The School District of Philadelphia

The Office of Research and Evaluation

Race to the Top

The Educator Effectiveness Project

2013

The School District of Philadelphia
Race to the Top
Year 1 Evaluation Report

Prepared for:
The Office of Human Resources

Naomi Wyatt
Chief Human Resources Officer

Brett Shiel
Deputy, Teacher Effectiveness

Tennille Peeler
Program Director, Race to the Top

By:
The Office of Research and Evaluation

Katherine Stratos, M.S.
Senior Research Associate

Tonya Wolford, M.Ed., Ph.D
Deputy

August 2013

Table of Contents

List of Tables and Figures

Executive Summary

Background: In December 2011, Pennsylvania was awarded $41,326,299 under the federally-funded Race to the Top (RTTT) grant program. The primary objective of Pennsylvania's Race to the Top Local Education Agency (LEA) Grant program is for participating local entities to adopt and implement Pennsylvania's Educator Effectiveness instrument and to use the evaluation process and results to inform local decisions regarding professional development and staff retention in support of student achievement.

Methods: The evaluation of RTTT consists of two major components: an implementation study (formative) and an impact evaluation (summative). Data sources include program documents, observations, focus groups and feedback surveys.

Findings: Major components of the District's RTTT program fall into three categories:

1. **State-mandated activities** – program components that are mandated by the state with state-established goals, including:
    - Training and professional development on the educator effectiveness system and tools.
2. **Required District-response activities** – program components that are necessary for the District to accomplish in order to implement the new educator effectiveness system, but that have not been explicitly defined by the state as grant requirements, including:
    - Implementation of a technology solution;
    - Establishing student-teacher linkages for Pennsylvania Value-Added Assessment System (PVAAS) teacher specific growth data; and
    - Development of elective data-student learning objectives.
3. **Supplemental District activities** – program components conceived by the District that are designed to strengthen and enhance implementation of the new educator effectiveness system, including:
    - Instructional coaching; and
    - Inter-rater reliability certification.

Professional Development and Training: The goals established by the State for training 25 percent of central office administrators and 50 percent of principals on the new educator effectiveness system were met, but the goals for training 50 percent of teachers and 25 percent of specialists on the system were not met, due to scheduling challenges as well as challenges associated with staff layoffs. However, program staff plan to have all teachers and specialists trained early in the subsequent year. The goal to pilot the observation process on 10 percent of teachers was not met.

Technology Solution: A technology system was purchased that will automate the educator evaluation process and streamline educator ratings. The Pearson Educator Development Suite (EDS) will allow observers to capture observation data electronically, aggregate collected and external data, track progress in line with educator frameworks and connect educators to professional development content through an integrated video library.

Training for principals is currently underway, and the tool is expected to go live at the beginning of the 2013-2014 school year.

PVAAS Teacher Specific Growth Data: As part of the new educator effectiveness system, teachers will now be held accountable for the achievement of their specific students. The State has used RTTT funding to enhance its data systems so that it is able to link teachers to their respective students. This system was piloted in Spring 2013. Principals and teachers from three School District of Philadelphia (SDP) schools participated in the pilot.

Elective Data/Student Learning Objectives: LEAs will be responsible for the development of Student Learning Objectives (SLOs) to measure student progress in non-tested grades and subjects. Pennsylvania Department of Education (PDE) has created a rubric that outlines general requirements for developing SLOs and has provided training to districts, which RTTT representatives from SDP have attended. Elective Data/SLOs will go into effect as part of teacher evaluation in 2014-2015. As of the end of Year 1, planning for the development of the SLOs had not formally begun.

Instructional Coaching: Five instructional coaches were hired to serve as liaisons between teachers, administrators, schools and central administration, and to provide support through an array of activities that are designed to build leadership and improve teacher instructional capacity and student learning. Coaching support is planned to begin in SY 2013-2014.

Inter-rater Reliability Certification: Initially, inter-rater reliability certification through Teachscape for all observers was articulated as a requirement of the State, however that requirement was soon made obsolete. The District plans to work with the State to receive Teachscape licenses in order to certify observers, though this endeavor is in the very early stages.

Discussion and Recommendations: Throughout Year 1 of implementation, the District has been successful in ensuring that principals and school leadership teams have been trained on the upcoming changes to state-mandated teacher, principal and specialist evaluation that will go into effect in the 2013-2014 school year. The District also has a well-established plan to train all teachers on the same materials at the start of the 2013-2014 school year, even though the goal to train 50% teachers in Year 1 was not met.

The most significant implementation challenges during Year 1 were the result of unclear or shifting leadership in charge of critical decisions associated with changes to educator effectiveness practices introduced as a result of Act 82. Many of the activities associated with the RTTT grant are inextricably linked with local policy decisions at the District level, and throughout Year 1 of RTTT implementation, the individuals responsible for those policy decisions were not the same individuals responsible for implementing RTTT program activities, leading to a disconnect between programmatic decisions and larger District decisions. Ongoing collective bargaining agreement negotiations have also made it difficult for the program office to anticipate or establish changes to teacher evaluation practices.

In August 2013, the District hired a Deputy in charge of teacher effectiveness as well as a new Chief of Human Resources, and those individuals should work closely with both the RTTT staff and District leadership to align long-term District strategies with shorter-term programmatic activities. Because most of the grant requirements were established by the State for Year 1, and speak to only the professional development and training components of the grant, it is important that the educator effectiveness and RTTT teams work with Office of Research and Evaluation (ORE) staff to establish local programmatic goals for Years 2 and 3 of implementation, as well as long-term anticipated impact that is clearly defined and measurable. This should include further defining the program design and activities for the next two years of the grant, such as activities related to coaching and inter-rater reliability.

The RTTT and educator effectiveness teams must continue to work closely with Information Systems, as well as with the State's PVAAS resources, to implement the PVAAS Teacher Specific Growth measures and roster verification process with the highest level of accuracy and fidelity.

It is important that the RTTT program team and the educator effectiveness team also work closely with District leadership and appropriate Curriculum, Assessment, and Technology staff to begin defining a process for establishing Student Learning Objectives that align with the District's vision, mission and Action Plan.

Through the course of implementing these new initiatives and changes, communication with major stakeholders, including appropriate central office departments, principals, teachers, union partners, and parents should be clear, frequent, and easily accessible through a variety of channels.

Background and Introduction

On July 2, 2009, the United States Secretary of Education, Arne Duncan, addressed the National Education Association with the last in a series of speeches about the four core reforms embodied in the *American Reinvestment and Recovery Act* leading up the release of $5 billion in competitive grants. This fourth and final speech focused on the quality of the education workforce – teachers, principals, and education support professionals. Duncan concluded his speech by presenting a challenge: "to make sure every child in America is learning from an effective teacher- no matter what it takes." He asked the audience to join President Obama and himself in "a new commitment to results that recognizes and rewards success in the classroom and is rooted in our common obligation to children."[1]

Districts and states across the country are engaging in efforts to redesign educator evaluation systems, motivated by two main factors. First, teachers generally did not receive meaningful feedback on their

[1] Duncan, Arne. "Partners in Reform." Address by the Secretary of Education to the National Education Association. 2 July 2009. Speech.

instructional practices and had little guidance about what was expected of them in the classroom. Second, traditional teacher evaluation systems did not differentiate among high- and low- achieving teachers.[2]

Race to the Top Grant

On February 17, 2009, President Obama signed into law the *American Recovery and Reinvestment Act of 2009* (ARRA), historic legislation designed to stimulate the economy, support job creation, and invest in critical sectors, including education. The ARRA provided $4.35 billion for the Race to the Top Fund, a competitive four-year grant program designed to encourage and reward States that are creating the conditions for education innovation and reform; achieving significant improvement in student outcomes, including making substantial gains in student achievement, closing achievement gaps, improving high school graduation rates, and ensuring student preparation for success in college and careers; and implementing ambitious plans in four core education reform areas:

- Adopting standards and assessments that prepare students to succeed in college and the workplace and to compete in the global economy;
- Building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction;
- Recruiting, developing, rewarding and retaining effective teachers and principals, especially where they are needed most; and
- Turning around the lowest-achieving schools.[3]

In 2010, the U.S. Department of Education (ED) awarded Race to the Top Phase 1 and Phase 2 grants to 11 States and the District of Columbia. In 2011, the ED awarded Phase 3 grants to seven additional States which were finalists in the 2010 Race to the Top Phase 2 competition. [4] [5]

[2] Sartain,Lauren, Sara Ray Stoelinga, and Eric R. Brown. "Rethinking Teacher Evaluation in Chicago: Lessons Learned from Classroom Observations, Principal Teacher Conferences, and District Implementation." N.p., Nov. 2011. Web. June 2013. <http://ccsr.uchicago.edu/sites/default/files/publications/Teacher20Eval20Report20FINAL.pdf>.

[3] Race to the Top Executive Summary http://www2.ed.gov/programs/racetothetop/executive-summary.pdf

[4] Race to the Top Pennsylvania Report Year 1: 2012 www2.ed.gov/programs/racetothetop/performance/pennsylvania-year-1.pdf

[5] Also in 2011, USDE made nine awards under the Race to the Top – Early Learning Challenge to improve quality and expand access to early learning programs, and close the achievement gap for children with high needs. In 2012, four more states received Early Learning Challenge grants. Additionally, in 2012, USDE made awards to 16 applicants through the Race to the Top – District competition to support local education agencies (LEAs) implementing locally developed plans to personalize and deepen student learning, directly improve student achievement and educator effectiveness, close achievement gaps, and prepare every student to succeed in college and careers.

Pennsylvania Context and Legislation

> *Race to the Top States are developing comprehensive systems of educator effectiveness by adopting clear approaches to measuring student growth; designing and implementing rigorous, transparent, and fair evaluation systems for teachers and principals; conducting annual evaluations that include timely and constructive feedback; and using evaluation information to inform professional development, compensation, promotion, retention, and tenure decisions. In addition, Race to the Top States are providing high quality pathways for aspiring teachers and principals, improving the effectiveness of teacher and principal preparation programs, and providing effective support to all educators.*
>
> Race to the Top Program, Executive Summary

The Pennsylvania Department of Education (PDE) designed its Race to the Top Phase 3 application to accelerate key aspects of the State's strategic plan for education. As one of seven States to receive a Race to the Top Phase 3 grant, Pennsylvania received a total of $41.3 million over four years.

Pennsylvania's current system of evaluation (prior to 2013-2014) includes two ratings for educators – "satisfactory" or "unsatisfactory." Statewide results show that 99.4 percent of all teachers and 99.2 percent of all principals who were evaluated during the 2009-10 school year received a "satisfactory" rating. Despite these results, student growth on national assessments has been relatively stagnant.[6] In a 2011 testimony on teacher and principal evaluation systems, Carolyn Dumaresq, then Deputy Secretary for Elementary and Secondary Education for the Pennsylvania Department of Education said how these statistics show the need to have "a broad, multi-measure evaluation system to measure performance and effectiveness. We will be better able to gauge our educators' levels of performance and also allow them opportunities for development or guidance with an effective evaluation system in place in order to target an improvement plan."

As part of its comprehensive and coherent approach to education reform, the State is committed to improving educator effectiveness, and has completed its third year (SY 2012-2013) of the continued development and implementation of a new teacher, specialist (non-classroom teacher) and principal evaluation system that evaluates educators' professional practices and incorporates student performance results as a significant factor.

Pennsylvania has contributed to the tool's implementation by providing professional development in the use of the new evaluation systems, including how to utilize the information to improve teacher and principal effectiveness. The State is also working to improve access to data that can be used to inform instruction.

In 2010, PDE launched the development of its teacher evaluation system, starting with the selection of a teacher practice observation tool based on Charlotte Danielson's Framework for Teaching. A pilot was conducted in Spring 2011 with four LEAs and one Intermediate Unit (IU), and expanded to over 100 LEAs in SY 2011-2012. As part of its Race to the Top plan, the State finalized its classroom teacher observation rubric in Summer 2012, with updates made to the rubric based on lessons learned in the first two pilots. As of SY 2012-2013, the rubric has been rolled out to all districts, including The School District of Philadelphia (SDP).

[6] *HB 1980 - Educator Evaluation System*, House Education Committee Cong. (2011) (testimony of Carolyn Dumaresq, D. Ed., Deputy Secretary for Elementary and Secondary Education). Print.

Also in Year 1, the State continued its pilot of the principal evaluation rubrics. The pilot began in 27 LEAs in SY 2011-2012 and was expanded to 237 LEAs for SY 2012-2013, including SDP. During summer 2012, the rubrics were revised based on lessons learned from the pilots and will be rolled out to all LEAs in SY 2013-2014.

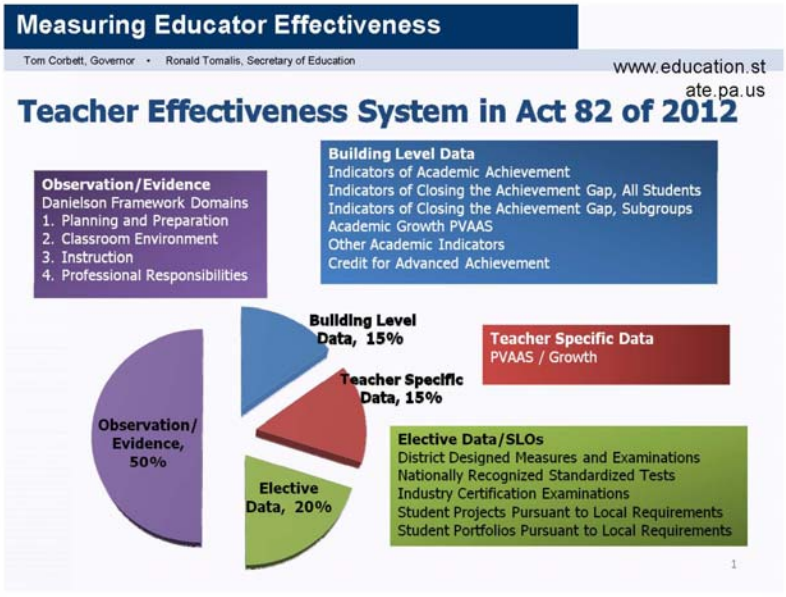
Act 82 of 2012 (H.B. 1901)

In June 2012, House Bill (H.B.) 1901 was passed into law requiring that 50% of an educator's (including teachers, principals, and specialists) overall evaluation score be based on multiple measures of student performance, with the remaining portion of the overall rating based on measures of professional practice such as observations.

As part of this policy, the number of possible rating categories was expanded from two: satisfactory, unsatisfactory to four: distinguished, proficient, needs improvement, and failing. An overall performance rating of distinguished or proficient will be considered satisfactory. An overall performance rating of needs improvement will be considered satisfactory, except that any subsequent overall rating of needs improvement issued by the same employer within 10 years of the first overall performance rating of needs improvement where the employee is in the same certification shall be considered unsatisfactory. An overall performance rating of failing shall be considered unsatisfactory. An overall performance rating of needs improvement or failing will require the employee to participate in a performance improvement plan.

Teacher evaluation results for educators in tested grades and subjects will be based on Observation/Evidence (50%), Building Level Data (15%), Elective Data/ Student Learning Objectives (SLOs) (20%), and Teacher Specific Data based on PVAAS/Growth (15%), shown in Figure 1.

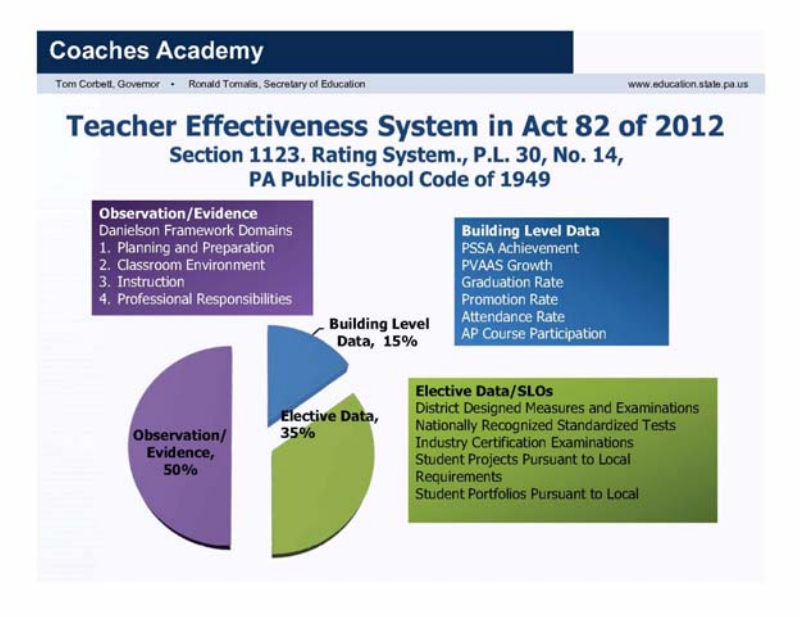
Figure 1. Teacher Effectiveness System in Act 82 of 2012.



Source: Pennsylvania Department of Education www.education.state.ps.us

Evaluation results for teachers in non-tested grades and subjects will include Observation/Evidence (50%), Building Level Data (15%), and Elective Data/ SLOs (35%), as shown in Figure 2.

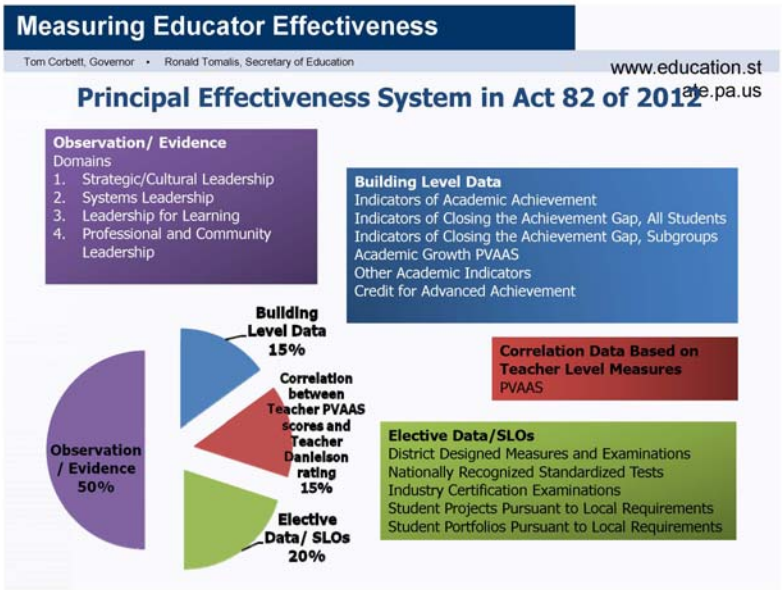
Figure 2. Teacher Effectiveness System in Act 82 of 2012 – Non-tested Grades and Subjects



Source: Pennsylvania Department of Education www.education.state.ps.us

Principal evaluation results will be based on Observation/Evidence (50%), Building Level Data (15%), Correlation Data Based on Teacher Level Measures (15%), and Elective Data/ SLOs (20%), as shown in Figure 3.

Figure 3. Principal Effectiveness System in Act 82 of 2012.



Source: Pennsylvania Department of Education www.education.state.ps.us

Ratings for Specialists, which includes non-teaching professional employees who do not provide direct instruction to students, will be based on Observation/Evidence (80%) and Student Performance of All Students in the School Building in which the Nonteaching Professional Employee is Employed (20%), as shown in Figure 4.

Figure 4. Non Teaching Professional Employee Effectiveness System in Act 82 of 2012.

Source: Pennsylvania Department of Education www.education.state.ps.us

Teachers will begin receiving ratings based on the new evaluation system in SY 2013-14 and specialists and principals will receive ratings beginning in SY 2014-15. Once ratings have been given, LEAs can begin to use evaluation results to inform personnel decisions.

The law states that three years of student growth data must be used to inform the student growth portion of the teacher evaluation for teachers in tested grades and subjects. Three years of data will be available for the first time in SY 2015-2016, and until student growth data is available, observation and/or Elective Data/SLOs will be used instead. During the first year of implementation in 2013-2014, only Observation Data and Building Level Data will be used in a teacher's rating[7].

The State has developed a value-added model to measure student growth on State assessments. LEAs will be responsible for the development of Student Learning Objectives (SLOs) to measure student progress in non-tested grades and subjects. In an effort to support LEAs in the development of SLOs, PDE has created a rubric that outlines general requirements for developing SLOs and has indicated that it will provide training to districts.

In order to implement the full evaluation system, including multiple measures of student performance, the State has been improving its data systems to be able to link teachers with their respective students. PDE piloted its enhanced data system for the first time in Spring 2013. SDP participated in this pilot.

[7] LEAs have the option to use the SLO measure for 2013-2014 ratings, if they are prepared to do so.

Local Context and SDP Program Design

The Race to the Top program requires that LEAs work with the State to implement all or significant portions of the State's Race to the Top plan. In spring 2012, Pennsylvania initiated a RTTT Phase 3 LEA Grant Program, with the goal of working with participating entities to adopt and implement Pennsylvania's Educator Effectiveness Instrument and use the evaluation process and results to inform local decisions. Approximately 50% of Pennsylvania RTTT award was designated specifically for formula grants to eligible LEAs in support of this goal.

The School District of Philadelphia applied for and received an allocation amount of $11,112,128 for a three-year period (SY 2012-2013, 2013-2014, 2014-2015), which the School Reform Commission voted to accept on August 16, 2012.

The program required appropriate professional development training on the new educator effectiveness system for all LEA supervisors (all staff who contribute to the evaluation of staff) as well as for all teachers and specialists (staff who are evaluated but do not contribute to the evaluation of other staff). The performance measures, which awardees must agree to meet or exceed by the end of the school year indicated, are shown in Table 1.

Table 1. Race to the Top Performance Measures

	SY 12/13	SY 13/14	SY 14/15
SUPERINTENDENTS & CENTRAL OFFICE STAFF			
Trained on Educator Effectiveness Principal Instrument.	25%	50%	75%
PRINCIPALS			
Trained on the Educator Effectiveness Teacher Instrument.	50%	100%	-
Trained on the Educator Effectiveness Specialist Instrument.	25%	50%	75%
Trained on the Educator Effectiveness Principal Instrument.	25%	50%	75%
Evaluated using the Educator Effectiveness Principal Instrument.	-	10%	50%
TEACHERS			
Trained on the Educator Effectiveness Teacher Instrument.	50%	100%	-
Evaluated using the Educator Effectiveness Teacher Instrument.	10%	50%	100%
SPECIALISTS			
Trained on the Educator Effectiveness Specialist Instrument.	25%	50%	75%
Evaluated using the Educator Effectiveness Specialist Instrument.	-	10%	50%
ALL EDUCATORS			
Using the professional development modules associated with the Educator Effectiveness Instrument within the SAS portal.	10%	30%	50%

Source: Race to the Top LEA Grant Guidelines

Race to the Top grant monies awarded to LEAs may be used for any expense incurred through participation in related professional development provided by 1) the Intermediate Unit or approved provider or 2) the grantee and/or asynchronous SAS modules, as well as for costs associated with the implementation of the Educator Effectiveness Instrument post professional development. Each LEA program design must demonstrate consideration to the following areas:

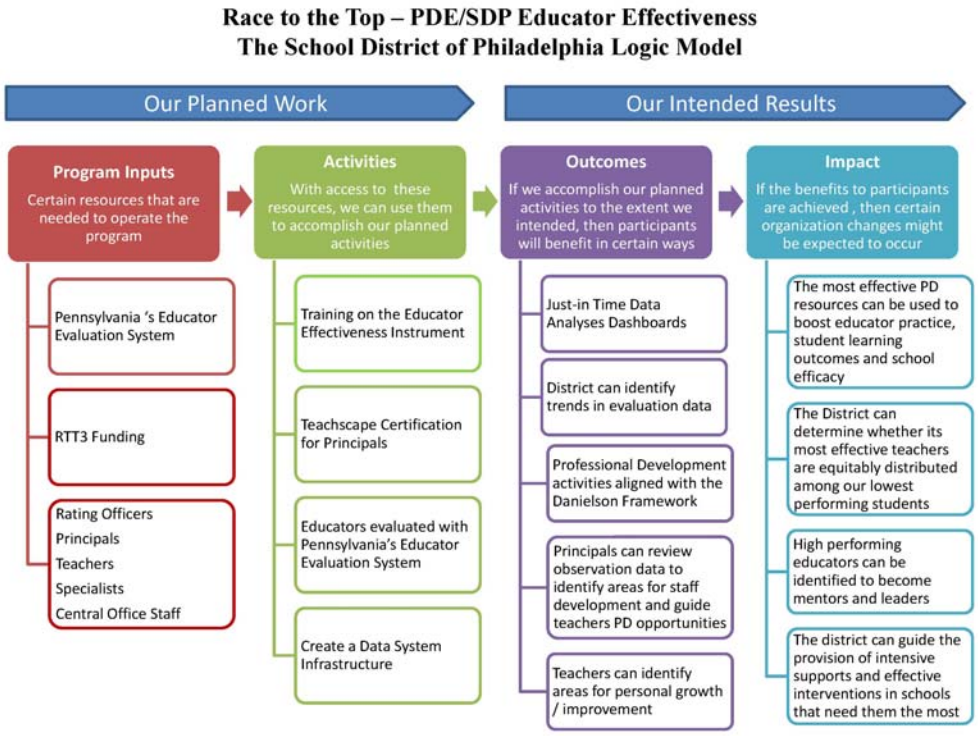
- Improving educator effectiveness in schools, resulting in all schools having highly qualified teachers;
- Guiding the provision of intensive supports and effective interventions in schools that need them most, thereby improving achievement in low-performing schools;
- Enhancing the collection and use of data to improve student learning and college and career-readiness
- Assisting in progress towards the adoption and implementation of PA Common Core standards; advancing performance in assessments, and enriching teaching and learning particularly in STEM education.

The District's program design is geared specifically to The School District of Philadelphia's Action Plan v1.0, Strategy 4: To Identify and Develop Committed, Capable, People. The project design reflects this core philosophy and approach and includes the following features as its major activities:

- Increasing knowledge of the educator effectiveness systems and processes through professional development and training;
- Creating a more useful, streamlined, and efficient educator rating system through the purchase and utilization of an Educator Effectiveness Technology Solution;
- Establishing student to teacher linkages for calculating PVAAS Teacher Specific Growth Data;
- Increasing Instructional Capacity through academic coaching;
- Ensuring fair, valid, and reliable observation data through inter-rater reliability certification for Principals;
- Establishing Student Learning Objectives that align with District objectives; and

A basic logic model is included below in Figure 5 to highlight inputs, processes, short-term outcomes, intermediate outcomes, and long-term outcomes. Program inputs refer to the components of the program design, processes refer to the methods of delivery and instruction, and the outcomes related to performance measures and data sources that will assess the impact of the program. The model will be developed further as the project unfolds and additional factors are identified.

Figure 5. Logic Model of Race to the Top Program Inputs, Activities, Outcomes, and Impact.



Existing SDP Systems and Procedures

This section outlines the existing teacher evaluation procedures in place at SDP as of the 2012-2013 school year. The School District of Philadelphia currently employs a Professional Growth System (PGS) that has been co-constructed with the Philadelphia Federation of Teachers (PFT). One element of the PGS is the Professional Development Plan (PDP) for professional employees. Professional employees in their PDP years must complete a designed plan in collaboration with the principal, with content that aligns with the school's data and goals reflected in the school's Action Plan.[8] SDP employs a differentiated model of supervision, in that employees operate under a different observation structure depending on the number of years they have been with the District. A tenured employee may either be in a Formal Observation Year or a Professional Development Plan (PDP) Year. The differentiated schedule of supervision is shown in Table 2.

[8] Professional Growth System Handbook.2012-2013 Edition.The School District of Philadelphia in Partnership with the Philadelphia Federation of Teachers. https://cc.philasd.org/service/home/~/Professional%20Growth%20 System%20Booklet%208%2024 %2012.doc-4.pdf?auth=co&loc=en_US&id=43013&part=2

Table 2. School District of Philadelphia Differentiated Supervision Schedule for 2012-2013

Years	4	5	6	7	8	9	10	11	12	13	14
Type	P	P	X	P	P	X	P	P	X	P	P

Years	15	16	17	18	19	20	21	22	23	24	25
Type	X	P	P	X	P	P	X	P	P	X	P

Years	26	27	28	29	30	31	32	33	34	35	36
Type	P	X	P	P	X	P	P	X	P	P	X

Years	37	38	39	40
Type	P	P	X	P

X=Formal Observation Year

P= Professional Development Plan (PDP) Year

Source: School District of Philadelphia Professional Growth System Handbook 2012-2013 Edition.

The following timeframe is used for the development and monitoring of the PDP:

- June-October: PDP collaborative meeting no later than October 20th
- Mid-year review meeting no later than January 15th
- End of the year review meeting no later than June 15th.

Goals included in the PDP are required to be:

- Specific, with outcome that show progress over time;
- Measurable;
- Attainable within the PDP cycle;
- Relevant to the school data and approved Action Plan; and
- Timely, so that progress can be assessed during the appropriate review dates

Peer Assistance and Review (PAR) is a component of the PGS and is a mandatory program for all non-tenure, first year teachers and for tenured teachers who have been rated unsatisfactory in the previous school year. PAR is also available to teachers on an as needed basis. For example, participation in PAR may be requested by a tenured teacher who believes that his/her teaching competence will benefit from PAR or by a principal for tenured teachers who are in their PDP years. The PAR program has ongoing leadership provided by a Panel of eight members, four of whom shall be selected by the Federation and four of whom shall be selected by SDP. The panel is divided into PGS Pairs consisting of one Federation appointed member and one District appointed member. PGS Pairs meet regularly with Consulting Teachers to review the work of Consulting Teachers and the progress of teachers assigned to the PGS Pair, to evaluate teachers and make retention recommendations to the PGS Panel as a whole. The Panel makes discretionary decisions regarding all components of the PGS, including but not limited to:

- Determining eligibility for the PAR Program,

- Monitoring the overall progress of teachers participating in the PAR,
- Making retention recommendations for new teachers and tenured teachers participating in PAR, and
- Creating and monitoring processes of the SDP.

For tenured teachers who are in a formal observation year, as well as non-tenured teachers, the schedule of observations is shown in Table 3 below.

Table 3. School District of Philadelphia Observation and Rating Schedule for 2012-2013.

Employee Type	Employee Status	Previous Year's Rating	Observation and Rating Schedule
Temporary Professional Employee(non-tenure)	1st Year Teacher	**N/A**	**1 Formal Observation** by May 25 of school year **Multiple Informal Observations** **One Rating** by June 1 of school year
	2nd and 3rd Year Teachers	**Satisfactory**	**Four Formal Observations** (1 and 2 by Jan 31; 3 and 4 by June 8) **Multiple Informal Observations** **Two Ratings** by Jan 31 and June 18 of school year
		Unsatisfactory	**Four Formal Observations** (3 by school administrator, 1 by assistant superintendent) **Multiple Informal Observations** **Two Ratings** by Jan 31 and June 18 of school year
Professional Employee (tenure)	Not in PDP Year	**Satisfactory**	**Two Formal Observations** 1st by Jan 3, 2nd by June 8 of school year **Multiple Informal Observations** **One Rating** by June 18 of school year
	In PDP Year	**Satisfactory**	**No Formal Observations** (special observation status, if applicable) **Multiple Informal Observations** **One PDP** post-conference by June 13 of school year **One Rating** by June 18 of school year
	Rated 'U' from Previous Year	**Unsatisfactory**	**One Formal Observation** by May 25 of school year **Multiple Informal Observations** **One Rating** by June 1 of school year *Final recommendation to PAR Panel for Retention or Dismissal*

Source: School District of Philadelphia Professional Growth System Handbook 2012-2013 Edition.

As defined in the PGS guidebook, formal classroom observations and informal walkthroughs are an essential component for professional growth and development. Prior to the observation, administrators are to hold individual pre-observation conferences with each teacher. Teachers are to complete the pre-observation form prior to the conference with the administrator and then meet to discuss the form and prepare for the observation. After the formal observation, administrators will hold individual post-observation conferences with each teacher, using the post-observation form.

If a principal has concerns about the performance of a tenured teacher who is not currently in a formal observation year, the principal may request that the PAR Panel place the teacher on Special Observation Status (SOS). See Appendix A for the full Special Observation Status procedure.

Evaluation Research and Methodology

The RTTT evaluation consists of two major components: an implementation study (formative), and an impact evaluation (summative). The implementation component examines the extent to which the various features of the program are rolled out with fidelity or appropriate adaptations. Multiple data sources and analytical techniques are being utilized to conduct the evaluation, including document analysis, direct observation, feedback surveys, interviews, and focus groups. A range of programmatic data (qualitative and quantitative) is being collected and analyzed.

Through Year 1, the program office has relied on the specific objectives defined by the State for training requirements. For Year 2, the program office will need to further define specific objectives on what the program proposes to do in the realms of professional development and training, certification, coaching, technology solution implementation, PVAAS teacher specific growth, and elective data/SLOs.

While the primary purpose of the evaluation is to explicate the implementation of RTTT and the impact of the new rating system on educator practice, the secondary goal is to examine linkages to student achievement. If the premise of new methods of teacher evaluation is based on improving teacher practice, and improvements in teacher practice are linked to improved student outcomes, student performance is likely to improve with an improvement in teacher performance. To test this, we will analyze data about teachers' practice and student test results. A series of multivariate modeling empirical relationships between principal, teacher, and student outcomes will be used in subsequent years.

Core Research Questions

For the first year report, the focus will be on the first question. In subsequent years, as the new rating system is implemented, the remaining core research questions will be addressed.

How is the RTTT program implementing components of the program, including:

- Professional Development and Training;
- Implementation and Utilization of a Technology Solution;
- PVAAS Teacher Specific Growth data;
- Elective Data/Student Learning Objectives (SLOs)
- Coaching and Support; and
- Inter-rater Reliability Certification.

Did the new rating system impact teacher and principal performance and practice? If so, how?

Have possible changes in principal and teacher practice impacted student achievement? If so, how?

Table 4. Data collection for 2012-2013 (Year 1)

Data Type	Data Source
Interviews & Surveys	Teachers and Principals involved in RTTT initiatives such as trainings and pilot studies
Observations	Planning Meetings
	Training Sessions
Documents	Sign-in Sheets/Attendance Lists
	Observation Tracking Forms

Year 1 Findings

This section summarizes the findings on the research questions related to program implementation in the RTTT program. In Year 1, the components of the program that are being addressed are professional development and training, coaching and support, implementing a technology solution, inter-rater reliability, and establishing student-teacher linkages for PVAAS teacher specific growth data.

Professional Development and Training

The performance measures relevant to professional development and training were established by the State prior to the District's acceptance of the grant funding. An overview of the District's performance in meeting Year 1 training goals is shown in Table 5.

Table 5. District Performance on Year 1 Program Goals

	Year 1 Target	Year 1 Actual	Target Met?
SUPERINTENDENTS & CENTRAL OFFICE STAFF			
Trained on Educator Effectiveness Principal Instrument.	25%	100%	●
PRINCIPALS			
Trained on the Educator Effectiveness Teacher Instrument.	50%	83%	●
Trained on the Educator Effectiveness Specialist Instrument.	25%	82%	●
Trained on the Educator Effectiveness Principal Instrument.	25%	82%	●
Evaluated using the Educator Effectiveness Principal Instrument.	-	-	-
TEACHERS			
Trained on the Educator Effectiveness Teacher Instrument.	50%	7%	■
Evaluated using the Educator Effectiveness Teacher Instrument.	10%	1%	■
SPECIALISTS			
Trained on the Educator Effectiveness Specialist Instrument.	25%	2%	■
Evaluated using the Educator Effectiveness Specialist Instrument.	-	-	-

●= Target Met ■=Target Not Met

Superintendents &Central Office Staff

The goal of training 25% of Superintendents and Central Office staff on the Principal Effectiveness Instrument was exceeded, as each of the seven Regional Assistant Superintendents received training, which was provided by the State over a two-day period on October 15 and 16, 2012 in Harrisburg, PA.

Principals

The trained Assistant Superintendents were then required to provide turn-around training to principals on the Educator Effectiveness System. These trainings were delivered on two occasions. The first training was delivered to principals on January 17, 2013 at seven locations throughout the district, facilitated primarily by a trained Assistant Superintendent, with support from an ORE Research Associate. A total of 241 principals attended one of the seven trainings. The second was delivered primarily to assistant principals on February 4, 2013 at 440 N. Broad Street, facilitated by an Assistant Superintendent, a principal and an ORE research associate. A total of 103 assistant principals attended the training.

The Office of Research and Evaluation administered a survey at the conclusion of the trainings to solicit feedback regarding the effectiveness of the trainings, as well as to collect some additional information regarding principals' and assistant principals' perceptions of SDP's current rating system, their own professional practice, and SDP's professional growth system. A total of 318 participants (92%) completed the survey. Overall, the majority of participants felt that the presentation(s) was communicated clearly and effectively, answered questions they had about the Principal Effectiveness System, and gave them a clear understanding of how they will be rated using the new system. These results are shown in Table 6.

Table 6. Survey Responses: Effectiveness of Principal Evaluation System Training

Effectiveness of the Training	The following questions ask about the effectiveness of the training you just received.			
To what extent do you agree or disagree with the following statements?	**Strongly Disagree**	**Disagree**	**Agree**	**Strongly Agree**
The presentation content was communicated clearly and effectively.	1%	1%	51%	48%
The presentation content answered questions I had about the Principal Effectiveness System.	1%	8%	62%	30%
The presentation content gave me a clear understanding of how Principals/APs will be rated using the new system.	1%	9%	60%	30%

In a series of questions related to SDP's current rating system, the majority of respondents replied that SDP's current rating system provides them with individualized useful feedback from their rating officer, and is used to facilitate meaningful dialogue with and among educators. The majority also feel that the process of being rated is professionally meaningful and assists with the improvement of their practice, and positively impacts outcomes for students. These results are shown in Table 7.

Table 7. Survey Responses: Principals' Perceptions of School District of Philadelphia's Rating System

SDP's Rating System	The following questions ask about your experience during the most recent FULL rating period (2011-2012).				
To what extent do you agree or disagree with the following statements?	**I was not rated**	**Strongly Disagree**	**Disagree**	**Agree**	**Strongly Agree**
Under the rating system and supervision structure I am provided with individualized and useful feedback from my rating officer/supervisor.	5%	3%	8%	61%	22%
SDP uses educator evaluations to facilitate meaningful, growth producing dialogue with and among educators.	2%	4%	16%	68%	10%
The process of being rated is professionally meaningful to me and assists me with the improvement of my practice.	3%	2%	9%	61%	26%
I feel that SDP's process of evaluating and rating principals positively impacts outcomes for students.	2%	2%	15%	63%	17%

When asked a series of questions about their professional practice, the majority of principals and assistant principals responded that they often collaborate with other education professionals around improving student outcomes, use data and feedback to improve their practice as well as teaching and learning at their school. The majority of respondents said that they often access other resources outside of SDP in an effort to improve their practice, though 21% rarely or never do so. The majority responded that they feel a sense of responsibility for student outcomes at their school, and feel that SDP holds them accountable to those outcomes. These results are shown in Table 8.

Table 8. Survey Responses: Principals' Professional Practice.

Principals' Professional Practice	The following questions ask about your current professional practice.			
In the past year, how often have you engaged in the following activities?	**Never**	**Rarely (1-2 times per month)**	**Sometimes (3-4 times per month)**	**Often (7-8 times per month)**
Collaborated with other education professionals around improving student outcomes.	1%	12%	38%	49%
Used data and feedback to improve my own practice through an on-going process of planning, assessment and improvement.	0%	10%	41%	49%
Used data and feedback to improve the teaching and learning at my school through an on-going process of planning, assessment and improvement.	0%	6%	35%	59%
Accessed other resources outside of SDP in an effort to improve my practice.	1%	20%	48%	30%

To what extent do you agree or disagree with the following statements?	**Strongly Disagree**	**Disagree**	**Agree**	**Strongly Agree**
I feel a sense of responsibility for student outcomes at my school.	0%	0%	15%	84%
I feel that I am held accountable by SDP for student outcomes at my school.	0%	1%	27%	71%

Regarding SDP's professional growth system, the majority of total respondents felt that opportunities at SDP are aligned with the standards by which they are evaluated. However, when responses were broken down by role, a much higher percentage of assistant principals (38%) compared to principals (21%) do not feel that

opportunities for professional growth at SDP are aligned with their evaluation standards. The majority of respondents agree that there is alignment between SDP's goals, their school goals and their teachers' goals, and most believe that SDP uses high quality measures of educator and student performance. When broken down by role, however, 26% of assistant principals do not feel that there is alignment between SDP's goals, their school goals, and their teachers' goals, compared with only 13% of principals. The overall majority of respondents feel that they have easy access to examples of best practices in their field, and that SDP is effective in identifying opportunities to improve the practice of educators. However more assistant principals (31%) disagree that they have easy access to examples of best practices in their field compared with principals (19%). These responses are shown in Table 9.

Table 9. Survey Responses: School District of Philadelphia's Professional Growth System.

Professional Growth System				The following questions ask about SDP's CURRENT professional growth system.
To what extent do you agree or disagree with the following statements?	**Strongly Disagree**	**Disagree**	**Agree**	**Strongly Agree**
I feel that opportunities at SDP are aligned with the standards by which I am evaluated.	3%	22%	67%	9%
Currently, there is alignment between SDP's goals, my school goals and my teachers' goals.	0%	16%	70%	14%
Currently, I believe SDP uses high quality, valid, and reliable measures of educator and student performance.	1%	28%	66%	5%
I have easy access to examples of best practices in my field.	1%	21%	62%	16%
SDP is effective in identifying opportunities to improve the practice of educators.	2%	25%	65%	9%

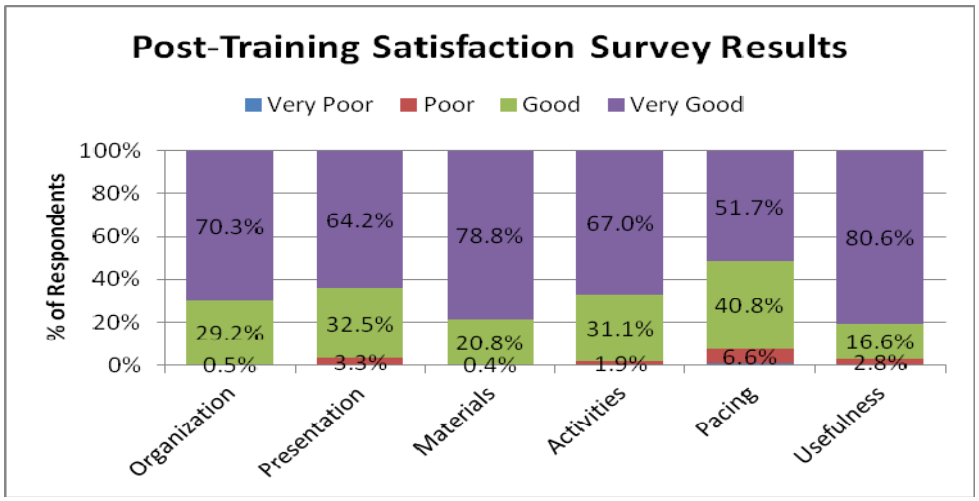
In order to provide an in-depth training on the Teacher Effectiveness System to principals, both to prepare them to implement the new system in the upcoming fall, as well as provide thorough turn-around training to their teachers, SDP contracted with The Danielson Group to provide a 3-day training on the Danielson Framework for Teaching clinical supervision model as it is intended to be implemented as part of the new educator effectiveness system. This training was held from June 26-28, 2013 at Lincoln High School, and was facilitated by trainers from the Danielson Group, with participants split into several individual groups, arranged by Principal Learning Team (PLT). A total of 345 principals and assistant principals attended the training. The content objectives for the training were:

- Learn the instrument and understand how it defines quality teaching;
- Learn observation skills;
- Learn how to apply the rubrics and score all dimensions accurately;
- Learn to minimize the impact of professional biases; and
- Prepare to train teachers on the framework material.

At the conclusion of the 3rd and final day of training, the Danielson Group administered a satisfaction survey, which was completed by 213 participants (62%). The survey addressed questions related to the organization, presentation, materials, activities, pacing and the usefulness of the PD to the participants in the form of a Likert scale (1=Very Poor to 4= Very Good). The participants were also asked to explain what aspect(s) of the training were most useful, and offer any revisions they would make to the training if given the opportunity.

Training feedback was very positive, with the vast majority of participants responding that the organization, presentation, materials, activities, pacing and usefulness of the presentation were very good. These results are shown in Figure 6 below.

Figure 6. Danielson Training Survey: Post-Training Satisfaction Survey Results.



The most useful elements of the training, according to the feedback, were learning about the different domains, components and elements included in the Danielson Framework for Teaching, the training activities (video clips, scenario cards, etc.), the acquisition of usable new coaching/teaching skills, the opportunity for collaboration, the ability to perform better evaluations, learning the distinction between evidence and opinion, and the increased knowledge of the scoring rubrics. The elements of the training that could use improvement, according to the feedback, including the pacing (which was sometimes cited as too fast, and sometimes as too slow), and the level of support provided (more collaboration, practice, and examples).

Teachers

The program fell short of its goal to train 50% of teachers on the Teacher Effectiveness System during Year 1, and trained approximately 7%, or 625 members of the teaching staff. Given the size of the District and the number of staff, the program office determined that the best model for training all 9,500 teachers on the Educator Effectiveness System would be with two back-to-back designated Professional Development (PD) days, with training conducted within each SDP school by the school's Principal, Assistant Principal, and other members of the schools' leadership teams. In order to conduct these trainings with fidelity, the District wanted to ensure that all Principals and Assistant Principals were fully trained in the Educator Effectiveness system, as well as prepared to conduct turn-around trainings on the system. Because the designated District PD days for teachers, as well as their content, are determined almost a full school year in advance, it became virtually impossible to conduct the intended district-wide teacher PD during the 2012-2013 school year. As a result, the District has fallen short of its required number of teachers trained for the 2012-2013 school year. However, the District has identified and earmarked November 5 & 6, 2013 as professional development days that will be solely dedicated to principals' turn-around trainings on the Teacher Effectiveness System to all of their teachers.

The teachers who have been trained were either part of a Cohort school, which required the principals to train the teachers this year in preparation to pilot the observation system, attended training for school-based teacher leaders facilitated by program staff, or attended a training for teacher coaches and consulting teachers facilitated by the Bucks County IU.

Feedback was collected after the training conducted for teacher coaches and consulting teachers by Bucks County IU on April 16 & 17, 2013. Participants were asked to record '2 remaining questions' that they had at the conclusion of the two-day training. Table 10, below, shows the six categories for which the '2 remaining questions' fell into, examples for each category, as well as the number of respondents.

Table 10. Survey Responses: 2 Remaining Questions after TC/CT Training

2 Remaining Questions Categories	Number of Respondents
1. Implementation of the new system and ongoing support • how the district will implement and support the system with limited time, staff, and resources while ensuring fidelity.	17
2. Training on the new Educator Effectiveness System • when and how teachers, principals, Consulting Teachers (CTs), Teacher Coaches (TCs) and other employee types would be trained in the new system and going forward, and how other stakeholder communities would be involved	10
3. Local Decisions and Processes • what the teacher evaluation process will look like in Philadelphia SD, • what the process will be for unsatisfactory teachers, and • what the District will choose as its Student Learning Objectives (SLOs)	9
4. Changing from a Punitive culture to a culture of growth and support • how leaders will be trained to change the climate and become better instructional leaders, • how long it will take to change from a punitive culture to a collaborative growth-focused culture, • how trust will be established between rater and rated, and • how culture change will be monitored and reinforced.	9
5. Multiple Measures and Summative Ratings • how building level data would impact teachers in high needs schools, • how scores will be translated into a final rating, • how the PVAAS three-year average will be calculated, and • how teachers' evaluations will be tied to principals' evaluations	6
6. Evaluation System for Non-Teaching Professional Employees and other employee types • how teacher coaches, ESOL/Special ed teachers, and long-term substitutes will be evaluated, • how data will impact non-instructional staff and how non-instructional staff will be professionally developed	6

The program also fell short of meeting its goal to pilot the observation process on 10% of teachers during Year 1, and piloted the system with approximately 75 teachers, or about 1%. Throughout the 2012-2013 school year, approximately 60 principals were trained over four half-days (two complete days) by the Bucks County IU in three cohorts on use of the Danielson Framework for Teaching as well as in the new rating system. These principals would be responsible for providing turn-around training and piloting the observation system on 10-20 of their teachers throughout the school year. There were two primary differences between the existing SDP observation structure and the pilot. The first is the forms being used. Currently, SDP uses a modified Danielson framework that has been modified in collaboration with the union. The pilot employed Danielson forms, including rubrics, feedback forms, etc. The second is the method of collaboration between principal and teacher that allows them to use discussion techniques to arrive on a mutually agreed-on score for the teacher.

Abundant evidence indicates that a thoughtful approach to teacher evaluation—one that engages teachers in reflection and self-assessment—yields benefits far beyond the important goal of quality assurance. Such an approach provides the vehicle for teacher growth and development by providing opportunities for professional conversation around agreed-on standards of practice[9]. The steps that follow represent the piloted supervisory system adopted by The School District of Philadelphia (SDP) aligned with Danielson's Framework for Teaching:

Step #1 Pre-Observation: Domain 1, Domain 4

- Teacher completes Lesson Plan Document and shares it with the Evaluator[10] two days in advance of the Pre-Conference.
- Teacher and Evaluator meet to discuss the upcoming lesson.
- Evidence is added to the lesson plan document that emerges from the pre-observation conference.

Step #2 Observation: Domain 1, Domain 2, Domain 3

- Evaluator arrives 5 minutes prior to the beginning of the lesson to 'walk the walls'.
- Evaluator observes the lesson and records evidence.

Step #3 Preparing for the Post-Teaching Conference: Domain 1, Domain 2, Domain 3, Domain 4

- Evaluator provides Teacher with completed observation form from Step #2 within 24 hours of observation.
- Teacher is provided with an opportunity to add evidence to the observation form that may have been overlooked by Evaluator.
- Teacher returns the observation form to Evaluator with any additions.

[9] Danielson, Charlotte. "Evaluations That Help Teachers Learn." *The Effective Educator* 68.4 (2011): 35-39. *www.danielsongroup.org*. Dec.-Jan. 2010/2011. Web.

[10] The term Evaluator refers to any individual certified to evaluate teacher effectiveness, such as a Principal, Assistant Principal, Assistant Superintendent, etc.

- Teacher completes the Teacher Self-assessment Rubric, highlighting the words in the rubric they think best characterize the evidence of their performance (may highlight phrases in multiple levels of the same component).
- Evaluator completes the Evaluator Assessment Rubric, highlighting the words in the rubric that best characterize the evidence of the teacher's performance.

Note: The Teacher and Evaluator do not need to meet during Step #3. With requisite training, the Teacher can engage in Step #3 independently or with the support of a coach.

Step #4 Post-teaching Collaborative Assessment: Domain 1, Domain 2, Domain 3, Domain 4

- Teacher and Evaluator meet to reflect on the lesson. Evaluator notes components of agreement on the rating form and then invites teacher to take the lead in discussing the other components where highlighted areas are not aligned.
- Components are collaboratively rated. Evaluator is the rater of record in the event of non-agreement. Evidence is the basis.
- Post-conference ends with Observation Summary being completed collaboratively. Either party can write.

Meetings were held monthly at Lincoln High School from March through June, 2013 for the principals participating in the pilot to provide feedback and receive updates regarding the current Educator Effectiveness/RTTT initiative and their involvement in the pilot. At the conclusion of the April 24, 2013 meeting, a paper-based survey was administered to collect feedback about principals' experiences in the pilot so far. Overall, principals expressed very high regard and praise for the structure of the observation framework model and appreciated the increased collaboration with teachers that it offered. They also felt as though they had received sufficient training on the process prior to piloting. However, when asked about challenges in implementing this process, nearly every principal cited concerns about the significant amount of time involved in conducting the full observation conference, formal observation, and post-observation conference with each teacher. The ability to find 30-45 uninterrupted minutes for each of these three critical elements of the framework was universally seen as a major challenge.

Principals were asked to provide feedback on the amount of training they received prior to participating in the pilot, the clarity surrounding the purpose of the pilot, and the instructions for implementing the pilot. Overall, most participants felt that they received just enough training on the Formal Observation Process prior to participating in this pilot. Nearly all participants felt that information regarding the purpose of the pilot was made clear and that instructions regarding how to implement this pilot were made clear (shown in Table 11).

Table 11. Principal Feedback: Training Provided for the Observation Pilot

Answer Options	Just enough training	Not enough training	Too much training	Response Count
How would you describe the amount of training you received on the Formal Observation Process prior to participating in the pilot?	17 (77%)	5 (23%)*	0 (0%)	22
Answer Options	**Agree**	**Disagree**	**Neither Agree nor Disagree**	**Response Count**
Information regarding the purpose of this pilot was made clear.	22 (96%)	0 (0%)	1 (4%)	23
Instructions regarding how to implement this pilot (process, forms, etc.) were made clear.	21 (91%)	1 (4%)	1 (4%)	23

* It is worth noting that at least two of the five principals who responded 'not enough training' did not receive training entirely as it was prescribed and intended. The training was delivered in four half-day sessions as part of a Cohort. One of these five principals attended two half days as part of Cohort I and two half days as part of Cohort II. This may have led to inconsistencies in the training, as different cohorts were taught by different instructors and may have had slightly different formats. Another principal attended only three of the four half day sessions.

Principals were asked to list up to three challenges they had faced so far in implementing this pilot. Nearly every principal listed time issues as a challenge. Specifically, principals noted the significant time involved in implementing the process with fidelity, difficulty coordinating common planning time when they could meet with teachers, and difficulty finding uninterrupted/undistracted blocks of time (30-45 minutes per conference) during the busy school day. Other challenges included difficulty finding volunteers to participate in the pilot and collecting Professional Personal IDs PPIDs from teachers who volunteered. These are both issues specific to the pilot process, which will not carry over into the actual implementation. Other principals expressed difficulty (both for them and the teachers) due to a lack of familiarity/comfort with the new tools and forms, as well as challenges for them in getting used to collecting strictly objective evidence during an observation. A few principals cited the change in mindset and dealing with the teachers' fear of something new as a challenge. These responses are shown in Table 12 below.

Table 12. Principal Feedback: Challenges in Implementing Observation Pilot

Answer Category	Response Count
Time Issues	21
Difficulty finding teachers to volunteer to participate	5
Lack of familiarity with the new instrument and forms for principal and teachers	4
Collecting PPIDs from teachers	3
Focusing on and recording evidence	2
Changing the mindset; fear of something new	2
Other	4

Principals were asked to list up to two questions they have about the Formal Observation process for teachers. The most common questions were regarding the process for an unsatisfactory teacher, how to arrive at a final rating for a teacher, and the observation process for a special education teacher. These responses are shown in Table 13.

Table 13. Principal Feedback: Questions during Observation Pilot

Answer Category	Response Count
What will the process be for teachers who are unsatisfactory?	4
How is a teacher's final rating calculated to determine the overall proficiency level?	3
How will the process be differentiated/adapted for special education teachers who are using scripted programs or involved in inclusion classes?	3
How to deal with 'prep payback issue'	2
Are PDPs still in place?	1
What happens when observer and teacher do not agree?	1
When will teachers receive official info from district?	1
How will teachers collaborate with other teachers?	1
Will PD be sustained?	1
Are all teachers being evaluated in 2013-2014?	1

Specialists

The District is experiencing a year of excessive turmoil and unrest, which has presented several challenges in implementing some of the Race to the Top Initiatives. During the 2012-13 school year, the district employed approximately 1,000 Specialists, 25% of whom were required to be trained during the 2012-13 school year, according to the RTTT grant requirements. In an effort to present a coordinated and well-organized training, the program office planned to facilitate separate trainings for each type of Specialist (i.e. a separate training for Counselors, School Psychologists, Nurses, Librarians, etc.), in order to personalize the training to the individualized frameworks and various responsibilities of each role. The largest group of Specialists was School Counselors, with approximately 390 schools counselors employed across the district, and the district planned to train all school counselors with a session on May 28, 2013, thus exceeding the 25% training requirement for Specialists. Unfortunately, as a result of the District's extremely dire financial situation, all School Counselors received notice in early May that they would be laid off at the conclusion of the school year. Presumably in large part because of this news, fewer than 50 school counselors attended the May 28th training, causing the District not to meet its RTTT requirements for training 25% of Specialists during the 2012-2013 school year.

Educator Effectiveness Technology Solution (Evaluation Monitoring Tool)

In Year 1, RTTT funds were budgeted to purchase an Evaluation Monitoring Tool that would automate the educator evaluation process and streamline educator ratings. The current rating systems and procedures consist of hand-written or typed rating forms administered to teachers by district administrators. The Evaluation Monitoring Tool would serve as the vessel by which the District would transition to an online system that provides access to observation data as well as overall rating data and links to resources to support educators in areas for growth relevant to goals set by teachers and administrators that are aligned with individual, school, and district-wide goals.

After a competitive bid process and a vote of acceptance, SDP selected NCS Pearson's Educator Development Suite (EDS) as the appropriate tool for this purpose. The three-year contract between NCS Pearson and SDP began in August, 2013.

The Pearson EDS supports a flexible, multiple-measures approach to the generation of educator effectiveness ratings and will help the District better measure, manage, mentor, and support their teachers—connecting strategic goals to educational standards and classroom activities to professional development opportunities and recommendations. The modular components of the EDS include:

- **Observation Tools** for capturing observation and perception feedback, whether through desktop or mobile devices.
- **Educator Reporting Tools** that aggregate data from external systems, as well as data gathered from our observation tools. This data is then used to generate educator effectiveness ratings that can be disaggregated to into a myriad of reports at the district, building, teacher, grade-level, and subject-level areas to inform and guide professional growth plans.
- **Educator Profiles** that let individual users track their progress in line with educator Frameworks, design their own professional growth plans, and access the District's online library of professional development resources with automated, intelligent PD suggestions based on educator reporting results.
- **Video Library** which contains roughly 2,000 research-driven professional development videos that are aligned with the teaching standards and present model teachers delivering current and relevant classroom and practices, as well as strategies for implementing many leading instructional programs and curriculum.

It is endeavored that the tool will be live by October 1 of the 2013-2014 school year, and that observers will exclusively use the tool to collect observation data (rather than using paper and pencil) for all observations beginning at the start of the school year and going forward. All principals and assistant principals have received new laptops through the Race to the Top grant for this purpose. Principals will receive an overview of the tool during the Leadership Conference week in the beginning of August, and training on how to use the tool is planned for the end of August.

PVAAS Teacher Specific Growth

As part of the State's education reform agenda committed to improving educator effectiveness, PDE is implementing new teacher and principal evaluation systems that take into account student achievement as a significant factor, as well as working to improve access to data that can be used to inform instruction. In order to implement the full evaluation system, including multiple measures of student performance, the State has used Race to the Top funding to enhance its data systems so that it is able to link teachers to their respective students, and is currently in the process of developing the direct teacher-student data linkages in its value-added model of student growth.

In the Spring of 2013, PDE extended an offer to all LEAs to participate in a one-time pilot of the PVAAS roster verification process[11] and system for PVAAS Teacher Specific Reporting. Participation in the pilot was the only opportunity for LEAs to experience the process, time, and resources needed to yield PVAAS Teacher Specific Reporting before it actually "counts" towards the first PVAAS 3-year rolling average in the teacher evaluation system. Participation in the PVAAS pilot would provide each LEA with information to plan for full implementation of the PVAAS roster verification process and system during 2013-2013. LEAs were able to choose their level of pilot participation, and could opt to include all schools in the LEA, some schools, or only one school, as well as which grade levels, subjects, courses, and number of teachers to include from each pilot school.

SDP opted to participate in the pilot with a selection of three schools that spanned all grade ranges from K-12.

The following were requirements and timeline for pilot participation:

- By March 1: Submit "Intent to Participate in Pilot" survey
- By March 15: Submit file of staff email addresses and PPIDs (required format)
- April 15-26: Participate in one roster verification training (https://pvaas.sas.com)
- April 29 - June 7: Complete PVAAS roster verification process
- May 13 - June 13: Provide feedback to PDE regarding PVAAS roster verification
- Fall 2013: Participate in one PVAAS Teacher Specific Reporting training
- Fall/Winter 2013: Provide feedback to PDE regarding PVAAS Teacher Specific Reporting

PVAAS Teacher specific reporting estimates the effect of a teacher's performance on the academic growth of a group of students, and provides reports based on the Education Value-Added Assessment System (EVAAS) methodology. The purpose of PVAAS Teacher Specific Reporting is three-fold:

- To provide a teacher specific growth measure to be used as part of Pennsylvania's Educator Effectiveness System
- To provide diagnostic feedback to teachers regarding their influence on the academic growth of students; and
- To provide data for teachers and administrators to guide discussions about a teacher's influence on the academic growth of groups of students.[12]

Teachers receiving PVAAS Teacher Specific Reporting are permanent or temporary professional employees who hold a valid teaching certificate and who have full of partial responsibility for content specific instruction of assessed eligible content as measured by PA's assessments (PSSA and/or Keystone exams), which may include teachers other than those who are the teacher of record.

[11] Roster verification is a local LEA process by which teachers and administrators document that teachers are linked accurately to students for the correct tested grade/subject/course for the proportion of time available to instruct each student (PVAAS Teacher Specific Reporting Pilot Guide).
[12] PVAAS Roster Verification and Teacher Specific Reporting Pilot Guide

The goal to provide PVAAs Teacher Specific Reporting based on accurate data will be accomplished via a PVAAS web-based roster verification system provided to LEAs by PDE. It is a secure, web-based system with double levels of authentication and login to ensure security of PVAAS Teacher Specific Reporting. All changes to any roster for an individual teacher are date and time-stamped to reflect who made the edit and when the edit occurred.

Elective Data/Student Learning Objectives

According the new educator rating system as defined by Act 82, twenty percent of a teacher's rating will be based on elective data, including measures of student achievement that are locally developed and selected by the school district from a list approved by PDE, including, but not limited to, the following:

- District-designed measures and examinations;
- Nationally recognized standardized tests;
- Industry certification examinations;
- Student projects pursuant to local requirements; and/or
- Student portfolios pursuant to local requirements.

Districts shall select and develop measures using a Student Learning Objective process, which is a process to document a measure of educator effectiveness based on student achievement of content standards. Student achievement can be measured in ways that reflect authentic learning of content standards, and SLOs are written to a specific teacher and a specific class/course/content area for which that teacher provides instruction.

According to guidelines established by PDE, these elective Student Learning Objectives (SLOs) should meet the following criteria:

- Specific: Specifies a statement of accomplishments to be achieved; outcomes are results-focused, not process-focused
- Measurable: Provides evidence that is measured by a valid and reliable approach
- Attainable: Attainable, realistic, ambitious, but achievable expectations
- Relevant: Related to professional and/or academic standards; supported by data related to student performance outcomes
- Time Bound: Uses at least two separate points in time in a given year

In teacher evaluations where there is no state assessment available (teachers who do not teach PSSA or Keystone assessed eligible content), SLOs will comprise 35% of the final rating. For teachers of assessed-eligible content (PSSA and Keystone courses), SLOs will comprise 20% of the final rating. Student Learning Objectives will be factored into a teacher's rating for the first time in SY 2014-2015, and districts will have SY 2013-2014 to select and pilot their SLOs. An example of a Student Learning Objectives template can be found in Appendix B.

Academic Coaching

The original budget proposed a coaching model that would include the following coaching support:

- **Principal Coaches (3 FTE)** to assist Principals in the understanding and practice of the evaluation process
- **Teacher Coaches (3 FTE)** to assist teachers in the understanding and practice of the evaluation process
- **Coaching Coordinators (2 FTE)** to be based at the Central Office

The Coaching structure was amended during Year 1, and five Instructional Coaches were hired in July, 2013. Coaching Coordinators will no longer be included in the program design. Principal Coaches have not been formally removed from the program design, but have yet to be hired.

The five Instructional Coaches who have been hired are all PA certified educators and have served in school-based teacher leader positions. They will serve as liaisons between teachers, administrators, schools and the school district, and provide support through an array of activities that are designed to build collective leadership and to continuously improve teacher instructional capacity and student learning. The five coaches will provide support to individual teachers using a coaching cycle model, as well as provide professional development on teacher effectiveness, the Danielson model, and areas of need that are identified based on observation results. Instructional Coaches will play a role that serves the following needs for teachers:

- Classroom Supporter: to increase the quality and effectiveness of classroom instruction through collaboration, co-planning, modeling, side-by-side teaching and effective feedback
- Instructional Supporter: to support the implementation of instructional strategies that are effective and that enhance student learning
- Curriculum/Content Facilitator: to promote the implementation of Common Core State Standards (CCSS)
- Data Coach: to ensure that student achievement data is used to drive instructional decisions
- Facilitator for Change: to engage teachers in reflective thinking and guide them towards reflective practices
- Professional Learner/Facilitator: to engage in opportunities for continuous learning in order to remain current on research based instructional practices, and to design and facilitate effective professional learning opportunities based on the expectations for professionals in the School District of Philadelphia
- Resource Provider: to enhance classroom instructional by providing a variety of resources that are geared toward teacher effectiveness and student achievement

Inter-rater Reliability

The initial Race to the Top LEA Grant Guidelines published by PDE state that in addition to the core professional development requirements, all supervisors must complete inter-rater reliability training to receive certification prior to using the Educator Effectiveness Instruments to evaluate staff performance. PDE partnered with Teachscape to develop the asynchronous training module to be available at no cost to the participant starting

2012/2013. However, this language was made obsolete in a Race to the Top LEA Grant FAQ document of March 19, 2013.

The purpose of certification through the Teachscape *Focus* system is to prepare and certify observers to conduct accurate and consistent evaluations of teacher practice at specific K-12 grade levels. PDE released 750 Teachscape Focus licenses to LEAs state-wide in September 2012, 550 in October 2012 and 450 in January 2013 on a first come, first-served basis. There were 52 licenses released to SDP principals/assistant principals, and of them only one passed the proficiency test to receive certification. The principal who received certification logged approximately 37 hours of online training. This included video-based training in all four domains of the Framework for Teaching, scoring practice with master-scored videos, and a proficiency assessment to measure ability to accurately score classroom teaching.

Although certification is no longer required by the State, SDP's RTTT program office has stated that it intends to require all observers to earn certification, and is working with the State to potentially procure Teachscape licenses for all principals at no cost to the District. However, the State has made it clear that any licenses that are purchased and assigned to the District that go unused will be charged to the District, at $399 per license.

Discussion and Recommendations

This section summarizes findings from Year 1 and highlights issues to consider for future years of the program.

Professional Development and Training

> *Classroom observations can be powerful tools for professional growth. But for observations to be of value, they must reliably reflect what teachers do throughout the year, as opposed to the subjective impressions of a particular observer or some unusual aspect of a particular lesson. Teachers need to know they are being observed by the right people, with the right skills, and a sufficient number of times to produce trustworthy results. Given this, the challenge for school systems is to make the best use of resources to provide teachers with high-quality feedback to improve their practice.*
>
> MET Project Policy and Practice Brief: Ensuring Fair and Reliable Measures of Effective Teaching

The District spent the majority of Year 1 training on the Educator Effectiveness system, with a focus on training observers on the observation process. Recommendations from the Gate's Foundation's Measuring Effective Teaching (MET) Study's Foundations of Observation report highlighted the most important part of prioritizing implementation on a tight timeline as making sure that principals "have a true understanding of the instrument", and that "the process is about helping people unlearn prior conceptions of teaching practice, as well as learn new ones." The district was wise to invest so much of Year 1 on thorough principal training, and should be encouraged to maintain this level of commitment to making sure that observers and teachers have a complete understanding of the processes and tools.

This training and pilot process were implemented during a difficult time, as a new collective bargaining agreement was being concurrently negotiated with the Philadelphia Federation of Teachers (PFT) local teachers'

union. Because of this, principals and teachers had several questions regarding any possible changes in the locally-decided observation system (such as the frequency of observations conducted, the differentiated supervision system, etc.) As Year 1 of training has concluded and negotiations are still underway, employees are expressing increasing urgency in understanding what the upcoming evaluation year will look like.

A primary finding while piloting the observation tool was that principals were not confident in their own ability to implement it with fidelity due to the large amounts of time involved. This is a significant issue, because success of the observation tool and system hinges on quality implementation, and "good tools that are poorly implemented will have little benefit" (MET Study, 2012).

Gathering Feedback for Teachers, a report published by the Bill & Melinda Gates Foundation as part of the Measures of Effective Teaching (MET) Project, emphasized the following six minimum requirements for high-quality classroom observations:

1. Choose an observation instrument that sets clear expectations
2. Require observers to demonstrate accuracy before they rate teacher practice
3. When high-stakes decisions are being made, multiple observations are necessary
4. Track system-level reliability by double-scoring some teachers with impartial observers
5. Combine observations with student achievement gains and student feedback
6. Regularly verify that teachers with stronger observation scores also have stronger student achievement gains on average[13].

In adopting the Danielson Framework for Teaching (FfT), the District has succeeded in meeting the first recommendation, as the FfT was one of five observation tools studied by the MET Project and determined to sufficiently define a set of teaching competencies and provide specific examples of the different performance levels on each.

The third recommendation, and its corresponding research, may have significant implications for SDP, as principal feedback suggests risk for the observation protocol not being implemented with fidelity. Clearly SDP is strained by limited resources; which means that it will be important for the District to ensure the reliability of its observation protocol while maximizing existing resources. Increasing the number of people who are trained and qualified to observe is one way to boost observation capacity (MET Study, 2013). Master teachers, instructional coaches, and supplemental digital video could allow multiple observers to view instruction. Furthermore, it may not be necessary for every observation to be equally long or comprehensive. A MET Study suggestions says that "teachers who have demonstrated basic skills could be the focus of more targeted observations aimed at higher levels of performance", and "systems may get a more complete picture of teacher practice if they have more frequent, shorter observations (ideally, by more than one person), rather than fewer longer ones".

[13] The second, fifth, and sixth recommendations will be discussed in detail in later sections.

In Denver, CO, teachers are observed four times each year, but only in two of those are they observed on all competencies in the system's instrument. During each of the other two, they are observed on two competencies representing specific areas of focus for the teachers. In Hillsborough Co. FL, the number of observations teachers receive each year is determined by the prior year's evaluation score, with those who receive the lowest ratings having 11 observations (including formal, informal, administrative, and peer) and those with the highest ratings having 5 observations.

Ultimately, the district must decide how to allocate time and resources to ensure reliable and useful classroom observations. Considerations such as how many lessons, of what duration, and conducted by whom are decisions that should be taken seriously and informed by what priority the district places on reliability considerations. Other factors, such as novice teacher status, prior effectiveness ratings, and an overall professional development strategy are all factors in defining an informed observation structure.

Educator Effectiveness Technology Solution (Evaluation Monitoring Tool)

In order to **improve** teacher effectiveness, we must first be able to **view** teacher effectiveness. The District is undergoing a massive culture shift in its move to electronically managed evaluations. However, the implementation of the Educator Development Suite (EDS) has brought to light several issues related to educator effectiveness and its implementation that have yet to be resolved at the conclusion of Year 1. In order to have the EDS configured to match local processes, these local processes and decisions need to be very clearly and definitively established, and many of them have not, which has contributed to significant delays in preparing to launch the EDS at the start of Year 2. The initial goal was for the EDS to launch for principal and teacher use no later than October 1, 2013, but at the time of this report, the project does not appear to be on target to reach that goal.

Many of these issues stem from the fact that the activities associated with the RTTT grant are inextricably linked with local policy decisions at the District level, and throughout Year 1 of implementation, the individuals responsible for those policy decisions have not been the same individuals responsible for implementing RTTT program activities, which has led to a disconnect between how to move forward in issues of teacher evaluation. For example, Act 82 mandates that every teacher receive, as part of his or her summative yearly evaluation score, a rating in each of four established Domains: Planning and Preparation, Classroom Environment, Instruction, and Professional Responsibilities. However, District policy, established in the Collective Bargaining Agreement between SDP and PFT dictates that a teacher receives a 'formal' observation only every third year. It is yet to be decided how a teacher will receive a rating in the four previously listed Domains every year without receiving a for-stakes observation in that year.

The lack of communication may be partially a result of significant staffing changes across the district through Year 1 of the RTTT grant, and until recently, a lack of clear leadership for who is responsible for issues related to teacher effectiveness. In August 2013, the District hired both a Deputy of Teacher Effectiveness and a Chief Human Resources Officer who will provide necessary leadership in this area. Prior to this time, this gap in

staffing presented significant challenges to critical decision-making required to continue moving this work forward.

Furthermore, SDP and PFT have been in ongoing contract negotiations throughout the past year, as the current collective bargaining agreement expires in August 2013. The negotiations have presented challenges in determining decisions related to negotiated items associated with educator effectiveness, particularly as it relates to teacher observation schedules and protocol.

Once the District has been able to arrive at several critical decisions and configure the EDS accordingly, principals and teachers will be trained on how to utilize the system. However, the program office has not yet established clear guidelines for how principals and teachers are expected to utilize the system once it is launched. For example, if they will require 100% utilization for all teacher evaluation data during the 2013-2014 school year, or if there will be a gradual rollout with some leniency for those who are slow to transition to using the EDS. It is recommended that the program office require 100% utilization in order to ensure consistency, but whatever is decided will need to be clearly defined and communicated to principals and teachers, including the consequences associated with non-compliance.

Moving forward, even the strongest evaluation system will not be particularly useful if it fails to provide supports for teachers to improve their practice. The information gained from measuring needs to be applied to professional ratings, human capital management, and progress monitoring to deliver better support and professional development. The greatest promise of a multiple measures system of evaluation lies in its use as a development tool, and for that promise to be realized, professional development will need to be individualized to meet teachers' specific needs. There has been minimal research published on the impact of professional development models which are explicitly aligned with teachers' evaluation results, however, what evidence is emerging suggests that individualized feedback to teachers can lead to better outcomes for students.[14]

As the program continues through Year 2 of implementation, and begins to collect robust teacher and principal evaluation data, program offices should work with research staff to identify methods of aligning individual, school, or district-wide professional development to targeted and innovated methods of professional development aligned with work-force needs.

PVAAS Teacher Specific Growth

As stated by Meyer and Dokumaci (2010), one of the main factors affecting the quality of any value-added system is "the availability and quality of longitudinal data on students, teachers, and schools, particularly the degree to which students, classrooms/courses, and teachers are correctly linked."[15] The accuracy of this input data also impacts the degree of acceptance of these models by those who are affected by them, specifically teachers.

[14] "Feedback for Better Teaching." Bill and Melinda Gates Foundation. Web. http://www.metproject.org
[15] Meyer, R. H. & Dokumaci, E. (2010). Value-added models and the next generation of assessments. Princeton, NJ:Educational Testing Service.

There is extensive debate and research on the validity and reliability of value-added models, however as states and districts have begun to tie teacher value-added data to high stakes accountability decisions such as teacher evaluation, retention, and dismissal, the quality of the input data used in these calculations has also become an issue. The problem of incorrect student-teacher linkage data has begun to be recognized by various stakeholders.

In a lawsuit against the New York City Department of Education's release of teacher data reports, the United Federation of Teachers (UFT) claimed:

> New York City teachers have found multiple mistakes in their reports, including reports on students and even entire classes which the teachers never taught. Other inaccuracies include the addition of students who were taught for part of the year by a different teacher, inconsistencies in accounting for students who need special help, and other issues with data collection and evaluation[16].

In a discussion of teacher performance pay, Burns and Gardner (2010) explain that issues of data quality must be addressed:

> The data obtained from district and state information systems also are often riddled with inaccuracies and errors that can wreak havoc on the operation of a performance pay program. In an analysis of data system quality, Battelle for Kids documented several common weaknesses, including too few data snapshots (data collected only once or twice a year does not accurately capture what is happening in schools), inaccurate course codes, errors with the unique student identification number (multiple students with the same number) and incorrect student-teacher linkages.[17]

Preliminary analyses using real data conducted by the Value-Added Research Center at the University of Wisconsin-Madison indicate that incorrect student-teacher linkages can cause severe errors in teachers' value-added scores.

In implementing the new teacher evaluation system, it is critically important that Philadelphia examine the capabilities of its data systems and technical infrastructure to ensure that they have the necessary capability to capture student-teacher linkages in the most accurate way possible. In addition to ensuring the district's technical capacity, the district must make every effort to effectively communicate to principals, teachers, and other stakeholders the importance of the accuracy of the data used in the roster verification process, and provide guidance for how they can assure the quality of the submissions that will be used in their evaluation. This will require the district to first establish internal guidelines, and then distribute them in a way that is easy for the non-technical audience to understand and use. The district should also implement a robust technical support plan during the spring roster verification windows so principals and teachers know where to go for help, should they have questions or issues.

[16]UFT to file suit to prevent release of incorrect teacher data. UFT press release, October 20, 2010, retrieved 7/15/13 at http://www.uft.org/press-releases/uft-file-suit-prevent-release-incorrect-teacher-data.
[17] Burns, S. F. and Gardner, C. D. Reforming Teacher Pay. The School Administrator March 2010,Number 3, Vol. 67, 15-20.

Elective Data / Student Learning Objectives (SLOs)

Identifying highly effective teachers of subjects, grades and students who are not tested with standardized achievement tests, such as teachers of art, music, physical education, foreign languages, K-2, high school, English language learners, and students with disabilities, can be one of the most challenging aspects of including student achievement and/or growth in teacher evaluation.

Most states and districts are considering a variety of assessment types to provide measures of student performance in non-tested grades and subjects. An analysis conducted by Katie Buckley of Harvard University and Scott Marion of the National Center for the Improvement of Educational Assessment has surveyed the existing approaches used to evaluate educators in non-tested grades and subjects and grouped commonly used measures into four categories:

1. Externally created *norm-referenced tests* (NRT), such as the Stanford-10 or Terra-Nova, and including standardized exams created for special populations, such as ACCESS for ELL students;
2. Externally created *interim assessments* such as Dynamic Indicators of Basic Early Literacy Skills (DIBELS) or Measures of Academic Progress (MAP);
3. National, state or district administered *end-of-course exams* that are standardized, such as the Advanced Placement (AP) exam or the New York Regents assessment; and
4. The use of *school- or teacher-developed measures,* including such tools as locally created end-of-course tests, common performance tasks or other curriculum-embedded assessments, and student portfolios.

One approach for incorporating student data in teacher evaluations for non-tested grades and subjects is to create and administer more tests in subjects and grades where there are not current large-scale tests (such as the PSSA and Keystone). States such as Delaware and large districts such as Hillsborough, FL and New York City are pursuing this avenue (Marion & Buckley, 2013). While this may seem like a reasonable approach, creating a testing program is difficult, and maintaining **high quality** testing programs is even more challenging (Marion & Buckley, 2013).

Using tests for such high stakes purposes as educator accountability requires very high standards of technical quality, including “a level of reliability necessary to support high sakes decisions, items and forms that meet content validity standards, and technically appropriate linking designs that ensure that scores across years or forms can be validly placed on the same scale” (Marion & Buckley, 2013). The State of PA has invested in establishing these levels of quality in order to use PSSA and Keystone data for the Teacher Specific Data portion of the evaluation. However, to do so at the district level for use in the Elective Data/SLO portion would require a robust internal psychometric capacity or contracting with a test development company or consultant, introducing significant ongoing costs. Furthermore, the potential to undertake such an effort does not promise any significant return on investment, and few would argue for such an increase in additional testing to begin with(Marion & Buckley, 2013).

Student learning objectives offer advantages over other the previously discussed methods of analysis and measurement for several reasons. They can be highly flexible in that they can be used across all grades and subjects with existing measures of performance or adapted to new assessment systems as they are developed. Furthermore, SLOs are designed to incentivize the positive practices of setting empirically-based goals for each student (or the class), monitoring the progress toward these goals, and then evaluating the degree to which students met the intended targets (Marion & Buckley, 2013).

The RTTT Technical Assistant (TA) Network defines SLOs as "a participatory method of setting measurable goals, or objectives, based on the specific assignment or class, such as the students taught, the subject matter taught, the baseline performance of the students, and the measureable gain in student performance during the course of instruction" (2010).

Marion & Buckley (2013) have outlined the following criteria specific to establishing Student growth objectives:

1. Each district shall develop a set of procedures for establishing and evaluating goals. These procedures shall include general district approaches as well as providing guidance for specific content areas.
2. Goals shall be established for each student and at the aggregate classroom level, such that individual students are ambitious and standards-based, while aggregating goals may be normative. They strongly suggest having aggregate goals focus on the full range of students rather than the simple class average.
3. Goals shall be based on data such as prior assessment/grade history and must reflect meaningful (e.g. college readiness) and measurable targets.
4. Multiple goals may be established for each student, but at least one of the goals shall be a long-term goal (e.g. a semester or year) in order to have a greater chance of detecting real change.
5. Goals shall be set by teachers in consultation with professional learning communities, a committee of peers, and/or principals. Goals should be made public, at least internally to other educators in the school and parents.
6. Progress toward and attainment of goals shall be determined by measures that are aligned with the learning targets and are technically appropriate to determine whether students have actually met the goals. In other words, a case in which the assessment should be avoided is if it is only least nominally aligned with the targets and at a level far below the actual goals so that one is unable to actually judge if the student met the goal.
7. The assessments used to measure the goals shall be reviewed by a committee of peers and administrators to judge their adequacy for evaluating student progress toward the goals.

Source: *Marion, Scott, and Katie Buckley. "Approaches and Considerations for Incorporating Student Performance Results From "Non-Tested" Grades and Subjects into Educator Effectiveness Determinations." National Center for the Improvement of Educational Assessment, 7 Sept. 2011. Web. July 2013. <http://www.nciea.org/publications/Considerations%20for%20non-tested%20grades_SMKB2011.pdf>.*

There are several limitations and challenges ahead in implementing locally-established SLOs in educator evaluation. For one, SLOs can only be as good as the quality of the goals set for each student and by the quality

of the measures used to evaluate the goals (Marion & Buckley, 2013). The district will need to make significant investments in establishing specific and aligned guidelines for SLOs as well as in professional development requirements for educators in order to be able to create the learning objectives, ensure that the performance goals set are attainable yet rigorous, and develop or select the appropriate measures for the goals. It is also important to recognize what is known as Campbell's Law – the potential for corruptibility for any quantitative indicator used in high stakes decision making. It is critical that the design of this system component recognize these threats and acknowledge them appropriately.

Marion & Buckley suggest that the use of SLOs as elective data in educator evaluation systems may "foster an internal locus of control in that educators would feel like they have more control over their evaluations compared to externally delivered results" as well as "promote a sense of fairness" in that all teachers within each school would be operating within the same framework and would all be responsible for designing and evaluating SLOs.

Although it is not listed as a State-approved form of Elective Data, it is important to acknowledge the momentum that has been gaining in recent years for utilizing student perception surveys as part of instruments used to inform and impact teacher ratings. Administration of such surveys has taken place district-wide in Denver, Memphis and Pittsburgh, and piloted state-wide in Georgia and North Carolina.

Research on the use of student perception surveys in K-12 education has not been extensive; however, studies consistently suggest that student surveys are a reliable measure of teacher effectiveness (Hanover Research, 2013). Survey designs have established content validity by developing research-based, relevant questions and structuring surveys around core constructs related to high-quality teacher attributes. Studies have shown that student surveys can accurately predict student achievement gains, are more accurate than other widely used instruments such as teacher observation tools alone, and were second only to previous test-score gains in predicting a teacher's ability to increase test scores.[18]

A research synthesis conducted by the National Comprehensive Center of Teacher Quality offers a number of advantages to school districts using student perception surveys, such as:

- Cost and time efficient;
- Can be collected anonymously;
- Require minimal training;
- Enable tracked changes over time; and
- Provide valuable feedback to teachers.[19]

[18] Ferguson, R. F. "Student Perceptions of Teaching Effectiveness." National Center for Teacher Effectiveness and the Achievement Gap Institute, Hardvard University. October 14, 2010, p.2.

[19] Goe,L.,Bell,C., and Little,O. "Approaches to Evaluation Teacher Effectiveness: A Research Synthesis." National Comprehensive Center for Teacher Quality. June 2008.

There are some potential disadvantages to using student surveys to rate teachers, primarily that survey results, like any quantitative measure, may be affected if they are used in high-stakes environments. One criticism from Michael Mendel of the United Federation of Teachers claims that student surveys would incentivize teachers to strive for classroom approval.[20]

Developing and validating an effective student perception survey is resource and research intensive, and because of that, many states and districts are purchasing existing surveys rather than creating their own. Two off-the-shelf surveys are specifically designed for teacher evaluation: The Tripod Survey and My Student Survey. The former, developed at Harvard, which is comprised of evidence-based questions, has been shown to reliably predict student achievement gains and has received support from the Bill & Melinda Gates Foundation's MET Project. The latter was developed by Vanderbilt researcher Ryan Balch, and has been shown to effectively relate high teacher ratings with high student performance. Both can be completed in under 30 minutes, are offered in print and online, and come with additional services such as logistics support and results analysis. Sample questions from the Tripod Student Survey and the My Student Survey can be found in Appendices C and D.

Research on the Tripod survey has determined that students are just as capable of identifying effective teaching as principals and other classroom observers. The Measures of Effective Teaching Project (MET), a research study funded by Bill & Melinda Gates Foundation to explore the strength of multiple measures, found that student surveys have an equally strong correlation to predicting student learning and growth as compared to classroom observations by trained evaluators. And because students spend time with their teachers many days throughout a year, student survey results are more reliable than the few classroom observations that take place each year.[21]

As of the 2012-2013 school year, Memphis City Schools was the only district identified that **requires** student perception surveys be factored into high-stakes performance reviews. Chicago Public Schools plans to count student feedback as 10% of teacher evaluations beginning in 2013. Some districts give teachers the choice of factoring surveys into their reviews. In Utah, David School District offers a student perception survey as an optional measure and Massachusetts plans to have an optional survey in place by the 2013-2014 school year.[22] Georgia, which piloted My Student Survey determined that student surveys would count for 10% of a teacher's evaluation in tested subjects and 40% for teachers in non-tested subjects.[23] Thomas Kane, a MET project coordinator attests that surveys should count for 20 to 30% of a teacher's evaluation, as he believes this is enough for teachers and administrators to take them seriously, but not enough to tempt teachers to cheat or pander to students.[24]

---

[20] "Student Perception Surveys and Teacher Assessments." Hanover Research. http://scee.groupsite.com/uploads/files/x/000/08f/0fb/Student%20Perception%20Surveys%20and%20Teacher%20Assessments%20-%20Membership%20%282%29.pdf.

[21] 'Measures of Effective Teaching." Bill & Melinda Gates Foundation. http://metproject.org

[22] Balch, R. "The Validatio9n of a Student Survey on Teacher Practice." Vanderbilt University. 2012. http://mystudentsurvey.com/wp-content/uploads/2012/06/Balch-Student-Surveys-2012.pdf

[23] Balch, R. Op. cit., p.6.

[24] "Ensuring Fair and Reliable Measures of Effective Teaching: Culminating Findings from the MET Project's Three-Year Study." Bill & Melinda Gates Foundation. January, 2013.

The MET Project researched the value of multiple measures of teacher effectiveness and found that although each measure has different strengths and weaknesses, the combination of three measures – teacher practice, student learning and growth, and student perception – results in the highest reliability and predictive power. These results are shown in Table 15 below.

Table 15. Table of the MET Project's findings on strengths and weaknesses of each measure

Tool for measuring teacher effectiveness	Predictive Power	Reliability
Classroom observation	Low	Medium/High
Value-added	High	Medium
Student surveys	Medium	High

Source:

All districts identified as having piloted or implemented student surveys administer them one or two times per year. Denver Public Schools (DPS) administers student surveys once in the fall, based on findings from the pilot:

- Student responses do not vary significant between fall and spring. DPS conducted surveys in both fall and spring, and found no statistical difference between the two;
- Surveys conducted only a few weeks into the school year were valid and consistent with surveys taken at other times; and
- Administering surveys early in the school year gives teachers time to view results and adjust their practice.[25]

Balch (2012) asserts that "there is no evidence on the number of times a teacher should be rated by their students each year", but speculates that "multiple surveys in one year could provide more reliable estimates and also reflect growth during the year."[26]

Including measures of student performance and student perception in educator evaluations is a very new endeavor. As such, regardless of the measures that are selected to serve as the Elective Data/SLOs, ongoing, formative evaluations must be put into place alongside new educator evaluation systems as they are piloted and through the first several years of implementation so that we are able to learn what is working well and what needs refinement or reconsideration (Marion & Buckley, 2013).

[25] "Frequently Asked Questions: New 2012-13 Pilot." Leading Effective Academic Practice, Denver Public Schools. http://leap.dpsk12.org/Resources/FAQs.aspx.

[26] Balch, R.

Academic Coaching

There is a growing interest in coaching, mentoring and peer-networking as means to enhance professional development. Mentoring may enhance individual, team and organizational performance through sharing and developing practice within an atmosphere of mutual trust and respect.[27]

It is unclear at this point exactly how the academic coaching model will be structured. Five academic coaches have been hired to support a full teaching staff of more than 9,000 educators, who are segmented into eight Professional Learning Networks. Roles and distribution of services will need to be clarified and defined, as it is currently unclear which teachers will receive coaching support, with what frequency, and how that coaching support will be structured and monitored.

A report published by The Center of American Progress has articulated the features of effective professional development as being highly targeted on relevant knowledge and practice, with close alignment between assessment (of teacher practice or child outcomes) and professional development supports intended to produce those practices and outcomes[28]. Yet even for approaches that do specifically articulate how and why they might work, such as coaching models, the challenge is scaling up: How can effective programs be delivered to large numbers of teachers? Even for professional development approaches that work with a few dozen teachers or a couple of schools, there is little evidence that they can be implemented in a standardized manner across teachers so that many thousands can improve their work systematically at the district-level.

If the District is to move forward for also providing coaching support for principals, that model will also need to be identified and fleshed out. One option is to use Observation Coaches to monitor the quality of the District's teacher observations by assigning coaches to monitor and mentor observers. This would still require the program office to delineate the score of tasks they will perform and the amount of time they are expected to devote to each task. Recommendations from the MET Study suggest that coaches selected to fill this role should have:

- Demonstrated expert-level observation skills
- The ability to effectively communicate rationales to support score decisions to observers
- An understanding of how to motivate observers to put aside personal scoring standards and adopt the standards ascribed by the scoring rubrics.[29]

[27] Rhodes, Christopher, Michael Stokes, and Geoff Hampton. *A Practical Guide to Mentoring, Coaching, and Peer-networking: Teacher Professional Development in Schools and Colleges*. London: RoutledgeFalmer, 2004. Print.

[28] Pianta, Robert C. "Teaching Children Well: New Evidence-Based Approaches to Teacher Professional Development and Training." Center for American Progress, Nov. 2011. Web. 1 Aug. 2013. <http://www.americanprogress.org/issues/education/report/2011/11/29/10663/teaching-children-well/>.

[29] Joe, Jilliam N., Cynthia M. Tocci, Steven L. Holtzman, and Jean C. Williams. "Foundations of Observation: Considerations for Developing a Classroom Observation System That Helps Districts Achieve Consistent and Accurate Scores." Bill & Menlinda Gates Foundation, 2013. Web. July 2013. <http://www.metproject.org/downloads/MET-ETS_Foundations_of_Observation.pdf>.

Inter-rater Reliability

> *Inaccurate classroom observations lead to mistrust and poor decisions. Ensuring accuracy is not just a matter of training. It requires assessing observes' ability to use the instrument at the end of the training. Moreover, it may be necessary to ask observers to redemonstrate periodically their ability to score accurately.*
>
> MET Project Policy and Practice Brief: Ensuring Fair and Reliable Measures of Effective Teaching

At the time of this report, SDP has not outlined a specific plan for implementing the inter-rater reliability requirement. The MET study identifies this type of certification as a critical element in implementing quality classroom observations with fidelity. Requiring observers to demonstrate that they can apply an observation instrument accurately and fairly before they rate teacher practice is cited as one of six minimum requirements for high quality classroom observations (MET Project, Policy and Practice Brief, January 2012). The report findings suggest that observers be expected to demonstrate their ability to generate accurate observations and should be recertified periodically.

Should this element remain part of the program design, it is imperative that plan be developed soon that addresses the inter-rater reliability requirements for principals. Should this be a mandatory requirement for observers, this will need to be addressed and negotiated with the principals union, a proposed timeline laid out, and a strategic implementation plan put into action.

Other States and Districts are using the Teachscape *Focus* product for inter-rater reliability and certification. The Hartford Public School District (HPS) has purchased a 3-year contract with Teachscape using RTTT funds and has negotiated with their union to make it a job responsibility and requirement that observers pass inter-rater reliability certification. HPS implemented the system in the summer of 2012 for its 130 administrators, who have until June 30, 2013 to earn certification, otherwise they lose their ability to conduct observations. According to HPS's Director of Performance Management, about 80 of the 130 administrators achieved certification easily, though it was stressful for them to get through the 30-45 hours of training online as well as a 3-6 hour proficiency test. As of April 2013, all but 2 or 3 administrators had received certification. Going forward, HPS will be implementing calibration and re-certification program, which has also been negotiated with the union.

In Memphis City Schools, observers are trained for two-days on the teacher effectiveness observation rubric, and then practice scoring independently over approximately three weeks, followed by a three-hour refresher session and a certification assessment.

The Teachscape *Focus* tool has been established as the observer certification instrument aligned with the Danielson Framework for Teaching. However, should procuring licenses for all observers become cost-prohibitive or if Teachscape's tool does not measure the full range of observer competencies deemed important, the District may consider developing its own certification test to assess how well observers are calibrated to the observation rubric.

Findings from The MET Project's study on ensuring high quality data collection during teacher observations "highly recommends" that uncertified observers do not provide observation scores for teachers until they have demonstrated proficiency. If the district moves forward in implementing a certification requirement in order to be a qualified observer, the study offers the following considerations and suggestions:

- **How long will the observer be qualified to provide scores before certification is required?**
  The following approaches may be considered until uncertified observers reach the desired level of proficiency before they take another certification test:
  a) Assigning a scoring expert as a coach to guide the observer through the areas of the instrument and rubrics that he or she does not understand well or applies inaccurately;
  b) Having the coach explore the possible influence of bias and personal preferences with the observer; and;
  c) Assigning a certified observer to pair up with the uncertified observer during live observations.
- **How will the District provide support to principals who have not yet passed?**
- **What happens when a principal fails the certification exam? If he/she is the school's only assigned observer, how will the school proceed with teacher observations?**
  The following are suggestions for developing contingency plans until the principal becomes certified, such as:
  a) Sending another certified observer in with the principal so two observers are in the classroom at the same time; and
  b) Videotape the lesson so the principal observes live, and a second certified observer can watch the video later, while using the opportunity to discuss the performance with another trained observer.
- **How much of the certification data will the District release and to whom?**
  The District will need to decide how much data to release to test-takers, as well as how much to release to the public, considering the following:
  a) *Pass-Fail Status* – at a minimum, this must be released to observers so they know whether they passed the test or need to take it again. It is important to take into consideration that if an observer fails and is asked not to perform any observations until they pass, this information will eventually become known to teachers, even though it is not published.
  b) *Actual Scores* – Releasing actual score information to observers will allow them to respond appropriately to their scores. On the other hand, observers may be tempted to share or discuss their answers with colleagues who plan to take the test, which may lead to problems with the integrity of the items used. Publically releasing information about individual performance is not recommended, however releasing summary statistics about observation scores may provide credibility for the district's observation system or help teachers understand how well trained the district's observers are in the observation instrument.
  c) *Performance on Individual Dimensions* – Releasing information on how candidates perform on individual components can help observers identify components they were discrepant on, to allow for further review.
- **How long will the observer be qualified to provide scores before recertification is required?**

Observer accuracy may slide after initial training and certification. In determining how often to reassess observers, districts will need to weigh the cost of calibration against the benefits in terms of accuracy and trust.

Whether or not the District decides to implement inter-rater reliability certification for observers, it should also be conscious of other ways in which observations can be prone to error over time, and implement safeguards to monitor observer behavior and ensure the quality of teacher observations is maintained over time. The following list of observer effects have been highlighted by the MET Study as important factors to monitor, identify, and attempt to minimize in order to maintain high levels of reliability and validity of observation score interpretation:

- Familiarity Bias: The personal and professional relationships that observers have with teachers they are observing can lead to more lenient or more severe scoring, and affect perceptions in ways that may or may not be obvious to the observer. The greater the level of objectivity and transparency in the observation process, the more confidence teachers are likely to have in scores and feedback. One or more of the following procedures can guard against familiarity bias:

    - To the extent possible, have an independent within-district observer available if a principal feels he or she can't provide an unbiased observation;
    - Build in quality checks by capturing a subset of the observations on video and having a second person or team review for accuracy; and/or
    - Make sure principals and observers focus only on evidence and monitor the influence of their professional preferences and relationships with the teachers they observe.

- Drift: A shift in the overall direction of an individual or group of observers' scoring over time toward greater leniency, severity, or accuracy.

    - Individual drift can be detected by periodically assigning observers to score a set of observations that has been vetted by an expert.
    - Group drift can be identified through *trend scoring*, when a group of observes score a set of observations that were scored by the same or an equivalent group of observers in the past.

- Halo and fatal-flaw effects: Trait carryover effects in which one salient trait or feature of an observed teacher influences the observer's overall judgment. When the domain comprises related but distinct traits, the observer who is influenced by these effects fails to judge individual traits by their own merit.

    - Halo effect is when observers make this error in the direction of higher than deserved scores (because A is good, then B and C must also be good)
    - Fatal-flaw effects is when lower than deserved scores are awarded (because A is weak, B and C must also be weak)

- Central tendency effect: When the scores of observers tend to cluster in the middle of the score scale out of the observer's reluctance to score at the more extreme ends of the scale. Possible causes include:

    - An observer's lack of confidence that he or she knows how to recognize performance at the tails;
    - An observer's fear of the consequences of awarding such score points, such as an effect on school culture;
    - An observer's belief that the performance do not exist or are extremely rare;
    - An observer's distorted perception of these performances as "fairytale" performances because they bring to mind such extremes; and/or
    - An observer's decision to "play it safe" because scores in the middle tend not to raise red flags.

Source: *Foundations of Observation: Considerations for Developing a Classroom Observation System that Helps Districts Achieve Consistent and Accurate Scores*

Conclusions

Throughout Year 1 of implementation, the District has been successful in ensuring that principals and school leadership teams have been trained on the upcoming changes to state-mandated teacher, principal and specialist evaluation that will go into effect in the 2013-2014 school year, and also has a well-established plan to train all teachers on the same materials at the start of the 2013-2014 school year, even though the goal to train 50% teachers in Year 1 was not met.

The most significant implementation challenges during Year 1 have been the result of unclear or shifting leadership in charge of critical decisions associated with changes to educator effectiveness practices introduced as a result of Act 82. Many of the activities associated with the RTTT grant are inextricably linked with local policy decisions at the District level, and throughout Year 1 of RTTT implementation, the individuals responsible for those policy decisions have not been the same individuals responsible for implementing RTTT program activities, leading to a disconnect between programmatic decisions and larger district decisions. Ongoing collective bargaining agreement negotiations have also made it difficult for the program office to anticipate or establish changes to teacher evaluation practices.

Now that the District has brought in a Deputy in charge of teacher effectiveness as well as a Chief of Human Resources, those individuals should work closely with both the RTTT staff and District leadership to align long-term district strategy with shorter-term programmatic activities. Because most of the grant requirements were established by the State for Year 1, and speak to only the professional development and training components of the grant, it is important that the educator effective and RTTT teams work with ORE staff to establish local programmatic goals for Years 2 and 3 of implementation, as well as long-term anticipated impact that is clearly defined and measurable. This should include further defining the program design and activities for the next two years of the grant, such as activities related to coaching and inter-rater reliability.

The RTTT and educator effectiveness teams must continue to work closely with Information System as well as with the State's PVAAS resources to implement the PVAAS Teacher Specific Growth measure and roster verification process with the highest level of accuracy and fidelity.

It is important that the RTTT program team and the educator effectiveness team also work closely with district leadership to begin defining a process for establishing Student Learning Objectives that align with the District's vision, mission and Action Plan.

Through the course of implementing these new initiatives and changes, communication with major stakeholders, including principals, teachers, union partners, and parents should be clear, frequent, and easily accessible through a variety of channels.

References

Balch, R. “The Validatio9n of a Student Survey on Teacher Practice.” Vanderbilt University. 2012. http://mystudentsurvey.com/wp-content/uploads/2012/06/Balch-Student-Surveys-2012.pdf

Burns, S. F. and Gardner, C. D. Reforming Teacher Pay. The School Administrator March 2010,Number 3, Vol. 67, 15-20.

Danielson, Charlotte. "Evaluations That Help Teachers Learn." *The Effective Educator* 68.4 (2011): 35-39.*www.danielsongroup.org*. Dec.-Jan. 2010/2011. Web.

Duncan, Arne. "Partners in Reform." Address by the Secretary of Education to the National Education Association. 2 July 2009. Speech.

Doyle, D., & Han, J.G. (2012). *Measuring teacher effectiveness: A look “under the hood” of teacher evaluation in 10 sites.* New York: 50CAN; New Haven, CT: ConnCAN; and Chapel Hill, NC: Public Impact. Retrieved from http://www.concan.org/learn/research/teachers/measuring-teacher-effectiveness

Ensuring Fair and Reliable Measures of Effective Teaching: Culminating Findings from the MET Project’s Three-Year Study.” Bill & Melinda Gates Foundation. January, 2013.

Feedback for Better Teaching. Bill and Melinda Gates Foundation. Web. http://www.metproject.org

Frequently Asked Questions: New 2012-13 Pilot.” Leading Effective Academic Practice, Denver Public Schools. http://leap.dpsk12.org/Resources/FAQs.aspx.

HB 1980 - Educator Evaluation System, House Education Committee Cong. (2011) (testimony of Carolyn Dumaresq, D. Ed., Deputy Secretary for Elementary and Secondary Education). Print.

Joe, Jilliam N., Cynthia M. Tocci, Steven L. Holtzman, and Jean C. Williams. "Foundations of Observation: Considerations for Developing a Classroom Observation System That Helps Districts Achieve Consistent and Accurate Scores." Bill & Melinda Gates Foundation, 2013. Web. July 2013. <http://www.metproject.org/downloads/MET-ETS_Foundations_of_Observation.pdf>.

Meyer, R. H. & Dokumaci, E. (2010). Value-added models and the next generation of assessments. Princeton, NJ:Educational Testing Service.

Pianta, Robert C. "Teaching Children Well: New Evidence-Based Approaches to Teacher Professional Development and Training." Center for American Progress, Nov. 2011. Web. 1 Aug. 2013. <http://www.americanprogress.org/issues/education/report/2011/11/29/10663/teaching-children-well/>.

Professional Growth System Handbook.2012-2013 Edition.The School District of Philadelphia in Partnership with the Philadelphia Federation of Teachers. https://cc.philasd.org/service/home/~/Professional%20Growth%20System%20Booklet%208%2024 %2012.doc-4.pdf?auth=co&loc=en_US&id=43013&part=2

Sartain,Lauren, Sara Ray Stoelinga, and Eric R. Brown. "Rethinking Teacher Evaluation in Chicago: Lessons Learned from Classroom Observations, Principal Teacher Conferences, and District Implementation." N.p., Nov. 2011. Web. June 2013. <http://ccsr.uchicago.edu/sites/default/files/publications/Teacher20Eval20Report20FINAL.pdf>.

Student Perception Surveys and Teacher Assessments.Hanover Research. http://scee.groupsite.com/uploads/files/x/000/08f/0fb/Student%20Perception%20Surveys%20and%20Teacher%20Assessments%20-%20Membership%20%282%29.pdf.

Race to the Top Executive Summary http://www2.ed.gov/programs/racetothetop/executive-summary.pdf

Race to the Top Pennsylvania Report Year 1: 2012 www2.ed.gov/programs/racetothetop/performance/pennsylvania-year-1.pdf

Rhodes, Christopher, Michael Stokes, and Geoff Hampton. *A Practical Guide to Mentoring, Coaching, and Peer-networking: Teacher Professional Development in Schools and Colleges*. London: RoutledgeFalmer, 2004. Print.

[1]UFT to file suit to prevent release of incorrect teacher data. UFT press release, October 20, 2010, retrieved 7/15/13 at http://www.uft.org/press-releases/uft-file-suit-prevent-release-incorrect-teacher-data.

Appendices

THE SCHOOL DISTRICT OF PHILADELPHIA
In partnership with
THE PHILADELPHIA FEDERATION OF TEACHERS

SPECIAL OBSERVATION STATUS

Special Observation Status (SOS) for Tenured Teachers not in Formal Observation Year
Rationale: If a principal has concerns about the classroom performance of a tenured teacher who is not currently in a formal evaluation year, s/he may request that the PAR Panel place the teacher on Special Observation Status (SOS). The request for SOS does not remove the teacher from writing or implementing the scheduled Professional Development Plan. SOS is not subject to appeal.

Steps:

- Principal or Assistant Principal must conduct at least three informal observations anytime between **September 1st and December 31st** and will provide written feedback to the teacher after each informal observation.
- If Principal or Assistant Principal's three informal observations note concern in the area of Domain II and/or Domain III found in the Formal Observation Tool, the principal will compile the evidence of these informal observations.
- Principal will meet with the teacher to inform her/him of these concerns review the evidence of the three informal observations and, inform teacher of the intent to request Special Observation Status.
- Teacher has the right to bring union representation to the SOS meeting.
- Principal completes Request for Special Observation Status form and a one-page narrative summary of informal observations with Domain II/or Domain III found in the Formal Observation Tool, including the recommendations and supports put in place during this time period. The teacher will receive a copy of these documents no later than five (5) working days after the meeting with the principal.
- Documentation should be sent to the Administration's PAR Co-Chair after the above steps have been taken, but not later than **January 10th**.
- PAR Panel or PAR Panel sub-committee, as designated by the PAR Panel, will review documentation and review evidence to support request.
- PAR Panel or designated PAR Panel sub-committee will respond to request within five (5) working days.
    - If Principal's paperwork is not completed, or does not meet deadline, or does not focus on Domains II or III, the request will not be considered.
    - If Principal's paperwork is completed, the following will occur:
- The Principal and teacher will receive a letter from the PAR Panel informing them of their decision to place the teacher in a Special Observation Status. The letter will include an explanation of the process.

continued on next page

SPECIAL OBSERVATION STATUS *(CONTINUED)*

- A Consulting Teacher will be assigned to the teacher and will assess the performance of the teacher.
- Based on the assessment, a support plan will be written and submitted to PAR Pair. If the teacher is in the Professional Development Plan (PDP year, s/he will continue with the Professional Development Plan. CT will continue support until May 31st.
- Principal will conduct one formal observation between February 1st and May 31st.
- CT will submit written statement of teacher's progress to PAR Panel by May 31st. The CT may report the following: (a) that the teacher is **performing below standards** in Domains II and/or III or (b) that the teacher is **approaching standards** in Domains II and/or III, or (c) that the teacher **meets the standards** in Domains II and/or III.
- Principal will submit formal observation by June 2nd. The PAR Panel will review the formal observation and the CT reports to make recommendations.
- If formal observation is unsatisfactory, and the CT reports the teacher is performing below standards in Domains II and/or III, the teacher will be rated unsatisfactory and placed in PAR for the next school year.
- If formal observation is satisfactory, and the CT reports the teacher is approaching standards or meets standards in Domains II and/or III, the teacher will be rated satisfactory and Special Observation Status is ended.
- If there is a discrepancy, the PAR Panel will review documentation.
- The CT will be required to provide supporting evidence to defend her/his report and the Principal will be required to provide evidence to support her/his formal observation.
- The PAR Panel will make the final decision regarding the status of the teacher.

STUDENT LEARNING OBJECTIVE (SLO) TEMPLATE

SLO is a process to document a measure of educator effectiveness based on student achievement of content standards. SLOs are a part of Pennsylvania's multiple-measure, comprehensive system of Educator Effectiveness authorized by Act 82 (HB 1901).

1. Teacher Information

1a.	**Teacher Name**	Mr. or Ms. Math, grade 4
1b.	**School Name**	___ Elem
1c.	**District Name**	_____SD

2. Setting

2a.	**Course Title(s)/ Targeted Content Areas**	Grade 4, mathematics, Geometry			
2b.	**Classroom Setting**	Grade Level(s)	4		
		Frequency of Classes/sessions	1X/day		
		Total Number of Classes/sessions	180		
		Average Number of students per class	25		
		Total Number of students (all classes/sessions)	25		
2c.	**Which students will be used for the SLO?**	__X__All ____Sample	**2d.**	**Rationale**	With accommodations and modifications all students will know and be able to apply the skills of the standard

3. Student Learning Objective (SLO)

3a.	**SLO Statement**	4th grade students apply geometric concepts to design a city based on multiple criteria as described in our Rubric	
3b.	**PA Standards**	CC.2.3.4A1 Draw lines and angles and identify these in 2-dimensional figures CC. 2.3.4.A.2 Classify two dimensional figures by properties of their lines an angles CC.2.3.4.A.3 Recognize symmetric shapes and draw lines of symmetry	
3c.	**SLO Rationale**	This SLO will provide students an opportunity to demonstrate mastery of fourth grade geometry standards within a real world application.	
3d.	**Timeline**	Date to submit Sections 1 through 8 of the SLO Template for approval	August of current school year
		Date to complete Section 9 of the SLO Template	Following the geometry unit

4. Performance Measure(s)		
4a.	**Name of the Performance Measure(s)**	Geometry Town Rubric
4b.	**Type of Performance Measure(s)** (Check all being used)	**____District-designed Measures and Examinations** **____Nationally Recognized Standardized Tests** **____Industry Certification Examinations** **__X__Student Projects** **____Student Portfolios**
4c.	**Performance Measure(s) Purpose Statement**	The Geometry Town Rubric provides an assessment of an individual student's application of essential concepts and competencies corresponding to the selected standards for the Geometry Town project.
4d.	**Metric Type** (Check one)	☐ Growth (change in student performance across two or more points in time) X☐ Mastery (attainment of a defined level of achievement) ☐ Growth and Mastery

5. Administration		
5a.	**Frequency of Administration**	1X/year
5b.	**Resources/Equipment Required**	Graph paper with 1 inch grid, construction paper, markers, crayons or colored pencils, rulers, protractors, yardstick
5c.	**Adaptations/Accommodations**	Extension: • Students may add a new geometric requirement to the city along with a written description of the addition to the map. • Invite an architect to the classroom to talk about planning and the models they build in their work • Encourage students to prepare a presentation to a "City Planning Committee" regarding their city plan. Students should try to persuade committee members to choose their plan Intervention: • Exemplar with similar but different requirements • Visual model (drawings) of key shapes (shapes labeled) • Pre-made 2-D shapes made available to some students, use of color coding, labeled angles, • Chunk directions into steps with just 2 -3 directions per page • Scale back requirements for students with IEPs according to their specially designed instruction (SDI) or Supplementary Aids and Services (SaS) • Collaborative partnerships
5d.	**Personnel**	Teacher, student, community member

6. Performance Data		
6a.	**Scoring Tool(s)**	Rubric
6b.	**Score Collection**	Total number of criteria for each student's final product

6c.	**Scoring Personnel**	Teacher
6d.	**Score Reporting**	**By class to principal**

7. Performance Indicator(s)		
7a.	**Classroom PI**	**All students score proficient in each of the 3 target areas**
7b.	**Targeted Population PI**	**All students in class**

8. SLO Expectations				
8a.	Failing: Fewer than 60% of students meet the performance indicator	Needs Improvement: 60% to 79% of students meet the performance indicator	Proficient: 80% to 89% of students meet the performance indicator	Distinguished: 90% to 100% of students meet the performance indicator
		----or----		
8b.	Failing: : Fewer than ___% of students meet ___ out of ___ performance indicators	Needs Improvement: : ___% to ___% of students meet ___ out of ___ performance indicators	Proficient: ___% to ___% of students meet ___ out of ___ performance indicators	Distinguished: : ___% to ___% of students meet ___ out of ___ performance indicators

Teacher Signature_____Date_____

Evaluator Signature_____Date_____

9. SLO Results Rating		
9a.	☐ Distinguished ☐ Proficient ☐ Needs Improvement ☐ Failing	**Notes/Explanation**

Teacher Signature_____Date_____

Evaluator Signature_____Date_____

Adapted for the Pennsylvania Department of Education

Student Achievement Measures for Student Learning Objectives (SAM-SLO)

Quality Assurance Checklist and SLO Rubric

Section 1: Teacher Information

- ☐ 1a. Teacher's full name (First, MI, Last) is stated.
- ☐ 1b. School's full name (if serving more than one school-list all) is stated.
- ☐ 1c. District's full name is stated.

Section 2: Setting

- ☐ 2a. All applicable courses/content areas included within the SLO are stated.
- ☐ 2b. All applicable grade levels included within the SLO are stated.
- ☐ 2b. Frequency of classes/sessions is described.
- ☐ 2b. Total number of classes/sessions is calculated.
- ☐ 2b. An average number of students in the class or instructional setting is predicted.
- ☐ 2b. If multiple sections are offered, a total number of students for this course is predicted.
- ☐ 2c. The representative student sample for this SLO is stated. (All sections, sample populations, etc.)
- ☐ 2d. A rationale for the student population selected for this SLO is stated.

Section 3: Student Learning Objective (SLO)

- ☐ 3a. The SLO is specific and answers the questions: Who? What? How?
- ☐ 3a. The SLO is measurable and results-focused (vs. process-focused).
- ☐ 3a. The SLO is ambitious but attainable.
- ☐ 3a. The SLO is relevant, meaning linked directly to the targeted content standards.
- ☐ 3a. The SLO is measured by a valid and reliable approach.
- ☐ 3a. The SLO is supported by data related directly to student performance outcomes.
- ☐ 3a. The SLO is fair and unbiased.
- ☐ 3a. The SLO supports instructional outcomes for students.
- ☐ 3b. The SLO supports and is aligned to content-specific PA Standards.
- ☐ 3c. The rationale statement articulates why the SLOs was being developed.
- ☐ 3c. The rationale statement identifies what the SLO is designed to measure.
- ☐ 3c. The rationale statement provides insight into how the overall results will be used.
- ☐ 3c. The rationale statement is concise and free of technical jargon.
- ☐ 3d. The dates for submission of sections 1-8 to the evaluator prior to implementation are provided.
- ☐ 3d. The date established when the SLO results are translated into a Results Rating.

Section 4: Performance Measure(s)

- ☐ 4a. The name(s) of the performance measure(s) is (are) stated.
- ☐ 4b. The type of performance measures is identified.
- ☐ 4c. The performance measure's purpose statement is provided.
- ☐ 4d. The metric type associated with the performance measure is identified.

Section 5: Administration

- 5a. The frequency each performance measure is administered is stated or referenced.
- 5b. Unique requirements, conditions, and equipment needed to administer the performance measures are stated or referenced.
- 5c. Unique adaptations or accommodations of the performance measures are stated or referenced.
- 5d. Any additional personnel needed to administer the performance measures (beyond the educator) are stated or referenced.
- 5d. The ability for an equivalent professional peer to administer the performance measure is stated.

Section 6: Performance Data

- 6a. Scoring aspects, needed rubrics, measures, etc. have been explained.
- 6a. Guidelines include step-by-step directions on how to document scores, interpret and record results.
- 6b. Descriptions of tools to collect multiple student performance measures is described.
- 6b. Methodology to evaluate student performance measures is described.
- 6b. When appropriate, data are collected in a manner that a "baseline" reference point can be established and a "target" data goal can be set.
- 6c. Any additional personnel needed to administer the performance measures (beyond the educator) are stated or referenced.
- 6c. The ability for an equivalent professional peer to administer the performance measure is stated.
- 6d. Descriptions of work samples, exemplars, databases or other artifacts that will be presented to the principal or evaluator to support the SLO Expectations and Results Rating are described.

Section 7: Performance Indicators

- 7a. Expected levels of achievement on the student performance measures are described.
- 7a. The selected baseline indicators are linked to the overarching SLO.
- 7a. The selected baseline indicators are fair and unbiased.
- 7a. The selected baseline indicators are independent of each other.
- 7a. The selected baseline indicators reflect the targeted content standards.
- 7a. The selected baseline data are verifiable.
- 7a. The selected indicator targets are linked to the baseline data.
- 7a. The selected indicator targets are specific and measureable.
- 7a. The selected indicator targets are realistic and attainable.
- 7b. Expected levels of achievement on the student performance measures for unique populations are described.

Section 8: SLO Expectations

- 7a. The percent of students expected to meet the the Performance Indicator is determined for each of the four teacher effectiveness descriptors.

---or---

- 7b. The percent of students expected to meet a defined quantity of Performance Indicators is determined for each of the four teacher effectiveness descriptors.

Performance Indicator Rubric

DIMENSIONS	GREEN LEVEL	YELLOW LEVEL	RED LEVEL
Specific Specifies a statement of accomplishment to be achieved; outcomes are results-focused, not process-focused	SLO is clearly articulated and targets the needs of a specific group; clearly points towards applicable performance and/or learning outcomes.	SLO is articulated but not specific or does not address the targeted performance and/or learning outcomes.	SLO is unfocused, does not address appropriate standards, or is activity-based.
Measurable Provides evidence that is measured by a valid and reliable approach	SLO is measurable because it incorporates valid and reliable performance measures. Measures use high quality data sources.	SLO is measurable but uses performance measures with unknown validity and reliability characteristics. Measures are based upon known data sources.	SLO is not a clear outcome that can be reliably measured; data quality is not sufficient to support valid interpretation of the observed scores.
Attainable Attainable, realistic, ambitious, but achievable expectation	PI is a rigorous and achievable outcome that delineates the "minimally acceptable improvement" based upon past data, research, and professional judgment.	PI is an achievable outcome that delineates the "minimally acceptable improvement" based mostly upon professional judgment and anecdotal data.	PI may be an achievable outcome but does not delineate the "minimally acceptable improvement" or a rigorous standard.
Relevant Related to professional and/or academic standards; supported by data related to student performance outcomes	SLO is clear and directly aligned to school, student, and/or position priorities. Indicators are directly linked to relevant student outcomes.	SLO is aligned to school, student, and/or position priorities. Indicators are tangentially relevant to student outcomes.	SLO is not aligned to school, student, and/or position priorities. Indicators are not relevant.
Time Bound Uses at least two separate points in time in a given year	PI clearly identifies and establishes benchmarks for two or more points in time within a given calendar year OR a clear, date-specific target for demonstrations of mastery.	PI identifies baseline and target time frames within a given calendar year OR a general timeframe for demonstrations of mastery.	SLO PI does not consist of two points in time within a calendar year OR the target timeframe for demonstrations of mastery are implied.

The 7 C's and Sample Tripod Survey Items		Corresponding Performance Indicators from the New York State Teaching Standards
CARE	• I like the way my teacher treats me when I need help. • My teacher in this class makes me feel that he/she really cares about me. • My teacher really tries to understand how students feel about things.	**IV.1a.** Teachers are caring and respectful in their interactions with students. **IV.1d.** Teachers create a climate of acceptance and respect.
CONTROL	• My classmates behave the way my teacher wants them to. • Students in this class treat the teacher with respect. • Our class stays busy and does not waste time.	**II.6c.** Teachers organize and effectively use time to achieve learning goals. **IV.3a.** Teachers establish, communicate, and maintain clear standards and expectations for student behavior. **IV.3d.** Students exhibit respectful classroom interactions.
CLARIFY	• My teacher explains difficult things clearly. • My teacher knows when the class understands, and when we do not. • I understand what I am supposed to be learning in this class. • If you don't understand something, my teacher explains it another way.	**II.3b.** Teachers adapt instruction in response to various levels of student understanding. **III.2a.** Students understand directions and procedures. **III.2d.** Students understand lesson content through a teacher's use of multiple modalities, such as oral, written, graphic, kinesthetic, and/or tactile methods. **III.2e.** Teachers adjust communication in response to student needs. **III.6c.** Teachers adjust the pace of instruction, focus of instruction, and method of delivery based on students' progress.
CHALLENGE	• My teacher wants us to use our thinking skills, not just memorize things. • My teacher asks students to explain more about answers they give. • In this class, my teacher accepts nothing less than our full effort. • My teacher doesn't let people give up when the work gets hard.	**II.2c.** Teachers provide opportunities for students to engage in individual and collaborative critical thinking and problem solving. **III.3a.** Teachers articulate high expectations for all students. **III.3c.** Teachers challenge and support all students by incorporating various instructional strategies, experiences, and resources. **IV.2a.** Teachers encourage students to set high standards and expectations for their own performance.
CAPTIVATE	• My teacher makes lessons interesting. • Homework helps me learn. • I like the ways we learn in this class. • My teacher makes learning enjoyable.	**III.4c.** Teachers incorporate motivating and meaningful opportunities in instruction to engage students in learning experiences. **IV.2c.** Teachers promote students' curiosity and enthusiasm for learning. **IV.2d.** Students are actively engaged in learning.
CONFER	• My teacher asks questions to be sure we are following along when he/she is teaching. • My teacher wants us to share our thoughts. • Students get to decide how activities are done in this class. • My teacher respects my ideas and suggestions.	**II.5a.** Teachers determine current levels of students' understanding and knowledge of content through questioning techniques, discussion, and other methods. **III.2c.** Students' comments and questions are acknowledged and utilized to advance learning. **III.6b.** Teachers seek and provide feedback during and after instruction. **IV.2e.** Students openly express their ideas.
CONSOLIDATE	• My teacher takes the time to summarize what we learn each day. • My teacher checks in to make sure we understand what he/she is teaching us. • When my teacher marks my work, he/she writes on my papers to help me understand how to do better. • We get helpful comments to let us know what we did wrong on assignments.	**II.5c.** Teachers design learning experiences that connect students' prior knowledge and instruction to new content. **III.6a.** Teachers utilize various types of formative assessment during instruction to monitor and check for student understanding and assess progress. **V.2b.** Teachers provide timely feedback to engage students in self-reflection and self-improvement.

What are some examples of questions on the survey?

Category 1: Presenter – Ability to present information and structure lessons

- *When explaining new skills or ideas in class, my teacher tells us about common mistakes that students might make*
- *At the end of each lesson, the teacher reviews what we have just learned*

Category 2: Manager – Ability to manage a classroom and foster productivity

- *My teacher corrects students when they do not follow the rules of the class*
- *We are learning or working during the entire class period*

Category 3: Counselor – Awareness of student need and teacher-student relations

- *My teacher shows respect for all students*
- *My teacher notices when I am not participating in class*

Category 4: Coach – Providing feedback and challenging students

- *My teacher gives us guidelines for assignments so we know how we will be graded (grading rules, charts, rubrics, etc)*
- *I have to work hard to do well in this class*

Category 5: Motivator: Engaging and investing students in learning

- *My teacher has us apply what we are learning to real-life situations*
- *My teacher encourages me to share my ideas or opinions about what we are learning in class*

Category 6: Content Expert: Knowledge of subject and encouraging student thinking

- *My teacher is able to answer students' questions about the subject*
- *After asking us questions, my teacher lets us think for a few seconds before we have to answer*