Schools

Pittsburgh's Community Alliance for Public Education

Pittsburgh's New Teacher Improvement System

Helping teachers help students learn

Authors:

Amy Scott, MSPPM, Director of Research and Data Analysis, A+ Schools

Richard Correnti, Ph.D., University of Pittsburgh School of Education, Learning Research and Development Center

Pittsburgh's New Teaching Improvement System
Helping teachers help students to learn

After nearly five years of collaboration, Pittsburgh Public Schools (PPS) provided teachers with a comprehensive report card of how effectively they are reaching students earlier this month. The "no-stakes attached" report card being issued—the Educator Effectiveness Report—is just one piece of an $80 million effort embarked upon by the Pittsburgh Federation of Teachers (PFT), PPS, and local educational professionals to create a system of constant feedback and growth that will improve classroom teaching and educational outcomes for students.

How'd we get here?

A strong process of intentional collaboration between the PFT, PPS, teachers, and principals has produced a strong teaching evaluation system.

2008: Less than 15% of PPS teachers surveyed strongly agree that "Teacher evaluation in my building is rigorous and reveals what is true about teachers' practice."

2008-2009: "PFT/PPS RISE Leadership Team" formed to develop a new observational system of evaluation.

2009 -2010: A+ Schools rallies community support for PFT and PPS co-authored Empowering Effective Teachers Plan, which gets funded by Federal, State and foundation money to create a comprehensive system of teacher improvement.

2010-Present: Different measures are piloted and teacher feedback is obtained in order to create strong measures that teachers can use and trust.

August 2013: For the first time, teachers get a no-stakes report combining information from the various piloted programs to give them a comprehensive picture of how they're doing and provides feedback on how to improve.

What have we learned?

Improving teaching quality helps students: If we want to close achievement gaps and ensure all of our children graduate ready for college or career training, great teaching is the one thing schools provide that will most contribute to a student's academic success.

Pittsburgh's evaluation system is really a teacher improvement system: The critical partnership between the PFT and PPS and the careful piloting of different measures to get feedback from teachers and principals about what works has made this system strong and focused where it should be: helping all teachers get better.

Evaluation is important and should be bolstered by improved conditions that support effective teaching: These supports include quality professional development and training, positive learning environments, and strong school leadership.

No system is perfect and improvements should continue to be made: Because it is critical to maintain a teacher evaluation system that is reliable, accurate, and provides meaningful feedback for improvement, PPS and the PFT should be willing to refine and improve the system based on growing research about teaching and learning.

We all have a role to play to improve our schools: Teachers and principals alike want families and community to be involved. This brief provides a variety of tools to have constructive conversations with students, teachers, principals and other parents.

www.aplusschools.org

A+ Schools
Pittsburgh's Community Alliance for Public Education

Outline - Pittsburgh's New Teaching Improvement System

Introduction

Why is great teaching so important for Pittsburgh's students?

"My children had a favorite teacher…who took a lot of extra time to do work outside of the standard curriculum and the parents knew it. She taught my children how to communicate. Their writing developed because she took the time to give them extra feedback and provide practical advice on how they could improve."

- Maria Searcy,
   PPS Parent

Public education has yet to fulfill the promise of Brown v. Board of Education to improve educational outcomes for black and brown students, with severe costs to our nation.[1] Persistent gaps in student achievement result in part from our system, which frequently relegates low-income students of color to schools where "academic expectations and opportunities [are] considerably lower than what we expect of other students."

As Pittsburgh's advocate for educational equity and excellence in public schools, A + Schools champions changes in policies and practices that will accelerate academic achievement for students —especially black students. Because teachers are schools' most important contributor to students' academic and personal success, A+ Schools has been an advocate for teaching improvements so that each student experiences great teaching in every classroom, every day.

---

[1] According to the National Dropout Prevention Center, each year's class of dropouts will cost the country over $200 billion during their lifetimes in lost earnings and unrealized tax revenue (Catterall, 1985). In addition, the estimated tax revenue loss from every male between the ages of 25 and 34 years of age who did not complete high school would be approximately $944 billion, with cost increases to public welfare and crime at $24 billion (Thorstensen, 2004).
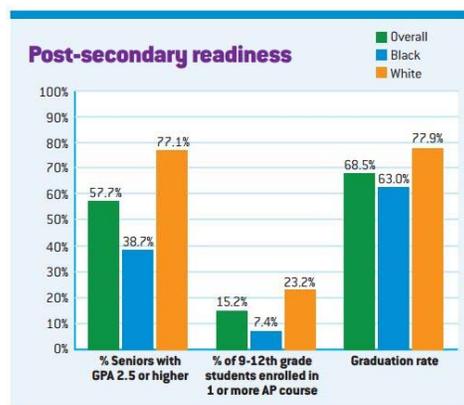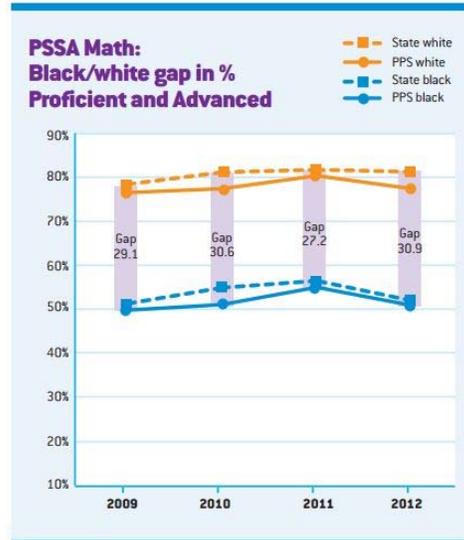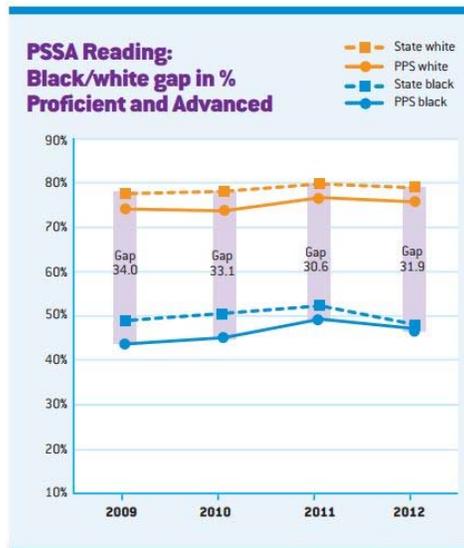
Pittsburgh Public Schools (PPS), like most school districts, has huge disparities in achievement between its students of color and its white students. On average since 2008, more than 75% of white students have tested proficient or advanced on state assessments, but less than 50% of black students have tested similarly. In addition, Pittsburgh graduates 63% of its African-American students compared to 77.9% of its white students and it is estimated that only 44% of black males successfully complete high school.[2] These numbers are discouraging and undoubtedly result from many factors stemming from systemic inequities that are expressed inside and outside of the classroom. What's encouraging, however, is that the most effective teachers in PPS produce gains in student achievement that, if accumulated over several years without decay, could erase the achievement gap between black and white students.[3] While not all teachers can perform in the top 5%, improving teaching through evaluation, quality feedback, and support to teachers can help raise achievement for all students and start to reduce gaps between black and white students.

Effective teaching defined and measured in Pittsburgh – a critical moment in time

> **In 2011 School Works interviews, teachers unanimously agreed that their success was tied to student success.**

How does Pittsburgh Public Schools support its teachers to improve their practice and deliver excellent education to their students? Until a few years ago, PPS did not formally define the attributes of an effective teacher. Educators received minimal feedback and a binary evaluation score of satisfactory or unsatisfactory. Through this system, 99% of teachers received a satisfactory rating with little to no feedback on what they were doing well and where they could focus to improve.

After years of working together with teachers, principals, and national education experts, a common standard has emerged:

> **An effective teacher is a professional, who knows his or her subject and teaches it well, inspiring and engaging all students as individuals to fulfill their personal and career goals, and accelerating learning so that all students are Promise-Ready.[4]**

This definition acknowledges the complexity of teaching, which involves not only delivering content knowledge to students but also building relationships with them to prepare them for the future.

Along with defining effective teaching, a dedicated group of professionals in Pittsburgh has worked for five years to create a thoughtful, comprehensive system of evaluation. This system was designed to provide teachers with information from multiple sources that they can use to make meaningful adjustments to each aspect of their practice mentioned above. It is meant to ensure that teachers gain a better understanding of what great teaching is and get the feedback they need to help their students succeed.

---

[2] A+ Schools Report to the Community (2012); Schott Foundation for Public Education (2012).
[3] Mathematica Policy Research, Inc. (2010).
[4] Pittsburgh Public Schools (1/3/2013).

This summer marks an important moment in time regarding teacher evaluation in Pittsburgh: for the first time, teachers will receive an evaluation based on a combination of three different lenses of effective teaching: classroom observations, student growth measures, and student perception surveys. Teachers have already been receiving feedback in each of these measures for the past 1-3 years, but now they will see all three together in one place.

Under Pennsylvania state law (Act 82), beginning in 2013-14, teacher evaluation must be based on professional practice and student outcomes. The information given to teachers this summer therefore gives teachers a full year before their evaluations count to pinpoint areas of strengths and weaknesses and improve their practice. The timing also allows the community and school district to know more about where and how to focus supports for our teachers.

Teacher evaluation – an important part of supporting great teaching

In this paper we have chosen to focus on Pittsburgh's system of teacher evaluation because it is a necessary—though, by itself insufficient—tool for growing and supporting teachers. We also know that teaching improves with high quality teacher training and professional development programs, positive school learning environments, nurturing school leaders who create strong instructional cultures and systems that make teaching quality count when deciding who teaches and where they teach.[5]

Understanding and improving supports for Pittsburgh's teachers

No system is perfect, and with so much at stake for students, any system to provide teachers with good feedback must continue to grow stronger. Teachers should always have access to the most meaningful information and feedback needed to improve their practice and advance their students' success. While evaluating what high-quality effective teaching looks like comes with many complexities, practitioners and researchers alike agree on some overarching criteria that a strong system of evaluation should strive to achieve:

- **Reliability** - whether the measures used are consistent across time and evaluators
- **Accuracy**- whether we are measuring what we believe we are measuring
- **Feedback for Improvement -** provide timely and meaningful feedback to teachers so they can refine, grow, and master their practice

The remainder of this paper will focus on:
1. how Pittsburgh has begun to use multiple measures,
2. what the multiple measures are,
3. what we know about our teaching corps based on results in each measure,
4. the degree to which each measure meets the criteria above,
5. and what can be done to strengthen the system and make it useful to teachers and students.

---

[5] L. Darling-Hammond, A. Emrein-Beardsley, E.H. Haertel, and J. Rothstein (2011).

Pittsburgh's new teacher evaluation system is designed to rely on multiple measures, or lenses, of effective teaching, which currently include classroom observations (Research-Based System of Evaluation, or RISE), student growth measures (value-added measures, or VAM), and student perception surveys (Tripod Student Survey). There is widespread consensus that the best teacher evaluation systems use multiple measures. Multiple measures serve a dual function:

PPS is proposing the following approach to reaching a combined measure of effective teaching:

50% Student Outcomes

5% Building-Level Results *PPS School VAM*

50% Classroom Observation *RISE/EIP*

30% Teacher-Specific Data *PPS Teacher VAM/RISE 3f*

15% Elective Data *Tripod Student Survey*

Source: Empowering Effective Teachers Education Committee Update (January 3, 2013).

1. They provide a check and balance across the different measures and help to improve the accuracy of any one measure used. Consistently positive or negative scores across multiple measures help to underscore the strength of evidence contributing to the overall evaluation.
2. Even more importantly, more information across measures helps maintain the focus of the evaluation system: feedback for the purpose of improvement. Including various sources of information in the evaluation system helps teachers know how they can respond to their evaluation to improve.
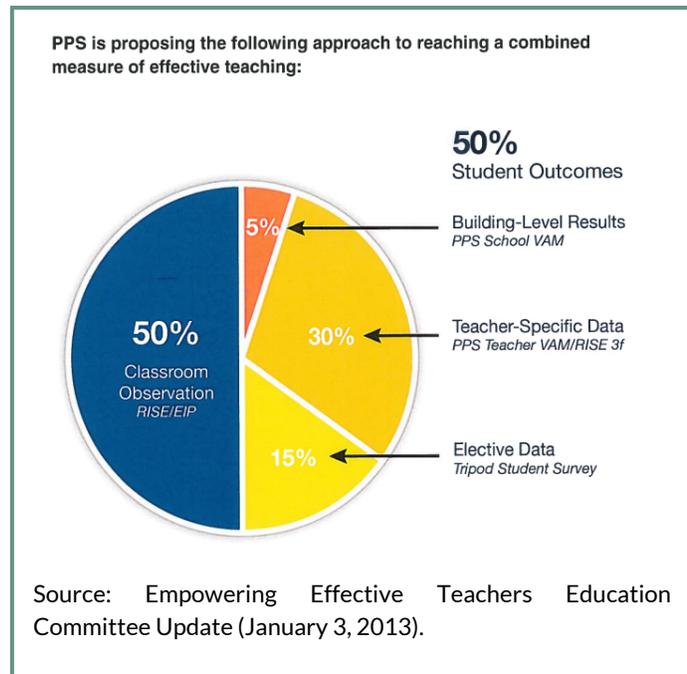
The importance of using multiple measures has been underscored by teacher unions and educational researchers alike:

> *When evaluating a complex activity—arguing a legal case, performing a surgical procedure or teaching a lesson—multiple measures examine diverse facets of the activity from many perspectives. These different perspectives are essential if we want to get an accurate representation of what we wish to assess. The sum of the various criteria that go into the assessment, as well as the multiple ways of looking at the criteria, allow us to make a considered judgment of a complex activity.[6]* – The American Federation of Teachers

> *Effective [teacher evaluation] systems have developed an integrated set of measures that show what teachers do and what happens as a result. These measures may include evidence of student work and learning, as well as evidence of teacher practices derived from observations, videotapes, artifacts, and even student surveys.[7]* – Linda Darling-Hammond et al.

---

[6] American Federation of Teachers, p 5.
[7] L. Darling-Hammond, A. Emrein-Beardsley, E.H. Haertel, and J. Rothstein (2011), p10.

Classroom observation – looking at great teaching through the lens of classroom practice

In order for students to grow and succeed, they need supportive teachers who set high expectations and clear learning goals and who give them continuous feedback and opportunities to revise their work. Teachers also deserve supports for understanding their strengths and areas for improvement, and they should be held to high standards that link clearly defined practices with positive results for students. Classroom observation is one important way to evaluate and give timely feedback to teachers about how they can improve their practice.

In 2011, most teachers surveyed by A+ Schools agreed that formal classroom observations help them to improve their practice (83%) and contribute to student achievement (79%). 66% of principals interviewed in 2013 believe RISE helps them differentiate supports for teachers at different levels.

What does classroom observation look like in Pittsburgh?

As we noted earlier, Pittsburgh teachers used to receive a rating of either satisfactory or unsatisfactory at the end of the school year. In 2008-09 the District, PFT, and a team of teachers and principals developed a much-improved system for teacher observation and evaluation, the Research-based Inclusive System of Evaluation (RISE). Starting in 2010-11, RISE became the official evaluation system for teachers. Now in 2013-14, the new teacher evaluation system will combine RISE scores with student growth data and student perception surveys.

RISE is based on Charlotte Danielson's Framework for Teaching, a set of components that are aligned to standards of the Interstate New Teacher Assessment and Support Consortium (INTASC), adopted by over 40 states for licensing new teachers.[8] Twenty-four components are organized within four domains: Planning and Preparation, The Classroom Environment, Professional Responsibilities, and Teaching and Learning.

Through RISE, school administrators and teacher leaders formally observe teachers at least once during the year, as well as visit classrooms for informal observations. Each formal observation includes a pre-conference and post-conference, where teachers receive scores (distinguished, proficient, basic, or unsatisfactory) in up to 24 components as well as feedback for improvement.

What do we know about Pittsburgh's teaching force based on classroom observations and feedback?

Nearly 85% of all teachers (about 1,700) received final RISE ratings in 2011-12. Teachers received a rating of distinguished, proficient, basic, or unsatisfactory for up to 24 different RISE components. Based on district-level averages of teachers' final scores, we can understand some helpful trends:

---

[8] L. Darling-Hammond et al (2011).

"A great teacher is someone who you know cares about your kids. And you know that by having conversations with them. You know that by when your children come home from school and they're excited about that subject, about that person."

- Tracey Reed Armant,
  PPS Parent

- **Most teachers show proficiency in their classroom observation evaluations.** An average of 78% of teachers received proficient ratings across all components, which allowed school and teacher leaders to focus resources on teachers who needed extra help.[9]

- **Teachers are performing well in the areas of professionalism, communicating with students, and organizing physical space.** At least 95% of teachers who were rated on these components received a proficient or distinguished rating. The greatest percentage of teachers received distinguished ratings in the following components: creating a learning environment of respect and rapport (17%) and communicating with families (16%).

- **There are some areas where teachers can improve.** About 1 in 3 teachers (35%) received a score of basic in component 3b. Using questioning and discussion techniques. This means teachers may ask lower level questions and not always engage all students in discussions or clarify their misconceptions.[10] About 1 in 5 teachers scored basic in 1f. Designing ongoing formative assessments (22%) and 3d. Using assessment to inform instruction (23%). This means teachers may communicate assessment criteria only to some students, inconsistently use results to differentiate instruction to students at different levels, or give verbal or written feedback that is not helpful to student learning.[11]

- **Teachers have an opportunity to focus their professional growth in just one area.** Nearly 600 teachers participated in Supported Growth Projects (SGPs), which are individualized independent professional development projects in specific areas that these teachers want to focus on for an academic year.[12]

- **Other teachers need intensive support to improve their practice and be more effective at raising student academic achievement.** Based on evidence that they needed additional support, principals placed roughly 7% of teachers (over 130) on Employee Improvement Plans (EIP) for intensive support during 2011-12. The District reports that many teachers placed on EIPs since 2009 have left the District either as a result of ratings, resignation or retirement, while many others have improved and are no longer participating in the EIP process.

---

[9] In 2011-12 individuals working on Employee Improvement Plans (EIPs) were not part of the RISE process and therefore did not receive ratings for any of the RISE components. Thus, this distribution of observation results does not represent the results of the totality of the PPS teacher workforce and is not fully representative of observation of teacher practice in the District.

[10] Pittsburgh Public Schools (8/11/2011); footnote 9.

[11] Pittsburgh Public Schools (8/11/2011); footnote 9.

[12] 20%: Using assessment to inform instruction, 15%: Using questioning and discussion techniques, 15%: Engaging students in learning, 8-10%: System for managing students' data and Communicating with families
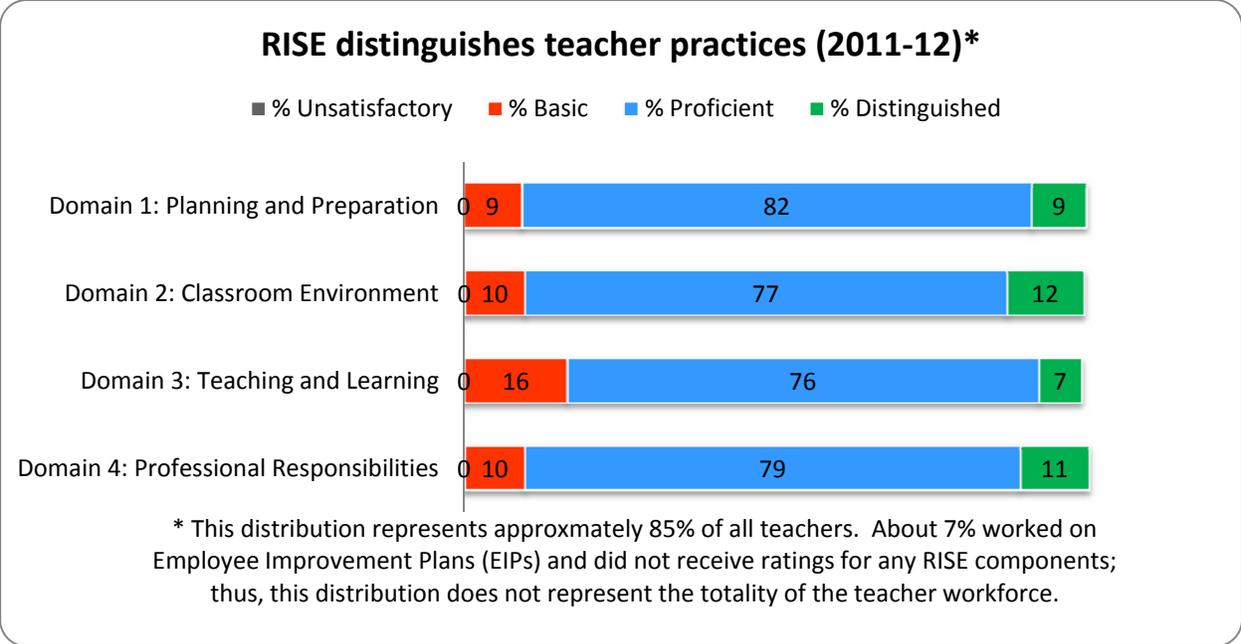
RISE distinguishes teacher practices (2011-12)*

■ % Unsatisfactory　■ % Basic　■ % Proficient　■ % Distinguished

Domain 1: Planning and Preparation	0　9　82　9
Domain 2: Classroom Environment	0　10　77　12
Domain 3: Teaching and Learning	0　16　76　7
Domain 4: Professional Responsibilities	0　10　79　11

* This distribution represents approximately 85% of all teachers. About 7% worked on Employee Improvement Plans (EIPs) and did not receive ratings for any RISE components; thus, this distribution does not represent the totality of the teacher workforce.

Figure 1: Teacher Performance Across RISE Domains

How does PPS' implementation of RISE measure up to criteria for a strong evaluation system?

Our students benefit the most from teacher evaluation systems implemented in accurate and reliable ways that help our teachers to improve. Best practices have been identified from research on building these systems, including the Measures of Effective Teaching (MET) Project which studied multiple classroom observation systems and student surveys relative to value-added estimates of student academic growth for 3 years with nearly 3,000 teacher-volunteers in public schools across the country.[13]

Criteria for strong evaluation system	What PPS does well with respect to RISE	How PPS could strengthen RISE's contribution to the system
Reliability	Every evaluator participates in the two-year Instructional Quality Assurance and Certification (IQAC) Process to ensure reliability and quality instructional feedback and support.	PPS must ensure that evaluators remain reliable over time by mandating ongoing recertification. Multiple observers should evaluate each teacher, especially if high-stakes decisions about employment will be linked to scores.

---

[13] Kane, T., & Staiger, D. O. (2012).

		Observers with no personal relationship to teachers should observe a subset of teachers to ensure system reliability.
	Pittsburgh teachers receive 2 formal observations a year before receiving tenure.	Pittsburgh teachers should receive more than one formal observation a year regardless of tenure status.
Accuracy	IQAC Process to certify evaluators, as described above.	Although there is evidence to suggest it, PPS should regularly verify that teachers with higher RISE scores also have higher achievement gains on average.[14]
	The rubric clearly articulates the meaning and critical attributes of each domain and component.	What the observation rubric focuses on and how information about teaching is collected in classrooms needs to continually be refined to more accurately track the dimensions of teaching that researchers find to be most highly predictive of student performance. This system must be dynamic in order to reflect what we continue to discover about how students learn.
Provides timely and meaningful feedback	Evaluators provide feedback to teachers as part of a post-observation conference.	The degree to which teachers find the feedback meaningful varies and should be tracked.[15]

---

[14] For example, Mathematica Policy Research Inc. examined the construct validity of the most recent value-added estimates and is expected to release a report on that validity study in the fall of 2013.
[15] Based on responses from teachers during A+ Schools' School Works community action research in 2011.

Student Growth – looking at great teaching through the lens of gains in student performance

Teachers cannot control the level of academic proficiency their students start with when they first step foot in the classroom,   However, teachers should be recognized for their ability to help students grow, regardless of their starting points. Value-added models attempt to identify the degree to which a school or teacher contributes to growth in student performance beyond what is expected. The idea behind these models is that given a student's unique characteristics (including prior achievement) we

"It's all about trying to figure out how do I meet the student[s] on their level."

- **Harold Michie, Teacher**

can predict an expected score for each student under the assumption that s/he made average gains for students with similar characteristics. A teacher's "value-added" for an individual student could be calculated based on a simple formula: the student's actual gains (the difference between current and prior achievement) minus the expected gains (see Figure 2). As the number of students taught by a given teacher increases the accuracy of our judgments about a teacher's "value-added" estimate improves.
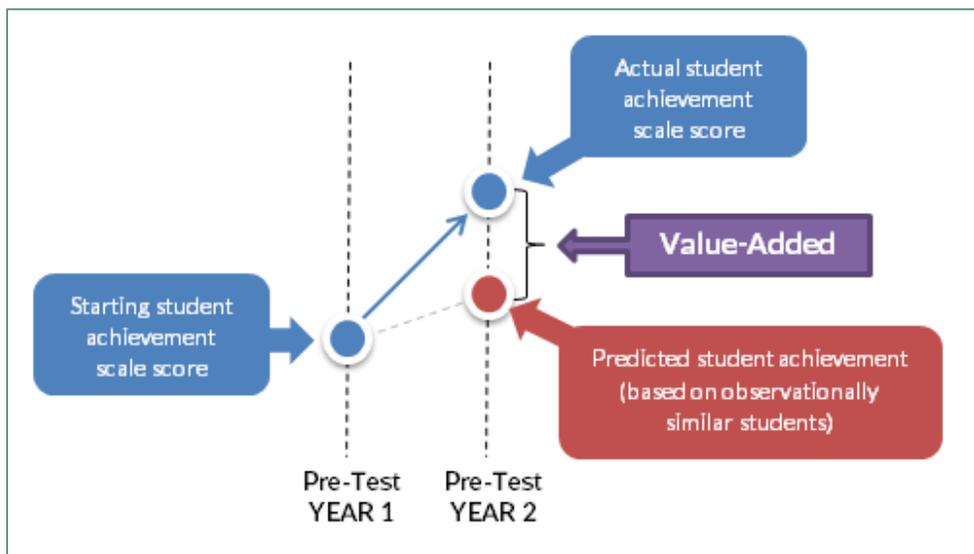

Figure 2: A teacher's value-added is the difference between a student's actual achievement and his/her expected achievement as predicted by a model.

Unpacking value-added models

In recent years the lexicon of "value-added models" (VAM) has come to mean different things to different people. Value-added models are not well understood and are often confusing because they vary widely in their application and use. While some scholars strongly advocate for their increased usage,[16] other scholars are concerned they will lead to more harm than good.[17] The truth is complicated and usually lies somewhere in the middle.

Attempting to use VAM to measure teachers' contributions to their students' academic growth has proven to be a challenging yet worthwhile task, given the fact that researchers for many years have compared student outcomes from different populations even while acknowledging the great limitations in doing so. To accurately understand them we must analyze each application of value-added models along at least two dimensions: 1) the strength of the data as well as the model applied to the data, and 2) how the results of the value-added model are employed in an evaluation system. Below we assess how PPS uses data and statistical models to generate a value-added estimate and how they propose to use these estimates as part of the new teacher evaluation system.

How strong is Pittsburgh's model for measuring student growth?

With the advent of Act 82 in Pennsylvania, it is mandatory for districts to include student performance in teacher evaluations. A fundamental question is whether value-added models perform better or worse than alternative ways of incorporating student achievement in teacher evaluations. Despite some of the popular rhetoric against them, value-added models are often *far preferable* to examining *unadjusted* student outcomes, given that they can account for the student characteristics that are outside of a teacher's control and that they try to understand student growth in addition to students' ending status.

Limitations of value-added models and guidelines for usage: Value-added models assume that the difference between actual and expected gains can be attributed to the effectiveness of the teacher. Teacher effects are most certainly contained in the estimates, but so is some measurement error, and it is unknown how much noise there is relative to the signal in these measures. Additionally, the "effects" attributed to the teacher are naturally due to some combination of the classroom context (effects over which the school and teacher have no control) and the instructional practice of teachers (effects that represent unique contributions of teachers). Thus, while it is important to hold teachers and schools accountable for student learning, this is balanced by the concern that accountability be practiced as fairly as possible.

Other limitations:
- separating teacher effects from school effects, using just value-added models, is nearly impossible, though the models try to parse the variance in order to generate estimates of these effects separately
- only teachers in tested grades and subjects obtain value-added estimates; thus, accountability to the particular student test scores used to calculate value-added estimates applies to less than half of all PPS teachers
- missing data could be problematic if the missing data pattern is not random.

School districts intending to use value-added estimates to measure teaching effectiveness should take care to calculate estimates by using large amounts of data over multiple years and by choosing a statistical model that will ensure reliable and consistent estimates. The law of large numbers suggests that with greater numbers of students passing through a single teacher, the estimate of a teacher's impact on student performance, relative to similar classrooms, can become very reliable. School districts should not only average across a classroom of students taught by a teacher in a given year but also include multiple cohorts of students across years, allowing for the possibility that a teacher may have been randomly assigned a tougher than normal class in a given year.

---

[16] Sanders, Wright and Horn (1997); Gordon, Kane and Staiger (2006).
[17] Darling-Hammond et al. (2012); Rothstein (2010).

Although there are different methods for calculating value-added estimates, most states and districts using VAM have settled on using models that are covariate-adjusted,[18] where a student's current achievement is predicted by prior achievement variables from last year and other student background characteristics.[19] Important features of Pittsburgh's model include the following:

- separate analyses are conducted by grade level and subject
- multiple test scores (e.g., math, reading and science) are used to predict each single-subject outcome (e.g., reading)
- estimates are averaged across up to three years, increasing precision of the estimates
- adjustments are made for a host of student background characteristics
- a calculation called a "quadratic polynomial predictor" is included to guard against floor and ceiling effects in which students scoring at either end of the distribution may appear to grow slower or faster, on average, than those in the middle of the distribution

The models used to calculate value-added estimates for PPS thus appear to be acceptable, if not standard, practice across the country.

The models themselves are generated not by PPS personnel, but rather by third-party experts at Mathematica Policy Research. This ensures no inter-personal biases creep into the estimation procedure, and it also ensures the technical quality of the analyses. The relationship between PPS and Mathematica is longstanding, and they have been collaborating over these models for several years. Thus, far from an abrupt response to recent legislation, the use of value-added estimates in PPS has been carefully planned. More importantly, because they have had several years to examine the models, current estimates are more precise because they rely on multiple years of data.

Finally, recent validity studies of VAM estimates have demonstrated correlations with the observation scores (RISE rubric) and the student survey (Tripod) commensurate with the Measures of Effective Teaching (MET) study undertaken by the Gates foundation. [20] This is significant because the MET study is the largest study conducted to date correlating different subject-specific observational measures of teaching with VAM estimates. Thus, not only do the statistical analyses conducted for PPS have face validity compared with other places instituting similar policies across the country, but other attempts to validate value-added estimates in PPS also conform to other national efforts at validation.

How does Pittsburgh incorporate value-added estimates into its new teacher evaluation system?

Given the limitations described in the box, scholars have been vocal about the cautions for using value-added estimates as the sole or even primary mechanism for making high-stakes decisions about teachers.[21] Perhaps the most important facet of the district's evaluation system (and the mandates of Act 82) is that evidence from value-added estimates can constitute no more than half of a teacher's evaluation. Teacher evaluation systems are too new to know the appropriate weight that should be given to value-added estimates, but there is near universal consensus that

---

[18] Wiley (2006)

[19] Given longitudinal data, researchers often examine a different variant of value-added models examining students' growth trajectories. In these models each student is their own control and teacher value-added estimates are comprised of the accumulation of "deflections" from students' individual growth trajectories.

[20] Kane, T., & Staiger, D. O. (2012).

[21] Harris (2011); Baker et al. (2010); Raudenbush (2004); McCaffrey et al. (2003)

evaluation systems should not rely on VAM alone since they do not give a full enough picture of a teacher's practice.[22] Pittsburgh Public Schools has proposed having a combination of school- and teacher-level VAM constitute 35% of a teacher's final evaluation.

What do we know about Pittsburgh's teaching force based on student growth measures?

For the past two years, 35-40% of teachers in PPS have been receiving VAM scores that include between one and three years of previous data for their students. As we mentioned in the beginning of this brief, this summer teachers with VAM scores based on three years of previous student data will receive a score combining their VAM, RISE, and Tripod scores as a preview for how their final evaluations will be structured for the 2013-14 school year. Teachers will have an opportunity to look at how their VAM scores may be reflective of the practices highlighted through RISE and Tripod, and begin to utilize this information to improve their practice. At the writing of this report, there was no meaningful way for us to report teacher VAM across the district. For teachers for whom there is no available VAM data or for whom VAM scores are not available based on three years of previous student data, PPS and PFT have elected to use evidence of student growth as collected through RISE 3f: Assessment results and student learning.

How does PPS' implementation of VAM measure up to criteria for a strong evaluation system?

Criteria for strong evaluation system	What PPS does well with respect to VAM	How PPS could strengthen VAM's contribution to the system
Reliability	PPS' measures rely on multiple years of data (up to three years) in order to create more reliable estimates.	PPS should continue to monitor the strength of their data and their models relative to the most recent research findings.
Accuracy	PPS' measures take into account • extensive student background characteristics (i.e., special education, special services, income, etc.) • differences in growth across the spectrum of achievement (floor and ceiling effects) PPS' measures use multiple test scores in different subjects to predict each single-subject outcome.	PPS should ensure that the tests used to calculate value-added estimates are well-aligned with curriculum and that the test broadly represents the skills teachers focused on over the year. Although there is preliminary evidence to suggest it, PPS should regularly verify that teachers with higher scores in other measures (classroom observations and student perception surveys) also have higher achievement gains on average.[23]

---

[22] Scherer (2012).

[23] For example, Mathematica Policy Research Inc. examined the construct validity of the most recent value-added estimates and is expected to release a report on that validity study in the fall of 2013.

Provides timely and meaningful feedback	PPS' value-added measures are designed to be used in combination with other measures to maintain a focus on feedback for improvement.	PPS and school leaders, and teachers should find ways to leverage value-added scores with RISE and Tripod results to help teachers uncover strengths and weaknesses in their teaching and provide resources for improvement.

Student surveys – looking at great teaching through the lens of students' perceptions in the classroom

No one knows more about what happens in classrooms than students and teachers. Researchers over many decades have found that students will engage more deeply and master their lessons more thoroughly when their teachers care about them, control the classroom well, clarify complex ideas, challenge them to work hard and think hard, deliver lessons in ways that captivate, confer with them about their ideas, and consolidate lessons to make learning coherent.[24]

"I think the kids really tell you what's great teaching. So when I do an observation, I tend to look more at the kids than at the teacher."

- **Katy Carroll, Teacher**

What does getting feedback from students look like in Pittsburgh?

Since 2009, students have shared their perceptions through surveys called the Tripod Student Survey, starting with 300 classrooms in 2009-11 and expanding to include 1,800 teachers and over 3,000 classrooms for 2011-12 and 2012-13. The surveys differ based on student grade level (K-2, 3-5, 6-8, and 9-12), with different language, numbers of questions, and numbers of response choices. Questions are meant to draw out student perceptions within each of the 7 Cs mentioned above: care, control, clarify, challenge, captivate, confer, and consolidate.[25]

Tripod has proven to be a reliable, valid measure, with students generally agreeing on their perceptions of their teachers regardless of factors such as age or race. Developed by Harvard researcher Dr. Ron Ferguson beginning in 2000 and refined by K-12 teachers, administrators, and professionals at Cambridge Education, Tripod has been given to several hundred thousand elementary, middle, and high school students in hundreds of schools in the United States, Canada, and China.[26]

---

[24] R.F. Ferguson (2010).
[25] For a description of the 7 Cs, please consult our Tool Box at the end of this brief.
[26] R.F. Ferguson (2010).

What do we know about Pittsburgh's teaching force based on student perceptions?

Based on district-level summaries of students' favorable responses in each of the 7 Cs for each of the grade levels, we can understand some helpful trends:

- **Pittsburgh teachers fall close to national benchmarks for favorable responses broken down by the 7 Cs.** For each grade span, PPS teachers scored within 0-4 percentage points of the national average of overall scores. When looking at each of the 7 C categories, the range was within 8 points below and 5 points above national averages. More than anything, this means that the surveys have been administered as consistently in Pittsburgh as they have in other districts. It may also suggest that on average, Pittsburgh's teachers experience similar strengths and challenges as other public school teachers nationwide.

- **Teachers are compared with their peers who teach within the same grade spans.** This levels the playing field and ensures fairness in scoring, given that benchmarks are different for different grade spans. Results are confidential and teachers cannot see each other's individual scores.

- **On average, teachers received the highest percentage of favorable responses in the areas of Care and Clarify (elementary grades) and Challenge and Clarify (secondary grades).** Care means helping students feel emotionally safe and able to rely on the teacher as a dependable ally. Clarify involves behaviors that promote understanding, including explaining ideas in multiple ways to clear up confusion. Challenge means pressing students to work hard and think hard, confronting students if their effort is unsatisfactory.[27]

- **On average, our students give the least number of favorable responses in the areas of Control and Captivate.** Control means classroom management skills that teachers use to foster effective communication and focus and make the classroom calm and emotionally safe. Captivate means making instruction stimulating and relevant to students. This information suggests that when supporting teachers' professional growth, we should pay particular attention to these areas.[28]

- **Students in elementary grades give more favorable responses on average – a notable drop occurs between upper elementary and secondary middle school.** In Pittsburgh as in other cities nationally, students tend to have less favorable perceptions of their teachers regarding the 7 Cs as they transition from elementary to secondary schools. In Pittsburgh, there was a drop of 15 points between upper elementary (3-5) and secondary middle (6-8) school overall scores, from 67% down to 52% favorable responses. From this information we might suggest focusing supports relative to the 7 Cs for teachers in secondary grades.

- **Teachers and principals can use this information as a powerful tool for change.** Teachers have had confidential access to their students' Tripod survey responses for two years now, and this year they will see those scores pulled together with their classroom observation scores and student growth data. Additionally, principals have access to school-level Tripod data and can use it to focus supports for teachers in their building.

---

[27] R.F. Ferguson (2010).
[28] R.F. Ferguson (2010).

Tripod Student Survey: % Favorable Responses, 2011-12*
* administered to over 90% of teachers (1,800+)

Early Elementary (K-2)
- Consolidate, 78
- Confer, 75
- Captivate, 71
- Challenge, 82
- Clarify, 82
- Control, 53
- Care, 85
- 7 Cs Composite, 76

Upper Elementary (3-5)
- Consolidate, 68
- Confer, 62
- Captivate, 58
- Challenge, 77
- Clarify, 78
- Control, 47
- Care, 78
- 7 Cs Composite, 67

Secondary Middle (6-8)
- Consolidate, 56
- Confer, 47
- Captivate, 47
- Challenge, 66
- Clarify, 59
- Control, 39
- Care, 50
- 7 Cs Composite, 52

Secondary High (9-12)
- Consolidate, 54
- Confer, 48
- Captivate, 48
- Challenge, 62
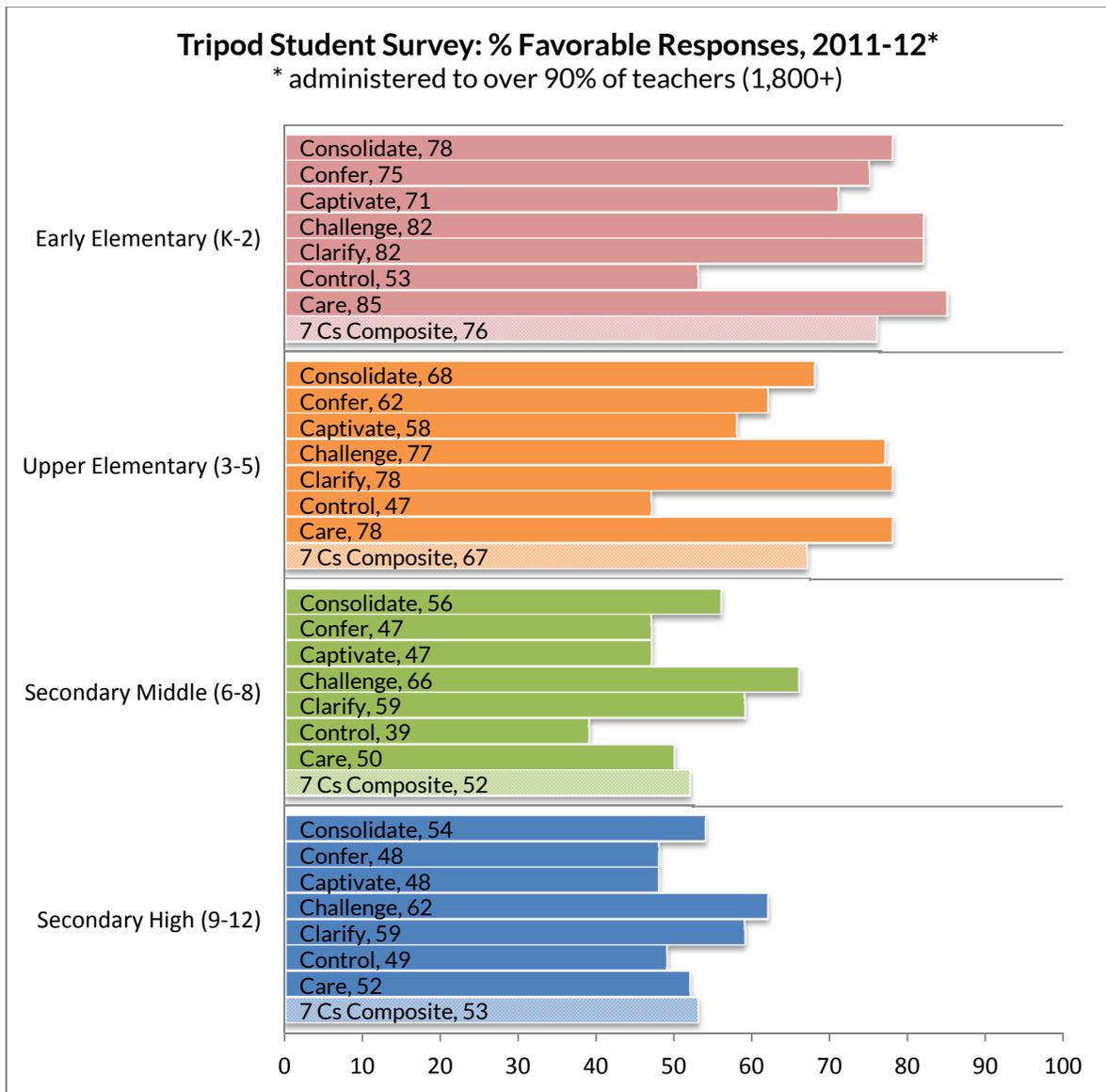- Clarify, 59
- Control, 49
- Care, 52
- 7 Cs Composite, 53

Figure 3: Tripod Student Survey Results, 2011-12. For a description of the 7 Cs, please see our Tool Box at the end of this brief.

How does PPS' implementation of Tripod measure up to criteria for a strong evaluation system?

Criteria for strong evaluation system	What PPS does well with respect to Tripod	How PPS could strengthen Tripod's contribution to the system
Reliability	Proctors (administrators and teachers) administer these surveys to students.[29] Tripod tells us how students—who spend the most time with teachers—experience significant aspects of classroom life and teaching practice and has been validated through over a decade of research and implementation in schools across the world.	PPS should ensure that surveys are administered consistently over time.
Accuracy	PPS proctors use a standard script from Cambridge Education in administering the Tripod Survey to students.	Although there is evidence to suggest it, PPS should regularly verify that teachers with higher scores also have higher achievement gains on average.[30]
Provides timely and meaningful feedback	Feedback, when used together with classroom observation ratings and value-added estimates, could be very meaningful in understanding how student engagement relates to classroom practice and results in growth.	Teachers receive confidential results months after students take the surveys – results should be made available to teachers as soon as possible after survey results are compiled.

---

[29] Pittsburgh Public Schools. *Professional Growth System: About the Tripod Student Perceptions Survey.*
[30] For example, Mathematica Policy Research Inc. examined the construct validity of the most recent value-added estimates and is expected to release a report on that validity study in the fall of 2013.

Conclusion

The Pittsburgh Public Schools and Pittsburgh Federation of Teachers have embarked on a long and important journey to build a high quality teacher evaluation system where great teaching is defined, identified, supported and nurtured to the benefit of students. This journey began where most districts are now – with a binary teacher evaluation that rated teachers as either satisfactory or unsatisfactory, provided very little feedback for improvement and little ability to differentiate excellence from average or from poor, and that ultimately left students and teachers ill served. Today, after five years of intensive partnership work, and collaboration Pittsburgh is among pioneering districts across the country building new, more robust, more helpful teacher evaluations built on research based best practices.

Pittsburgh's system compares favorably to national best practices and should be commended as such. We find it important that these measures have been piloted and studied prior to their application for decision-making. Teachers gave and received feedback on their RISE scores, value-added scores, and on the meaning of the Tripod survey. This feedback has not been tied to any accountability mechanism, but it will in the future. Compared with other districts across the country, PPS has introduced teachers to the evaluation system in an intentional and controlled way over time, not in one sweeping policy change.

But this journey is not over; there is a need to ***continue, accelerate, and strengthen*** the teacher evaluation system to be able to provide teachers, school leaders, and parents with the information they'll need to ensure every child benefits from great teaching and that student learning is accelerated. Our teachers and students will be best served if PPS continues to (1) improve and refine the current measures as well as seek out new ways to understand teaching and its effects on student learning, (2) search for meaningful measures to add to the system that can help teachers better understand what is effective, and (3) make sure that any improvements made stay focused on providing teachers quality feedback so they can better help students reach and exceed academic standards.

 Specifically, we urge PPS and PFT to address these areas to strengthen the new evaluation system:

Provide timely and meaningful feedback to teachers:
- PPS should track the extent to which teachers find feedback from multiple measures helpful and actionable for improving their practice and create mechanisms to adjust feedback accordingly
- **Tripod** – provide survey results to teachers during the school year so they have a chance to reflect and adjust their practice to act on their students' feedback

Ensure increased reliability of the system:
- **RISE** – ensure that evaluators remain reliable over time by mandating ongoing recertification, double-scoring some teachers with impartial observers, observing teachers

multiple times per year regardless of tenure status, and ensuring each teacher is evaluated by multiple observers

- **Tripod** - ensure that surveys are administered consistently over time

Ensure increased accuracy of the system:
- **RISE, Tripod, and VAM** - PPS should regularly verify that teachers with higher scores also have higher achievement gains on average
- **VAM** - ensure that the tests used to calculate value-added estimates are well-aligned with curriculum and that the test broadly represents the skills teachers focused on over the year

The next steps in this important journey must not only include the improvements suggested above, but also deliberate work to figure out how to make the evaluation system more *useful* to teachers and school leaders and more *beneficial* to students. Here are a few places to start:

- Use the information gathered from RISE and Tripod to design meaningful professional development for teachers at the building and individual teacher level. Our district level analysis suggests that professional development and support on the following dimensions would help teachers improve, but principals should work with teachers to identify professional development and support most aligned with individual teacher needs in their school:
    - Classroom Control – or management, including positive behavioral support and social and emotional learning support
    - Student Engagement – making instruction stimulating and relevant to students (Captivate); using questioning and discussion techniques and engaging students in learning
- Empower principals to use teaching effectiveness measures as well as site selection criteria in making decisions about who gets to teach in the school they are responsible for leading. Principals and teachers have more reliable, objective data about teaching than they have ever had, with principals receiving teacher-level data beyond RISE scores for the first time this year. Principals must be able to use this information to hire, retain, support and fire when necessary. Teachers should be assured that effectiveness counts and that their efforts to improve, not just longevity, are being valued.

Finally, strengthening our system of teacher evaluation and feedback is just one piece of the larger pie with which we must concern ourselves. We urge PPS to strengthen its work in these areas as well:

- **Fostering positive teaching and learning environments** – including stable work environments with opportunities for teachers to collaborate and learn from their colleagues & participate in leadership and decision-making
- **Nurturing school leaders who create strong instructional cultures that foster great teaching**[31] **and positive learning environments**

---

[31] The New Teacher Project (2012).

- **Continuing to engage teachers in innovation and reform**
- **Providing for the long-term financial sustainability of the district**
- **Allocating resources to deliver education more equitably to our most vulnerable students**

Great teaching ultimately results in our children succeeding -we must advocate that the new teacher evaluation system does what it's meant to do: help our teachers to refine and improve their practice, help teachers and other professionals to support their growth, and help ensure that every student benefits from great classroom instruction.

Tool Box - What can you do to support student success through great teaching?

If You Are An Advocate

Sometimes it's difficult to know how to support great teaching so we can help our students succeed. We've asked teachers, principals, counselors this question, and we've also taken a look at what should be in place along with an effective system of evaluation.

Invest time in schools. Visit them, connect them with resources, and spread the word about the importance of education – we heard this from principals, counselors, and teachers over the past four years of School Works interviews.

Join A+ Schools in advocating for the changes we lay out in this brief. Email info@aplusschools.org or call 412.697.1298 to join our growing list of community partners committed to equity and excellence for all of Pittsburgh's children.

If You Are A Parent

It is important to build strong relationships with your child's teacher and school leader so that you can be a powerful partner in your child's academic success—this means being empowered to ask questions about the quality and rigor of teaching and doing your best to help your child succeed.

> *As much as you may want to know, please DO NOT ask principals or teachers for evaluation scores! Pennsylvania Act 82 prohibits school districts from releasing teachers' evaluation scores to the public.*

Ask your child:
- What homework, tests, quizzes, and/or projects do you have this week?
- Do you feel comfortable asking questions in class and requesting support from your teacher?
- What subject(s) do you like best? What subject(s) do you have trouble with, and do you think you need support?
- Do you agree that...
  - ...your class stays busy and doesn't waste time?
  - ...your teacher explains difficult things clearly?
  - ...you like the ways you learn in this class?
  - ...you learn a lot in this class every day?

If you find your child doesn't agree with these statements, find a constructive way to discuss your concerns with the teacher and/or principal. Ask A+ Schools for help.

Ask your child's teacher:
- What are your goals or expectations for my child's performance in your class, and how can I help support my child in meeting these goals and expectations?
- At what grade or proficiency level is my child currently performing in core subjects (math, English, science, and history/social studies)? How do you know?

- o If your child is not performing at his/her grade level: What kind of supports will you be providing to bring my child up to the appropriate level by year's end? How can we work together to identify specific needs or supports that will help?
- o If your child is performing at grade level: How do you know throughout the year if my child is still on grade level? How can my child's learning be accelerated or enriched? How do you know if students are 'getting' the subject? How do you help struggling students?
- How will you communicate my child's progress to me, and how often should I expect an update?
- Ask to go over samples of your child's work with his or her teacher and ask about the quality of the work. How does it compare with the state and/or district standards? What could make it better?
- Does my child complete his or her homework, seem happy in school, and participate in class?

Ask your child's principal:
- How do you help teachers improve? Does your school have an area of focus for instructional improvement?
- How well do students in my child's grade level perform in this school? Are there patterns of performance, and if so, do you know why that's happening? What support do you provide teachers to help them improve achievement in that grade?
- How do you define an effective teacher? What is being done to ensure every child at this school is benefiting from excellent classroom instruction? How do you know its working?
- What do you do to support new teachers and retain effective teachers?
- What are some ways I can be a full partner in my child's academic success?

Talk to other parents about great teaching:
- **Improving teaching quality helps students**: If we want to close achievement gaps and ensure all of our children graduate ready for college or career training, great teaching is the one thing schools provide that will most contribute to a student's academic success.
- **Pittsburgh's evaluation system is really a teacher improvement system:** The critical partnership between the PFT and PPS and the careful piloting of different measures to get feedback from teachers and principals about what works has made this system strong and focused where it should be: helping all teachers get better.
- **Evaluation is important and should be bolstered by improved conditions that support effective teaching**: These supports include quality professional development and training, positive learning environments, and strong school leadership.
- **No system is perfect and improvements should continue to be made**: Because it is critical to maintain a teacher evaluation system that is reliable, accurate, and provides meaningful feedback for improvement, PPS and the PFT should be willing to refine and improve the system based on growing research about teaching and learning.

Tripod Student Survey – the 7 Cs

The following information was provided by Pittsburgh Public Schools and Ronald Ferguson, founder of the Tripod student survey.[32]

The Tripod student survey provides detailed information about students' experience in the classroom (7 Cs) and their engagement in learning from the students' own perspectives. The 7 Cs are the central constructs in the Tripod Project framework for measuring teaching effectiveness. Each construct is supported by research in peer reviewed publications that have appeared in education books and journals over the past several decades. The seven are the following:

1. *Care* pertains to teacher **behaviors that help students to feel emotionally safe and to rely on the teacher to be a dependable ally in the classroom.** Caring teachers work hard, and they go out of their way to help. They signal to their students, "I want you to be happy and successful, and I will work hard to serve your best interest; your success is an important source of my personal satisfaction." An example of a Tripod survey item measuring Care is: *"My teacher really tries to understand how students feel about things."*

2. *Control* pertains to classroom management. Teachers need skills to manage student propensities towards off-task or out-of-order behaviors, in order to foster conditions in the classroom that allow for effective communication and focus. Effective control helps to maintain order and supplements caring in making the classroom calm and emotionally safe from such things as negative peer pressures. An example of a Tripod survey item measuring Control is: *"Our class stays busy and doesn't waste time."*

3. *Clarify* concerns teacher behaviors that promote understanding. Interactions that clear up confusion and help students persevere are especially important. To be most effective, teachers should be able to diagnose students' skills and knowledge, and they need multiple ways of explaining ideas that are likely to be difficult for students to grasp. Teachers also must judge how much information students can absorb at any one time, and they should differentiate instruction according to individual maturity and interest. An example of a Tripod survey item measuring Clarify is: *"My teacher has several good ways to explain each topic that we cover in this class."*

4. *Challenge* concerns both effort and rigor -- pressing students to work hard and to think hard. . Challenging teachers tend to monitor student effort and to confront students if their effort is unsatisfactory. Students who do not devote enough time to their work or who give up too easily in the face of difficulty are pushed to do more. Similarly, students who do not think deeply or to reason their way through challenging questions are both supported and pushed. An example of a Tripod survey question measuring Challenge for effort is: *"In this class, my teacher accepts nothing less than our full effort."* A question measuring Challenge for rigorous thinking is: *"My teacher wants us to use our thinking skills, not just memorize things."*

5. *Captivate* concerns teacher behaviors that make instruction stimulating, instead of boring. Captivating teachers make the material interesting, often by making it seem relevant to things about which students already care. Brain research establishes clearly that stimulating learning experiences and relevant material make lessons easier to remember

---

[32] R.F. Ferguson (2010).

than when the experience is boring and the material seems irrelevant. Examples of questions concerning stimulation and relevance are: *"My teacher makes lessons interesting."* and *"[negatively worded]***I often feel like this class has nothing to do with real life outside school."**

6. *Confer* concerns seeking students' points of view by asking them questions and inviting them to express themselves. When students expect that the teacher might call on them to speak in class, they have an incentive to stay alert. In addition, believing that the teacher values their points of view provides positive reinforcement for the effort that it takes to formulate a perspective in the first place. Further, if students are asked to respond not only to the teacher, but to one another as well, a learning community may develop in the classroom, with all of the attended social reinforcements. An example of a question concerning *Confer* is: **"My teacher gives us time to explain our ideas."**

7. *Consolidate* is the seventh *C*. Consolidation concerns how teachers help students to organize material for more effective encoding in memory and for more efficient reasoning. These practices include reviewing and summarizing material at the end of classes and connecting ideas to material covered in previous lessons. Teachers who excel at consolidation talk about the relationships between ideas and help students to see patterns. There is a large body of evidence supporting the hypothesis that these types of instructional activities enhance retention by building multiple brain pathways for retrieving knowledge and for combining disparate bits of knowledge in effective reasoning. An example of a question concerning Consolidation is: *"My teacher takes the time to summarize what we learn each day."*

An in-depth look at the multiple measures used by PPS and what they tell us

As you can see in the chart below, each component of a teacher's final evaluation tells us different information. Using a combined measure makes the system stronger and more accurate, and more meaningful for teachers.

	Classroom Observation (RISE)	Student Growth (VAM)	Student Feedback (Tripod)	Indicator (7 Cs and Domains)
Set high expectations and clear objectives for all students	✓		✓	Challenge 1c – setting instructional outcomes
Use knowledge of students to design instruction	✓			1b – demonstrating knowledge of students
Show expertise and flexibility in the subjects they teach	✓			1a – demonstrating knowledge of content and pedagogy
Establish a culture of learning recognizing effort and quality	✓			2b – establishing a culture for learning
Develop and effectively manage classrooms	✓		✓	Control 2d – managing student behavior
Ask students questions and promote rich discussions	✓		✓	Confer 3b – using questioning and discussion techniques
Continuously assess student learning and provide feedback	✓			3d – using assessment to inform instruction
Engage students regardless of race or academic level	✓		✓	Clarify, Capitvate 3c – engaging students in learning; 3g – implementing lessons equitably
Show students they care about them			✓	Care
Help students draw connections between ideas			✓	Consolidate
Initiate frequent, positive communication with parents	✓			4f – communicating with families
Help their students grow academically	✓	✓		3f – impacts student growth
Understand how their practices contribute to student growth	✓	✓	✓	4a – reflecting on teaching and student learning

This information is based on research-based evidence of what effective teachers do:
- L. Darling-Hammond, A. Emrein-Beardsley, E.H. Haertel, and J. Rothstein (2011).
- R.F. Ferguson (2010).

References

A+ Schools (2012). *2012 Report to the Community on Public School Progress in Pittsburgh.* Pittsburgh, PA.

American Federation of Teachers. *A Guide for Developing Multiple Measures for Teacher Development and Evaluation*, 5. [PDF document]. Retrieved from http://www.aft.org/pdfs/teachers/devmultiplemeasures.pdf.

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., ... & Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers* (Vol. 278). Washington, DC: Economic Policy Institute.

Darling-Hammond, L., Emrein-Beardsley, A., Haertel, E.H., & Rothstein, J. (2011). *Getting teacher evaluation right: a background paper for policy makers.* American Educational Research Association and National Academy of Education.

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, *93*(6), 8-15.

Ferguson, R.F. (2010). *Student Perceptions of Teaching Effectiveness*. National Center for Teacher Effectiveness and the Achievement Gap Initiative at Harvard University.

Goddard, Y. & Goddard, R.D. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record 109*(4):877-896.

Gordon, R. J., Kane, T. J., & Staiger, D. (2006). *Identifying effective teachers using performance on the job*. Washington, DC: Brookings Institution.

Harris, D. N. (2011). Report to the Florida Education Association on Florida's New Teacher Performance and Compensation System.

Jackson, C.K. & Bruegmann, E. (2009). *Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers.* Washington, DC: National Bureau of Economic Research.

Kane, T., & Staiger, D. O. (2012). Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains. The Bill and Melinda Gates Foundation.

Mathematica Policy Research, Inc. (2010). *Estimating Teacher and School Effectiveness in Pittsburgh: Value-Added Modeling and Results.*

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating Value-Added Models for Teacher Accountability. Monograph*. RAND Corporation. Santa Monica, CA.

Pittsburgh Public Schools. (1/3/ 2013). *Empowering Effective Teachers Education Committee Update* [PDF document]. Retrieved from

http://www.pps.k12.pa.us/cms/lib07/PA01000449/Centricity/domain/19/committees/education/2013/01/121227_-_EET_EduCom_Final.pdf

Pittsburgh Public Schools (8/11/2011). *Pittsburgh RISE: Research-based, Inclusive System of Evaluation Version 10a. Pittsburgh Standards of Effective Teaching: Administrator's Formal Observation Rubric.* [PDF document].

Pittsburgh Public Schools. *Professional Growth System: About the Tripod Student Perceptions Survey.* [PDF document]. Retrieved from http://www.pps.k12.pa.us/cms/lib07/PA01000449/Centricity/domain/30/document%20library/Tripod%20faq.pdf.

Raudenbush, S. W. (2004). What are value-added models estimating and what does this imply for statistical practice?. *Journal of Educational and Behavioral Statistics*, *29*(1), 121-129.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, *125*(1), 175-214.

Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*(1), 57-67.

Scherrer, J. (2011). Measuring Teaching Using Value-Added Modeling: The Imperfect Panacea. *NASSP Bulletin*, *95*(2), 122-140.

Schott Foundation for Public Education (2012). *The Urgency of Now: The Schott 50 State Report on Public Education and Black Males.* Cambridge, MA.

The New Teacher Project. (2012). *Greenhouse Schools: How Schools Can Build Cultures Where Teachers and Students Thrive.*

Wiley, E. W. (2006). A practitioner's guide to value added assessment. *Educational Policy Studies Laboratory Research Monograph.Tempe: Arizona State University*.