the pilot year. The observation will permit us to document the training features (e.g., form, duration, collective

participation, focus on content, opportunities for active learning, coherence), dosage of training, and fidelity of implementation using a form that reflects the agenda, materials, and instructional strategies planned.

Teacher Evaluator Focus Groups. Just before the evaluator training starts and immediately after it ends, AIR will conduct structured focus groups with the teacher evaluators. Information about their backgrounds will be available through the training provider. The focus groups will glean participants' impressions of the training before they participate, to establish a baseline of understanding, and at the end of training, to establish whether changes can be documented (e.g., understanding of the framework, anticipated challenges, attitude and beliefs, perceptions of newly acquired skills and knowledge). In each implementation year, AIR will interview the teacher evaluators once, at the end of the school year, to track any changes in these perceptions.

Telephone Surveys of School Principals and Building Team Representatives. AIR will conduct two types of telephone surveys. One will gather principals' impressions of the implementation of the performance-based system at their school. The survey will seek feedback on the fidelity of implementation, impressions of the evaluators' performance, and impressions about the response of teachers to the evaluation system. A similar survey will be administered to members of each school's building-level training team. In the pilot, we will interview the principals and team members of each participating school. In each succeeding year, AIR will interview the same principals and team representatives once at the end of the school year.

Online Surveys of Teachers. AIR will develop and manage a brief online survey for teachers in the participating evaluation schools. In the pilot year, we will invite teachers to participate before the evaluation system components are implemented at their schools and then

after the evaluators have concluded the observations and conferences. The survey will gather teachers' pre- and post-implementation impressions of the intervention, attitudes and beliefs about the purposes and uses of the evaluation system, suggestions for improvement, and impressions of the feedback they received from the evaluators. AIR will use the same survey at the end of each school year to collect data on teachers' experiences with the evaluation system. As an incentive, we will give \$20 to each teacher who responds to each survey collection. To encourage full participation at the school level, the study team will administer a raffle with the names of schools twice during the grant period: at the beginning of Year 2 and in Year 3. An Apple iPad will be the raffle prize.

District Reports. AIR will create a data request form for each participating district. The district form will be used to collect data on school and student characteristics. In addition, districts will be asked, in their Memo of Understanding (MOUs), to provide annual results from their state assessments for the schools participating in the implementation study.

District leaders and AFT affiliate representatives will be interviewed and asked to report any changes to the policies and practices as a result of, or in concert with the implementation of the new evaluation system, such as identification of effective teachers, recruitment, retention, promotion, and award of incentives.

Teacher Performance Data. Data from each observation (two per teacher per year) will be gathered for the pilot and 3 implementation years, with the consent of teachers. Teacher evaluators will rate teachers on each domain of the Danielson framework.

Student Achievement Data. Student achievement data will be obtained for the students in each sample teacher's class for the pilot and the appropriate implementation years.

*Analysis and Reporting: Formative and Full Implementation Studies*--The dual purposes of the formative evaluation during the pilot are to “furnish information for guiding program improvement” (Scriven, 1991; Durlak and DuPre, 2008) and to test the data collection instruments to be used in the full-scale implementation evaluation. Each district will pilot the evaluation system in Year 1. The pilot will include all the components of the performance-based system: setting expectations for effective teaching, training evaluators, training school-based training teams, implementing observation and conference components, and developing strategic approaches for communication about the performance-based system. Using the tools described above, AIR will collect data about the training for evaluators and the implementation of the evaluation system. The observation data will be analyzed to report on fidelity of the evaluator and stakeholder training. Interview, focus group, and survey data will be analyzed to report on changes in beliefs and attitudes. Teacher ratings data will be used to report on use of the rating system. Descriptive and inferential statistical techniques will be used to calculate frequency, proportions, ranges, means, and so on. With data on school and district characteristics (e.g., size, poverty level) and teacher characteristics (e.g., number of years teaching), the study team will conduct a qualitative investigation of relationships between these characteristics and the implementation of the system, exploring variations across districts. We will analyze the data collected in teacher evaluator observations to investigate patterns of framework ratings and the precision with which the framework is used. In the pilot year, the analyses will not address outcome-focused research questions (i.e., changes in teachers meeting standards over time or increases in student achievement) because the data are treated as baseline only.

AIR will develop a brief report for AFT in spring of Year 1, as decisions are made to improve system components and processes and address variables that could affect outcomes. We will also participate in a debriefing meeting with representatives of participating schools.

The full-scale implementation evaluation will be conducted over Years 2–4 for each participating affiliate (RI and NY). As in the pilot year formative study, the implementation evaluation will collect data on each research question. The results of the implementation evaluation, reported in a comprehensive final report in Year 5, will be organized by research question and will track trends over time, variations in implementation, and variations in outcomes. We will use descriptive and inferential statistics to analyze school, district, and teacher characteristics and their relationship to implementation features and changes in attitudes and beliefs. The accumulated body of ratings from annual evaluations will yield descriptive data to analyze (1) evaluator precision in identifying effective teachers, (2) the percentage of teachers meeting standards over time, and (3) changes in student achievement over time (Kimball et al., 2004). Although improvements in teacher and student outcomes cannot be attributed with certainty to the performance assessment system, improved outcomes may provide evidence of promise for a future larger-scale effectiveness study.

**E. Strategy and Capacity to Further Develop and Bring to Scale (up to 5 points).**

**In determining the quality of the strategy and capacity to further develop and bring to scale the proposed project, the Secretary considers:(1) The number of students proposed to be reached by the proposed project and the capacity of the eligible applicant and any other partners to reach the proposed number of students during the course of the grant period.**

*This project will impact approximately 81,199 students over four years.* The project will be fully staffed by AFT teacher quality experts, NYSUT and RIFTHP staff who know the districts well and are experts in their field, and district- and school-level educators in 10 local education agencies. In addition, the project will benefit from the guidance of AFT’s Teacher Evaluation Advisory Panel (TEAP) (see Appendix H) of top experts in the field. All of these

parties have long-term established relationships with schools, districts, educators, policymakers, researchers and experts across the country and the world.

**(2) The eligible applicant’s capacity (e.g., in terms of qualified personnel, financial resources, or management capacity) to further develop and bring to scale the proposed practice, strategy, or program, or to work with others (including other partners) to ensure that the proposed practice, strategy, or program can be further developed and brought to scale, based on the findings of the proposed project.**

AFTEF has a multi-tiered strategy for bringing the project to scale. The E3TL project and its external evaluator will document the process of how to do this innovative work and identify successes and pitfalls. This information will help other districts across the country implement performance based teacher evaluation systems. This work is particularly timely because the federal policy climate is currently moving toward more comprehensive assessments of student and teacher learning, and many districts are eager to pursue this work.

AFT is in the unique position of being able to bring the work to scale via collective bargaining agreements in member districts. Based on lessons learned from the E3TL as well as future work, locals can memorialize the best practices for implementing and supporting a continuous system of teacher growth and evaluation in contract language.

Finally, the E3TL Consortium project design includes the piloting of a web-based application to support evaluator capacity. On-site training to certify evaluators can be timely and costly. This project seeks to test other methods that will still ensure accuracy but will be feasible and cost-effective as well. Data from this project will inform the use (including the validity and reliability) of online, web-based applications to certify teacher evaluators.

**(3) The feasibility of the proposed project to be replicated successfully, if positive results are obtained, in a variety of settings and with a variety of student populations. Evidence of this ability includes the availability of resources and expertise required for implementing the project with fidelity, and the proposed project’s evidence of relative ease of use or user satisfaction.**

AFTEF, with its history of working with affiliates to create teacher development and evaluation systems, has learned a great deal from these efforts, and will apply these lessons, along with those learned from the ongoing evaluation, to reach out to a broader array of school districts. The experience of working with a wide range of districts from this project will make it easier to replicate the model in districts of many different sizes and locations. In addition, knowledge gained from the external evaluation will provide recommendations on replicating the model.

The participating districts represent a good cross-section of the American public education system, which will aid in generalizing how effectively the system could be implemented in other districts. Some of the sites are large urban school systems that serve a large portion of poor and minority students and a few are small to mid-size urban and suburban systems that have relatively low poverty and few minority students.

Researchers have found that education reforms may not produce consistent result, often due to a lack of “fidelity of implementation” – meaning the strategies of a particular design were not delivered and implemented accurately and consistently (Berends, Bodilly, & Nataraj Kirby, 2002). Our evaluator will examine fidelity of implementation and make recommendations how to assure fidelity of implementation as the project scales up.

**The eligible applicant’s estimate of the cost of the proposed project, which includes the start-up and operating costs per student per year (including indirect costs) for reaching the total number of students proposed to be served by the project. The eligible applicant must include an estimate of the costs for the eligible applicant or others (including other partners) to reach 100,000, 250,000, and 500,000 students.**

With an estimated cost of approximately \$1.5 million per year for implementation of the E3TL project within the participating districts, the average cost per student per year is \$18.50. It is estimated that it would cost \$1.85 million per year to reach 100,000 students, \$4.625 million to

impact 250,000 students and \$9.25 million to reach 500,000 students. That figure should be reduced as the districts build capacity to further implementation efforts. Economies of scale also would likely bring the costs down significantly as numbers of students served rises. In addition, LEAs could redirect their current professional development dollars toward the system, so it is not an add-on cost.

**(5) The mechanisms the eligible applicant will use to broadly disseminate information on its project so as to support further development or replication.**

To share the lessons learned with the broader education community, AFTEF, along with NYSUT and RIFTHP plan to present webinars, seminars and reports. The AFT will continue to use its existing Teacher Evaluation Community web portal (<http://teacherevaluation.aft.org>) to disseminate information on the project. Ongoing updates and evaluation of E3TL will be posted, as well as blogs by various project stakeholders.

**Selection Criteria F: Sustainability (up to 10 points) In determining the adequacy of resources for the proposed project, the Secretary considers the following factors:(1) The extent to which the eligible applicant demonstrates that it has the resources, as well as the support from stakeholders (e.g., State educational agencies, teachers' union) to operate the project beyond the length of the Development grant.**

The AFTEF has the resources and stakeholder support to operate and sustain the E3TL Consortium beyond the length of the Development grant. The AFT has approximately 360 employees and annual budget of roughly \$172 million and has a healthy financial history that extends back to its founding in 1916. Implementing the E3TL in support of *A Continuous Improvement Model For Teacher Development and Evaluation* is a top union priority. Over the past few years, AFT has invested approximately \$775,000 to support an innovative teacher evaluation framework, both from its own funds and from major foundation donors including the Bill and Melinda Gates Foundation and the Carnegie Foundation.

The AFT represents nearly 3,000 local education agencies nationwide, 43 state educational agencies and more than 1.4 million members. The AFT has well-developed and

institutional structures and close working relationships with LEAs, many of which serve large populations of the most at-risk students in the country. These working relationships put AFT in a strategic position to impact hundreds of thousands of students, particularly those who have the greatest need for additional educational supports, including students eligible to receive free and reduced-price lunches, English language learners and those at risk of dropping out of school. Those relationships also will make it possible to support this innovative work beyond the length of the timeframe of this proposal.

AFTEF has letters of support from the 10 participating E3TL districts (See Appendix D). Those affiliates and their district partners that are not currently part of the E3TL Consortium will be eager to participate in a development and scale-up following completion of this project. In fact, if we can raise additional funds, we will include more district/union partnerships in the proposed work. Priority will be given to districts who meet the following criteria: the capacity to participate in the work of the project, a district/union formal agreement to participate, and a district wide student population where at least 40 percent are eligible to receive free and reduced-price lunches (FRPL) and at least 15 percent are ELLs.

**(2) The potential and planning for the incorporation of project purposes, activities, or benefits into the ongoing work of the eligible applicant and any other partners at the end of the Development grant.**

Improving teacher evaluation systems to support improve student learning is a top AFT priority. Randi Weingarten, the president of the AFT, has publicly announced the union's commitment to supporting the research and development of teacher evaluation systems that adhere to the standards previously mentioned. The strategies, products and lessons gleaned from the E3TL Consortium under this grant will be incorporated into the way AFT delivers professional development and technical assistance and material will be disseminated widely in

AFT's many widely read publications, and among other AFT communication and assistance tools.

Participating LEAs already have track records that highlight their commitment to improved teacher evaluation systems; all districts involved will be deeply invested in continuing to doing so by the project's end. In addition, the current policy climate will aid in these efforts. New and emerging federal and state mandates are making this work central to districts and their union partners.

**Selection Criteria G: Quality of Management Plan and Personnel (up to 10 points)**  
**In determining the quality of the management plan and personnel for the proposed project, the Secretary considers: (1) The adequacy of the management plan to achieve the objectives of the proposed project on time and within budget, including clearly defined responsibilities, timelines, and milestones for accomplishing project tasks.**

The E3TL Consortium is a partnership among four distinct entities. The AFTEF has responsibility to assure that all partnerships get the technical assistance they requires and that all tasks are accomplished in a timely, cost effective manner. The Rhode Island and New York state and local union leaders and their state and district administrative counterparts will develop and implement the evaluation system. The Danielson Group will assist the union/district partnerships in developing materials and training and the American Institutes of Research will evaluate the project. The AFTEF, NYSUT and RITHP have been working together for more than a year to implement comprehensive performance-based teacher development and evaluation systems in their respective states. The Danielson Group has an impressive track record of more than 20 years of work on performance-based teacher evaluations, particularly in regard to assessing instructional practice. The AIR has a long history of working with state and local districts to improve teaching and learning and, in particular, is highly skilled at evaluating professional development programs.

The E3TL Consortium has developed a comprehensive project management plan that details partners' responsibilities and milestones for accomplishing the project objectives on time and within budget. Angela Minnici, Associate Director at AFT, is the Principal Investigator and Project Manager and will be responsible for directing and managing the work. She will be supported by a Teacher Evaluation Advisory Group (TEAG), who will review plans and provide technical assistance as necessary to the affiliates. Charlotte Danielson and her colleagues will provide training and assistance with stakeholder awareness, evaluation skills and coaching in the states and districts. Colleen Callahan, Director of Professional Issues for RIFT, will be responsible for implementing the project activities in Rhode Island and Larry Waite, Manager of Educational Services for NYSUT, for doing the same in New York. They will oversee the work of the state and district development committees and will have a project coordinator assigned to each district to assure that the development of materials, training and the implementation of the system occurs in a timely manner and to address issues of implementation as they arise in the field. Meredith Ludwig will be responsible for overseeing AIR's formative and summative evaluations. The senior management team is comprised of Minnici, Callahan, Waite, Danielson and Ludwig.

### Timelines

There are eight major tasks to be accomplished in Year One: 1. The training materials for stakeholders and evaluators must be developed (union/district partners and Danielson Group); 2. the pilot schools must be selected (partnership and AIR); 3. Initial training of a stakeholders group must be undertaken (union/district partners and Danielson Group); 4. Evaluators must be trained and certified (union/district partners and Danielson Group); 5. The teacher evaluation must be implemented in pilot schools (union/district partnership); 6. Protocols for data collection

must be developed (AIR) and 6. Baseline data must be collected (AIR); 7. Materials must be refined as a result of pilot (union/district partners and Danielson Group); and 8. A committee must be formed to develop standards for assessing teacher effectiveness in dealing with LEP students and SWDs in mainstream classes (AFTEF, union/district partners and ELL, LEP and SWD education experts).

There are 6 major tasks to be accomplished in Year Two: 1. Full scale implementation of the evaluation systems in all districts (union/district partnerships) 2. Standards for effective instruction of LEP students and SWDs completed (union/district partners and national experts). 3. Materials developed for stakeholder and evaluator training on LEP students and SWD standards (union/district partners and national experts and Danielson Group). 4. Danielson Group and District partnerships refine materials based on full scale implementation experiences. 5. AIR continues data collection. 6. A conference on interim findings is held with all partners (AFTEF, Danielson Group, TEAG, union/district partnerships and AIR).

Year Three continues refining standards and materials based on the findings from Year Two. The LEP students and SWDs standards are implemented in original pilot schools. Materials are refined as a result of the pilot. AIR continues data collection, including student outcome data.

Year Four includes a full-scale implementation in all schools of the evaluation process including the standards for effectiveness in the instruction of LEP standards and SWDs. AIR continues to collect student outcome and implementation data. AFTEF, with assistance from AIR prepares a final report on findings of the study and recommendations for future action, including scale-up to additional districts/state-wide.

### Innovative Milestones

The E3TL Collaborative has clear and explicit milestones: 1. The development of original training materials for stakeholders and evaluators; 2. The creation of an ELL and SWD standards committee; 3. The collection of baseline data; 4. The pilot of training materials; 5. The revision of training materials; 6. Full-scale implementation of original training materials; 7. The creation of LEP students and SWDs teacher standards; 8. The development of training materials for stakeholders and evaluators on LEP students and SWDs standards; 9. The pilot test of those materials and their revision as a result; 10. Annual collection of outcome and implementation data; 11. Full scale implementation of evaluation system including the LEP students and SWDs component; 12 Annual and final project reports.

### Communication

The Principal Investigator will attend initial trainings and all joint RI and NY meetings. She will also visit all district sites at least once each year. To facilitate communication among the project entities the Project Investigator will hold monthly teleconference meetings with the principals—AIR, the Danielson Group, Colleen Callahan and Larry Waite—to assure the project is on schedule and to identify potential problems and develop solutions for them. Minutes of those meetings will be shared with the group to underscore activities and responsibilities to be accomplished during next month. Project principals will also submit quarterly progress reports. A similar structure will be developed by Callahan and Waite for monitoring activities, progress and potential problems in their district sites.

**(2) The qualifications, including relevant training and experience of the project director and key project personnel, especially in managing projects of the size and scope of the proposed project.**

**E3TL AFTEF Staff:**

**Principal Investigator and Project Manager: Dr. Angela Minnici**, will direct and manage the project. She is Associate Director of the AFT Educational Issues Department, and in that capacity currently provides support for the AFTEF Innovation Fund grantees in New York and Rhode Island. Prior to joining the AFT, Dr. Minnici was a senior researcher at the Center for Education Policy where she was principal investigator for two large-scale, multi-state and district studies that focused on the impact of federal and state legislation on teaching and learning. These studies were comparable in size and scope to the current E3TL efforts.

**E3TL Project Advisor: Dr. Joan Baratz Snowden** will serve as an advisor for the E3TL management team.

**E3TL Project Coordinator for New York: Melanie Hobbs** provides technical assistance to AFT locals on teacher evaluation and professional development and will coordinate project activities for the New York districts.

**E3TL Project Coordinator for Rhode Island: Justin Stone** provides technical assistance to AFT locals on teacher evaluation and implementing pay for performance systems and will coordinate project activities for the Rhode Island districts.

**AFT Teacher Expert Advisory Group (TEAG):** The project will also be guided by an advisory panel of experts in the field. See Appendix H for members and their bios.

**E3TL State Project Manager for New York: Larry Waite** will be responsible for overall management of project activities at the state level and will serve as the liaison between the local partnerships and the AFTEF. In this capacity, he will be responsible for communicating with the union/district partnerships, NYSUT, RIFTHP and AFT leaders and other project partners, coordinating all state partnership meetings, coordinating the work of this project with supports which will be available through other NYSUT Departments such as Field

and Legal Services and Communication and Media, monitor completion of projects activities in relation to prescribed time lines, assist in coordination of project evaluation activities within New York State, as needed, and submit all project program and fiscal reports pursuant to prescribed time lines.

**E3TL State Project Manager for Rhode Island: Colleen Callahan, Ed.D.**, will be responsible for overall management of project activities at the state level and will serve as the liaison between the local partnerships and the AFTEF. In this capacity she will be responsible for communicating with the union/district partnerships, RIFTHP, NYSUT and AFT leaders and other project partners, coordinating all meetings and other project related activities, monitoring completion of project activities in relation to prescribed timelines and agreements, assist in coordination of project evaluation activities within Rhode Island and submit all program and fiscal reports pursuant to prescribed timelines.

### **E3TL External Evaluators**

**Meredith Ludwig, Ed.D.**, will be the Principal Investigator of the evaluation. Dr. Ludwig has worked at the American Institutes for Research since 1996, focusing on the reform of teacher preparation in schools and departments of teacher education, as well as on the professional development (PD) needs of teachers. Among her many projects there, she directed two multi-year evaluations focused on teacher preparation issues for the U.S. Department of Education (Title II Partnership Program and Transition to Teaching Program (TTT)). She has been involved for many years in efforts focused on improving data collection and reporting on teacher quality and teacher effectiveness.

**Charles Storey, JD**, will be the evaluation's Project Director. Mr. Storey is a Principal Research Analyst who manages projects and conducts research in the areas of postsecondary

education and access to postsecondary education. He is the Project Director for the National Evaluation of the Alliances for Graduate Education and the Professoriate (AGEP) program for the National Science Foundation (NSF) and for the AGEP-Tracking Student Registry (AGEP-TSR) project. A former Program Officer in the U.S. Department of Education's Fund for the Improvement of Postsecondary Education (FIPSE), he was recently the Project Director for the Evaluation of the High School Grants Initiative of the Bill & Melinda Gates Foundation and recently completed work on a project identifying promising practices within the U.S. Department of Education's TRIO programs.

**Mengli Song**, Ph.D., will advise the project on all statistical analyses of the survey data. A Research Scientist at the American Institutes for Research (AIR), Dr. Song specializes in educational policy and politics, program evaluation, social network analysis, and quantitative research methods. She has been the task leader for the evaluation of the Bill & Melinda Gates Foundation's National School District and Network Program., and participated in the design of a longitudinal study of the program, and was the primary author of the data analysis plan for study.

**Dr. Terry Salinger** will serve as a Senior Advisor to the Evaluation. Dr. Salinger is a Managing Director and Chief Scientist for Literacy Research at AIR. She is currently the principal investigator for an IES-NCEE study of the extent to which pre-service teacher training programs include information about scientific research in beginning reading instruction and of the effectiveness of explicit instruction in literacy for adult ESL students.

### **The Danielson Group**

**Charlotte Danielson** will serve as a consultant for the overall project. She is a former economist and an educational consultant based in Princeton, New Jersey. She has taught at all levels, from kindergarten through college, and has worked as an administrator, a curriculum

director, and a staff developer. In her consulting work, Ms Danielson has specialized in aspects of teacher quality and evaluation, curriculum planning, performance assessment, and professional development.

**Albert “Duffy” Miller** will serve as project coordinator and lead trainer for the Danielson Group. He has 12 years experience as a classroom teacher and 16 years experience as a high school principal. He is the co-founder of Teaching Learning Solutions, and has since been working with school districts nationally supporting their improvement efforts and implementation of the Danielson *Framework for Teaching*. He is one of the charter members of the Danielson Group, and has worked on the development of programs relative to the use of the FfT with Educational Testing Service and Charlotte Danielson. He and Bernadette Cleland jointly proposed, supervised, and managed all aspects of the Danielson Group’s work with Prince George’s County Public School’s FIRST initiative, a project that implemented the use of the FfT in the Prince George’s County evaluation model.

**Bernadette Cleland** will serve as a material developer and trainer on this project. She has 17 years experience as a classroom teacher and 10 years as a middle and high school principal. She is a co-founder of Teaching Learning Solutions and has been working with school districts nationally supporting their improvement efforts and implementation of the Danielson *Framework for Teaching*. She has taught graduate level courses on supervision and evaluation, curriculum and instructional practices, and is currently one of two expert trainers of observers for the Understanding Teacher Quality research jointly supported by the Gates Foundation and Educational Testing Service.

## Bibliography

- Aaronson, D., Barrow, L. & Sander, W. (2003). Teachers and student achievement in the Chicago Public Schools. Working paper No. WP-2002-28. Chicago, IL: Federal Reserve Bank of Chicago.
- American Federation of Teachers. (1998) Survey of State Policies on Teacher Evaluation. Washington, DC: Educational Issues Policy Brief, AFT.
- American Federation of Teachers (2010). A Continuous Improvement Model for Teacher Development and Evaluation. Washington: DC: AFT.
- Archibald, S. (2007). How well do standards-based teacher evaluation scores identify high-quality teachers? A multilevel, longitudinal analysis of one district. Unpublished Ph.D. Dissertation, University of Wisconsin-Madison.
- Ball, D. L. (1996). Teacher learning and the mathematics reforms: What we think we know and what we need to learn. *Phi Delta Kappan*, 77(7), 500-08.
- Blunk, M. (2007, April). *The QMI: Results from validation and scale-building*. Paper presented at the annual meeting of the American Research Association, Chicago, IL.
- Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., and Hess, J. (2007). Examining district guidance to schools on teacher evaluation policies in the Midwest Region (Issues & Answers Report, REL 2007–No. 030). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest.
- Childress, S., D. Doyle & D. Thomas. (2009). *Leading for Equity: The Pursuit of Excellence in Montgomery County Public Schools*, Cambridge, MA: Harvard Education Press.
- Clotfelter, C.T., Ladd, H. F. & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, Madison, WI: University of Wisconsin Press, 41(4).
- Cohen, D. K., & Hill, H. C. (1998). Instructional policy and classroom performance: The mathematics reform in California (RR-39). Philadelphia: Consortium for Policy Research in Education.
- Cohen, D. K., & Hill, H. C. (2001). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record*, 102(2), 294-343.
- Cronin, L. & Capie, W. (1986, April). The influence of daily variation in teacher performance on the reliability and validity of assessment data. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

- Danielson, C. (2007). *Enhancing Professional Practice: A Framework for Teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. (2009). Teacher evaluation—Performance frameworks. In Hershberg, T., & Robertson-Kraft, C. (Eds.), *A grand bargain for education reform: New rewards and supports for new accountability* (pp. 59–73). Boston: Harvard Education Press.
- Darling-Hammond, L. & Snyder, J.(2000). Authentic assessment of teaching in context. *Teaching and Teacher Education*, 16(5-6), 523-545.
- Durlak, J.A. & DuPre, E.P. (2008) Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327 –350.
- Elmore, R. (2003). Bridging the gap between standards and achievement: The imperative for professional development in education. [Online.] Available: [http://www.ashankerinst.org/Downloads/Bridging\\_Gap.pdf](http://www.ashankerinst.org/Downloads/Bridging_Gap.pdf).
- Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Education Research Journal*, 38(4), 915-945.
- Grant, S. G., Peterson, P. L., & Shojgreen-Downer, A. (1996). Learning to teach mathematics in the context of systemic reform. *American Educational Research Journal*, 33(2), 502-541.
- Gordon, R., Kane, T. J. & Staiger, D. O. (2006). Identifying effective teachers using performance on the job. Hamilton Project Discussion Paper. Washington, DC: Brookings.
- Hanushek, E. A. (1992). The trade-off between child quantity and quality. *The Journal of Political Economy*, 100, 84–117.
- Hargreaves, A. and Fullan, M. G. (1992). *Understanding teacher development*. London: Cassell.
- Harris, D. & T. Sass. (2006). *Value-added Models and the Measurement of Teacher Quality*. Florida State University: Department of Educational Leadership and Policy Studies.
- Holtzapple, E. (2003). Criterion-related validity evidence for a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, 7(3), 207-219.
- Jerald, C. (2009). *Aligned by design*. Washington, DC: Center for American Progress.
- Joint Committee on Standards for Educational Evaluation. (1995). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.
- Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2010). Identifying effective classroom

- practices using student achievement data. NBER Working Paper No. 15803.
- Kane, T. J., Rockoff, J. E., & Staiger, D. (2006). What does certification tell us about teacher effectiveness? Evidence from New York City. NBER Working Paper Series, Vol. w12155.
- Kimball, S.M., White, B., Milanowski, A.T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education*, 79(4), 54–78.
- Kimball, S. M. (2002). Newport News teacher performance assessment system: A case study. Consortium of Policy Research in Education- University of Wisconsin. Retrieved April 12, 2010 from <http://cpre.wceruw.org/papers/Newport%20News%20TE%204-02.pdf>
- King, J. K., Malen, B., Baumann, P., Chen, E., Dougherty, A., Hyde, L., Jackson, C., Jacobson, R., & McKithen, C. (2010). Promises and perils of educator incentives: A formative evaluation of the Teacher Incentive (TIF) program in the Prince George’s County (MD) Public School System. Paper prepared for the 2010 Annual Meeting of the American Education Finance Association, Richmond VA.
- Knapp, M. S. (1997). Between systemic reforms and the mathematics and science classroom: The dynamics of innovation, implementation, and professional learning. *Review of Educational Research*, 67(2), 227-266.
- Koppich, J.E. (2009). Strategic management of human capital—Montgomery County Public Schools—SMHC in operation. Madison, WI: Consortium for Policy Research in Education (CPRE).
- Lieberman, A., & McLaughlin, M. W. (1992). Networks for educational change: Powerful and problematic. *Phi Delta Kappan*, 73, 673-677.
- Little, J. W. (1993). Teachers’ professional development in a climate of educational reform. *Educational Reevaluation and Policy Analysis*, 15(2), 129-151.
- Little, O., Goe, L., & Bell, C. (2009). A practical guide to evaluating teacher effectiveness. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from <http://www.tqsource.org/publications/practicalGuide.pdf>.
- Loup, K. S., Garland, J.S., Ellett, C. D. & Rugutt, J. K. (1996). Ten years later: Findings from a replication of a study of teacher evaluation practices in our 100 largest districts. *Journal of Personnel Evaluation in Education*, 10, 203-226.
- 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.

- McCutchen, D., Abbott, R. D., Green, L. B., Beretvas, S. N., Cox, S., Potter, N. S., Quiroga, T., & Gray, A. (2002). Beginning literacy: Links among teacher knowledge, teacher practice, and student learning. *Journal of Learning Disabilities*, 35, 69-86.
- McLaughlin, M.W. (1990). Embracing contraries: Implementing and sustaining teacher evaluation. In J. Millman & L. Darling-Hammond (Eds.) *The new handbook of teacher evaluation*. Newbury Park, CA: Sage Publications.
- Milanowski, A., Kellor, E., Odden, A., Heneman, H.G. III., White, B., Allen, James, & Mack, Kimberly. (2001, July) Final Report on the Evaluation of the 2000-2001 Implementation of the Cincinnati Federation of Teachers/Cincinnati Public Schools Teacher Evaluation System (Working Paper TC-01-3). Madison: University of Wisconsin, Wisconsin Center for Education Research, Consortium for Policy Research in Education.
- Milanowski, A. (2004a). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education*, 79(4), 33-53.
- Milanowski, A. (2004b). Relationships among dimension scores of standards-based teacher evaluation systems, and the stability of evaluation scores—student achievement relationships over time. *Consortium of Policy Research in Education- University of Wisconsin Working Paper Series*, No TC-04-02, April 2004.
- Mujis, D. (2006). Measuring teacher effectiveness: Some methodological reflections. *Educational Research and Evaluation*, 12(1), 53-74.
- National Commission on Teaching and America's Future (NCTAF) (1996). *What matters most: Teaching for America's future*. New York: Author.
- National Comprehensive Center for Teacher Quality (NCCTQ) (2009). Using performance-based assessment to identify and support high quality teachers. Washington, DC: Author.
- Nelson, H. (2009). *The relationship between teacher evaluations and value-added assessments of teachers*. Unpublished working paper. Washington, DC: American Federation of Teachers.
- Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237–257.
- Odden, A. (2004). Lessons learned about standards-based teacher evaluation systems. *Peabody Journal of Education*, 79(4), 126-137.
- O'Connor, R. E. (1999). Teachers learning Ladders to Literacy. *Learning Disabilities Research and Practice*, 14, 203-214.
- Pechone, R., Pigg, M. J., Chung, R. R., Souviney, R. J. (2005). Performance assessment and

- Electronic portfolios: Their effect on teacher learning. *The Clearing House*, 78(4).
- Rivkin, S., Hanushek, E. & Kain, J. (2005). Teachers, schools and academic Achievement. *Econometrica*, 73(2), 417-458.
- Sartain, L., Stoelinga, S. R. & Brown, E. (2009). Evaluation of the Excellence in Teaching pilot: Year 1 report to the Joyce Foundation. Chicago: The Consortium of on Chicago School Research at the University of Chicago.
- Sato, M., Wei, R. C. & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board certification. *American Educational Research Journal*, 45(3), 669-700.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Thousand Oaks, CA: Sage.
- Searfross, L. & Enz, B. J. (1996). Can teacher evaluation reflect holistic instruction? *Educational Leadership*, 53(6), 38-41.
- Shavelson, R. J., Webb, N. M., & Burstein, L. (1986). Measurement of teaching. In M.C. Wittrock (Ed.) *Handbook of research on teaching*, pp. 50-91. New York: Macmillan.
- Southard, M. & L.M. Dean. (2004). 2004 Evaluation Report Thinking Mathematics (TM): Title I and V. Leon County, FL: Leon County Schools.
- Southard, M. & L.M. Dean. (2007). Comparison Study Thinking Mathematics (TM): Title I and V. Leon County, FL: Leon County Schools.
- Stiles, K., Loucks-Horsley, S., & Hewson, P. (1996). Principles of effective professional development for mathematics and science education: A synthesis of standards (NISE Brief, Vol. 1). Madison, WI: National Institutes for Science Education.
- Talbert, J. E., & McLaughlin, M. W. (1993). Understanding teaching in context. In D. K. Cohen, M. W. McLaughlin, & J. E. Talbert (Eds.), *Teaching for understanding: Challenges for policy and practice* (pp. 167-206). San Francisco: Jossey-Bass, Inc.
- Toch, T. & Rothman, R. (2008, January). Rush to judgment: Teacher evaluation in public education. Washington, DC: Education Sector.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY: The New Teacher Project.
- Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Case studies for teacher evaluation: A study of effective practices*. Santa Monica, CA: RAND Corporation. Retrieved from <http://www.rand.org/pubs/notes/2007/N2133.pdf>.

Wright, P.M, E. Okunbor & S. Ross. (2009). Evaluation of the CPS-CTU Fresh Start Program: 2007-2008. Memphis: Center for Research in Educational Policy.

Wright, P., S. Ross, & A. Luttrell. (2005). Evaluation of the CPS-CTU Partnership School Program: 2004-2005. Memphis, TN: The University of Memphis, Center for Research in Educational Policy.