

**What Can Parents Tell Us About Teacher Quality?  
Examining the Contributions of Parent Perspectives in Comparison to a  
Portfolio of Alternative Teacher Evaluation Measures.**

Benjamin Master ([bmaster@stanford.edu](mailto:bmaster@stanford.edu))

*Stanford University*

## **Abstract**

Parent perspectives about their children's K-12 teachers have received increased attention in recent years as a potential performance measure for use in teacher evaluation systems. Parent evaluations of teachers may be valuable both because they may identify and provide feedback regarding important teacher contributions that support parent involvement in schooling, and because they may inform the assessment of teacher quality more broadly. However, relatively little research has explored the types of information that parents can provide about teachers, or the extent to which their perspectives provide distinct information in comparison to other teacher evaluation measures. I examine parent perspectives about teachers alongside a rich portfolio of alternative teacher evaluation measures including student surveys, formal lesson observations, and value added measures of teachers' student achievement gains. I find that parent perspectives offer a distinct and stable teacher performance measure over time, but one that is fairly one-dimensional. They are most closely aligned with student perspectives, but are also associated with specific components from formal lesson observation ratings, including teachers' success in fostering a positive classroom environment and encouraging student engagement. Parent perspectives are also predictive of teachers' math value added performance ratings in the same school year. However, they contribute little additional information for predicting value added ratings when considered in tandem with alternative measures, which suggests that their value as a proxy measure of teachers' value added to student achievement outcomes is limited in comparison to other evaluation measures.

## **Introduction**

The federal Race to the Top Initiative, as well as additional reforms enacted via waivers to federal No Child Left Behind legislation requirements, have focused unprecedented attention on the development of new teacher evaluation systems in K-12 schools. Evaluation systems may improve the quality of teaching via two key mechanisms. First, they may identify and promote effective teaching practices to help teachers to improve (Taylor and Tyler, 2011). Second, they may facilitate personnel practices and policies that support the retention of more effective teachers and the dismissal of less effective teachers, as well as more optimal assignment of teachers to jobs in which they can have the most positive effect (Boyd, Lankford, Loeb, and Wyckoff, 2010; Goldhaber and Theobald, 2011; Rockoff, Staiger, Kane, and Taylor, 2011).

In support of these aims, educators may benefit from employing measures of teacher quality that reflect the full spectrum of valued teacher contributions, and that also provide useful formative feedback that can help teachers to prioritize and improve their practice. Thus far, much of the emphasis in current reforms has been devoted to value-added measures of teacher effectiveness based on student test performance. While these measures assess a key schooling outcome, they provide little utility as formative tools for teacher development and are often unavailable in some subjects and grade levels. In addition, they may do a poor job of accounting for teacher impacts on student outcomes other than annual tested achievement (Rothstein, Ladd, Ravitch, Baker, Barton, Darling-Hammond, Haertel, Linn, Shavelson, and Shepard, 2010).

In light of the limitations of value added measures, educators and researchers have also devoted substantial attention to developing a broader portfolio of teacher quality measures for use in evaluation systems. Two prominent examples of this trend are surveys of student perspectives and formal lesson observations by trained observers. Lesson observation

instruments, such as the Framework for Effective Teaching (FFT), the Classroom Assessment Scoring System (CLASS), and the Protocol for Language Arts Teaching Observations (PLATO), offer rich information about teachers' behaviors in light of explicit performance standards that highlight desired teaching practices. For these reasons, they may be useful tools for aligning teacher practices and supporting formative staff improvement efforts in schools. High quality lesson observation measures also demonstrate modest correlations with teachers' value added to student achievement, particularly when instructional ratings are collected across multiple observations (Gates Foundation, 2012; Kane, Taylor, Tyler, and Wooten, 2011; Hamre and Pianta, 2005). Surveyed student perspectives tend to correlate more highly than lesson observations with teachers' value added to student achievement (Gates Foundation, 2012), and may also provide potentially valuable formative feedback to teachers regarding their students' classroom experience. However, student surveys may not be as readily accepted by stakeholders as legitimate for use in high stakes teacher evaluations, particularly when perspectives are gathered from very young children.

Parent perspectives about teachers have received less attention as a potential evaluation measure than either student perspectives or formal lesson observations, but there are several reasons to believe that they may add value when employed as part of comprehensive teacher evaluation systems. First, parent perspectives of teachers may provide useful information about a key aspect of teachers' work: engaging parents in their children's schooling experience. The inclusion of parent perspectives in evaluations may also incentive teachers to focus on this aspect of the job. Parental engagement in schooling is associated with both student achievement gains (Sheldon and Epstein, 2005; Powell-Smith, Stoner, Shinn, and Good, 2000) and with a range of other valued student outcomes and behaviors (Fan and Williams, 2010; Domina, 2005; Simon,

2001; Sirvani, 2007a). Moreover, some school-based interventions that encourage parent-teacher interaction have been shown to increase student motivation, engagement, and performance (Kraft and Dougherty 2012; Sirvani, 2007b; Bursztyn and Coffman, 2011). Second, aside from teacher contributions to parent engagement, parent perspectives about teachers may also provide useful information about teacher quality more broadly. Parents can draw from what they see and discuss with their child, as well as from interactions with their child's teacher, to inform their perspectives. Thus, they may possess unique information about teachers' overall competency.

Other potential advantages of leveraging parent perspectives in teacher evaluation relate to issues of implementation. Like student surveys, parent surveys are relatively inexpensive to administer and can provide timely evaluative feedback based on a single administration. Moreover, parent surveys are feasible to conduct across a wide range of subject areas and grade levels, and may be more readily accepted by teachers than surveys of very young students for use in high stakes evaluation. The broad applicability of parent surveys may be particularly useful in light of the limitations of value added measures, which are typically not available in all subject areas and grade levels. If parent surveys can serve as a partial proxy for other valued measures that may not be assessed for all teachers in an evaluation system, this would add another dimension to their utility. Finally, aside from the information parents may provide, parents constitute a key stakeholder group for schools and the act of including their perspectives in evaluation systems may by itself help to engender positive parent-school relations.

In light of these potential benefits of parent surveys, several states, including Connecticut, Utah, Idaho, and Florida, as well as many individual school districts nationwide have already begun using or piloting parent perspectives in teacher evaluation systems (Advocates for

Children, 2012). The rapid adoption of these measures highlights the need for additional research that can inform policy makers and school leaders about their potential utility and limitations.

### *Parent Perspectives for Teacher Evaluation*

Extant research on the use of parent perspectives for teacher evaluation is limited, but offers some insight into the type of information that they are likely to provide. For instance, research indicates that parent evaluations tend to be quite positive, but still sufficiently distributed to meaningfully differentiate between individual teachers. (Peterson, Wahlquist, Esparza Brown, and Mukhopadhyay, 2003; Peterson, 1987; Peterson, 1989). In addition, variation in parent ratings appears to reflect meaningful differences between teachers, distinct from school or respondent-level rating variance (Peterson et. al., 2003). In the same study, Peterson and colleagues find that different parent ratings of the same teacher also appear moderately stable from one year to the next, with a correlation of 0.41. However this finding is based on a small sample of just 28 teachers over two years, and it is unclear the extent to which that level of consistency reflects school-wide stability in parent ratings over time, versus stability in individual teacher performance relative to peer teachers within the same school.

In addition to research exploring those foundational characteristics of parent survey measures, a few studies have considered the degree of alignment between parent perspectives about teachers and other teacher evaluation measures. Most research in this vein compares parent perspectives with administrators' ratings of teachers, and suggests a limited and inconsistent degree of alignment between the two (Epstein, 1985; Peterson, 1987; Ostrander, 1995). For example, Epstein (1985) finds that administrators' ratings of teachers align well with parents' observations of teachers' explicit parent-engagement activities, but are less aligned with parents'

observations about teachers' general interpersonal skills. In contrast, Peterson (1987) finds that parent perspectives about teachers appear to align quite closely with student perspectives about teachers. This is not surprising, since parents' perspectives are likely to be largely informed by their own children's reports about their classroom experience.

Collectively, this body of evidence offers some guidance to educators and policymakers weighing the use of parent perspectives in teacher evaluations, but important gaps remain with respect to both the utility and limitations of these measures. First, previous research does not clarify the extent to which parent measures capture a stable attribute of individual teachers, distinct from stability in school-wide parent ratings of teachers. Second, the degree of overlap between parent measures of teachers and other common teacher evaluation measures has received relatively little attention. Additional research comparing alternative teacher performance measures can help to identify where each may provide either similar or distinct information about teacher practices, when used in tandem in comprehensive evaluation systems. Measures that are similar may potentially serve as partial proxies for each other when either is unavailable, while measures that are more distinct may provide unique feedback about teacher performance. Finally, no previous research has linked parent ratings of teachers with measures of teachers' impact on student outcomes, such as their value added to student achievement gains.

This study provides a substantial contribution towards advancing our understanding of parent evaluation measures, by addressing each of the aforementioned gaps in the prior research. I examine what parents can tell us about stable aspects of individual teacher quality – distinct from school quality – by testing how well parent ratings of a teacher in one school year predict their parent ratings in the next, using both overall and within-school relative teacher ratings. In addition, I compare parent ratings of teachers to an unusually broad range of alternative teacher

performance measures, including student ratings of teachers, lesson observation ratings of teachers, and measures of teachers' value added to student achievement. In the case of student ratings and lesson observation ratings, I identify which specific components from those measures are most closely aligned with parent ratings of teachers. This offers a richer view than prior research about the specific types of information that parents are more likely to pick up on. Finally, I examine the extent to which student ratings, lesson observation ratings, and parent ratings predict teacher value added measures, either individually or when considered in tandem. This is important because value added measures reflect one key outcome of teachers' work. The degree of alignment between parent ratings and teachers' student achievement gains can provide some external validation of the information provided by parents. Additionally, if using parent measures alongside other measures improves schools' ability to predict teachers' value added performance ratings, this may be of practical value to districts seeking to identify proxy measures in contexts where value added measures are unavailable.

In light of the previous discussion about parent perspectives, I specifically examine the following five questions of interest:

1. To what extent do parent perspectives about teachers provide meaningful differentiation between teachers, both within and across schools?
2. Do parent perspectives about teachers represent a stable attribute of teachers that is consistent across different school years and classrooms of students?
3. To what extent do parent perspectives align with other teacher evaluation measures such as student surveys, lesson observations, and teacher value added measures?
4. What specific questions or dimensions from student surveys or lesson observations are more or less closely aligned with parent perspectives?



5. When used in tandem with student survey or lesson observation measures, do parent perspectives help to improve predictions of teachers' value added performance ratings?

I address these questions by examining parent perspectives about teachers over time and alongside an unusually rich portfolio of alternative teacher evaluation measures. I find that parent perspectives offer a distinct and stable teacher performance measure over time, but one that is fairly one-dimensional. They are most closely aligned with student perspectives, but are also associated with specific components from formal lesson observation ratings, including teachers' success in fostering a positive classroom environment and encouraging student engagement. Parent perspectives are also predictive of teachers' math value added performance ratings in the same school year. However, they contribute little additional information for predicting value added measures when considered in tandem with alternative measures, which suggests that their value as a proxy measure of teachers' value added to student achievement gains is limited in comparison to other evaluation measures.

### **District Context and Data**

My data come from a network of public charter schools that operate alongside state public schools under a single centralized district management team (hereafter referred to as “the district”). The district includes a mix of elementary, middle, and high schools, attended by approximately 5,000-10,000 students. In order to help preserve district anonymity, I do not include precise details about the size or number of schools examined in this study.<sup>1</sup> District schools serve an over-subscribed, lottery-selected population of predominantly poor and minority students across grades K-12. Seventy-five percent of district students qualify for free or

reduced price lunch, and the student population is made up of eighty percent African American students and nineteen percent Hispanic students.

Beginning in school year 2010-2011, the district began piloting new teacher evaluation measures in eight of its schools. In the following school year 2011-2012, the district implemented a new evaluation system using these measures, and included all district schools and almost all of the approximately 450 full-time teachers in the district. The purpose of the teacher evaluation system is to both provide formative feedback to teachers regarding their performance, as well as to identify, formally recognize, and differentially compensate “expert” teachers. The system includes multiple measures of teacher quality, including parent surveys about teachers, student surveys about teachers, formal lesson observations of teachers’ instruction, and value added measures of student achievement gains in teachers’ classrooms. The inclusion of parent perspectives as part of teacher evaluations in the district reflects the importance that school and district leaders place on parent-teacher relations. District leaders express the explicit expectation that teachers will maintain regular contact with parents as part of their core work responsibilities.

In the rest of this section, I detail, in sequence, each of the individual measures that the district collects as part of its teacher evaluation system and that are considered in this study. These include parent surveys about teachers, student surveys about teachers, and formal lesson observations of teachers’ practice. Finally, I describe the available student achievement and demographic data from the district, which I use to create teacher value added measures.

### *Parent Surveys*

The district collected parent surveys regarding 96 teachers in school year 2010-11, and 444 teachers in school year 2011-12, encompassing all grade levels K-12 and all subject areas except

for music, physical education, and art. For each participating teacher, the district collected responses from a random sample of up to 32 families of students from across all of the teacher's classrooms. The average response rate for the surveys was approximately 80 percent. For those surveys with one (31 percent) or more than one (8 percent) unanswered survey item, I use an Iterative Chained Procedure to impute the missing data, using data from all of the non-missing survey items.

The questions on the parent survey were selected by district administrators, and include items addressing a range of topics, including the quality of communications between parents and the teacher, the teacher's professionalism in interactions with the parents, the perceived treatment of the parents' child by the teacher, and parents' assessment of the teacher's positive impact on their child, both academically and socially. Each of the 11 survey items are detailed in Table 1. All questions were rated on a five point Likert scale.

### *Student Surveys*

The district collected student survey data for 329 teachers in school year 2011-12, for teachers in grades 3 to 12. For each participating teacher, the district administered surveys in-school to a random sample of up to 24 students from across all of the teacher's classrooms. The average survey completion rate for students was approximately 90%. For those surveys with one (7 percent) or more than one (2 percent) unanswered survey item, I use an Iterative Chained Procedure to impute the missing data.

District administrators selected the questions on the student survey, borrowing partially from existing instruments that have been used in other research, including the Measures of Effective Teaching (MET) project (Gates Foundation, 2012). Five of the 13 student survey questions are

identical to survey questions from the MET work, while 2 others are very similar. These seven questions address multiple aspects of students' experience with their teachers, and are drawn from the areas labeled as "Challenge," "Control," "Care," and "Clarify" in the MET research. That work has also identified these questions as effective predictors of teachers' value added performance. Other questions on the student survey instrument address areas related to teachers' expectations for students, their recognition of student's work, and the students' general feelings regarding their teacher's class. The 13 items from the student survey are detailed in Table 3. All questions were rated on a five point Likert scale.

### *Lesson Observations*

Formal lesson observation data provide additional information about teachers' instructional practices in school year 2011-12. The district collected lesson observations for 485 teachers, in each of three windows in the Fall, Winter, and Spring of that school year, for all teachers in all subject areas and grades. To facilitate the observations, the district employed a locally-created observational rubric consisting of 16 individual dimensions of teachers' instructional practice, each rated on a five point Likert scale. Conceptually, the rubric's 16 dimensions are organized within 4 domains of instructional practice that relate to: 1) the degree of academic rigor and clarity in the lesson, 2) the design and delivery of core lesson elements, 3) the development of a positive classroom culture, and to 4) ensuring student learning occurs. A brief description of each of the 16 dimensions, and the domain (labeled "D1" to D4") that each is a part of, is included in Table 6. The average overall teacher rating score on the lesson observations was 2.77 (out of 5) across all rubric dimensions, with a standard deviation of 0.59.

District leaders designed the observation rubric after an extended process of engaging school leaders and teachers in conversation about what they considered to be the most important aspects of instruction. However, the rubric also references external research – particularly the work of Jon Saphier – and includes substantial overlap with some other prominent classroom observation tools. In particular, elements from the rubric that relate to areas such as high quality questioning by teachers, student participation and engagement, clear classroom procedures and routines, the structure and pacing of the lesson, and effective responses to student questioning are similar to those emphasized in the Danielson Framework for Effective Teaching’s “Instruction” domain (Danielson, 2007). Nevertheless, the articulation of specific observational indicators in the district’s rubric is unique. The district’s rubric is designed to assess a set of core instructional practices across all subjects and grade levels, and does not codify subject or grade-specific expectations for teachers’ practice.

Each rubric dimension references multiple indicators. Some indicators measure teacher practices, while others that assess student behavior that may result from teachers’ observed or unobserved efforts. For example, the dimension related to “student engagement” includes indicators for teachers’ appropriately-timed use of different categories of engagement techniques (e.g. alerting, enlisting, or engaging students), as well as indicators for the percentage of students that engage when the teacher asks a question of the whole class. Similarly, indicators for the dimension related to high quality student responses assess both how rigorously the teacher pushes students to improve their responses, as well as how seamlessly students respond to the teacher’s expectations. In contrast, a few dimensions focus just on teacher behaviors. For instance, indicators for the dimension related to “planned, rigorous questioning” include an

assessment of the quality of teachers' planned questions (according to specific criteria), and whether the teacher adapts to student responses when engaging in back and forth discussion.

Each teacher's three observational ratings include a mix of one or two 45-60 observations by a local school administrator with knowledge of that teacher's subject area, and one or two 60-90 observations by a team consisting of both a school administrator and a senior instructional leader from the district central office. All raters received some formal calibration training on the use of the observational rubric, including one or two group training sessions using video-recorded lessons. In addition, the dimensions of teaching emphasized in the rubric had been a mainstay of district professional development efforts for several years. However, the district did not formally test the reliability of all trained raters, and therefore the reliability of the resulting teacher lesson observation ratings is uncertain. In spite of this limitation, the lesson observation data represents a potentially rich source of information about teachers' specific instructional practices.

#### *Administrative Data*

District administrative data include a range of student achievement and demographic measures, and records of student and teacher assignment in school year 2011-12. Student demographic data include students' race, English language learner (ELL) status, special education status, free or reduced price lunch status, gender, grade level, and grade retention status. Spring state test data are available for students in grades 4-8, in math and language arts, from both school year 2011-12, and from the prior school year 2010-11. The data also identify students' teachers in math and language arts. In cases where students were taught by multiple teachers in the same subject area, a dosage variable specifies, for each student, the proportion of total instructional time spent across teachers.

## **Methods**

In this section, I first describe my approach to generating each of the teacher-level performance measures considered in the study. I then detail the methodology I employ to address my specific research questions regarding parent perspectives.

### *Generating Teacher Performance Measures*

*Parent and Student Perspectives.* I calculate teacher-level measures from parent and student survey questions by averaging the ratings given by respondents (students or parents) for each teacher, on each individual survey item. I then create an overall teacher score for each of the surveys by averaging teachers' scores across all of the individual survey items. Finally, I standardize teacher-level scores across all teachers in my sample.

In order to facilitate an analysis of teacher ratings separate from school-wide ratings of teachers, I also create measures that are relative to teachers' within-school peers. To do this, I first calculate a school-level mean survey score across all respondents. Subtracting this school level overall score from the equivalent teacher-level scores yields a within-school teacher-level rating, which I then standardize across the full sample of teachers.

*Lesson Observations Ratings.* To generate teacher scores for each lesson observation dimension, I calculate each teacher's average score across their three lesson observation data points. I standardized average scores on each dimension across my sample. I create a single aggregate observation rating for each teacher by generating a principal component factor from across the standardized teacher scores on the 16 dimensions, and standardize the resulting factor

across teachers. I also generate within-school relative teacher lesson observation scores, in a manner similar to that described for my survey measures.

*Value Added Measures.* I construct value added measures of teacher performance for school year 2011-12 for a total of 146 math and language arts teachers whose students are in grades 4 through 8. I begin by standardizing student test scores within each grade and subject using a sample that includes both district student test scores and student test scores from large external benchmark districts within the same state. I then estimate teacher value added measures using a two-step, aggregated residuals procedure similar to the one detailed by Ehlert, Kohlert, Parsons, and Podgursky (2012). Conceptually, this method aims to assess teachers' relative performance in comparison to other, "similarly circumstanced" teachers, in this case according to the observable characteristics of their students and classrooms. Because teacher-student assignment data is only available for the district (rather than from the external benchmark districts), teacher value added measures are generated relative to a within-district sample.

The two steps of value added estimation proceed as follows:

$$(1) y_{ijt} = \beta_1 y_{ij(t-1)} + \beta_2 y_{ij'(t-1)} + X_{it} \delta + \gamma_{ijt} \sigma + \varepsilon_{ijt}$$

In this first step (1), standardized test scores for student  $i$  in subject  $j$  (math or language arts) in the current year  $t$  are modeled as a function of test scores from the prior year ( $t - 1$ ) in both the same subject  $j$  and alternate subject  $j'$ , as well as a vector  $X$  of student  $i$ 's observable variant and invariant characteristics in the current year, a vector  $\gamma$  of the peer<sup>2</sup> average of the same observable student characteristics, and an error term. Student observable characteristics include indicators for each student's race, English language learner (ELL) status, special education status, free or reduced price lunch status, gender, each grade level, and for whether the student was retained in their grade level from the prior year. Peer average characteristics include the



average peer student test scores in both the same and the alternate subject, as well as the average of peer student demographic characteristics including race, gender, lunch status, special education status, and ELL status.

In the second step, I compute the residuals,  $\hat{\sigma}_{ijt}$ , based on students' predicted test performance, and aggregate those residuals up to students' teachers to create estimated teacher effects,  $\varphi$ .<sup>3</sup> Because many teachers in this district co-teach their students in either the same or different classrooms, I weight the student residuals according to a dosage  $\theta$ , that corresponds to the proportion of each student's time that is spent with each teacher in a given subject.

Aggregating student residuals for each teacher in the manner just described provides a within-district relative ranking of teacher performance, but I am also interested in identifying teachers' performance relative to peers at the same school. To assess the latter, I generate within-school teacher value added measures by aggregating student residuals at the school level, and subtracting the resulting school value added measure,  $\omega$ , from each teacher measure  $\varphi$  within that school.

### *Investigating Parent Perspectives About Teachers*

*Differentiation and Stability.* My first two research questions relate to how parent measures differentiate between individual teachers, and whether they do so by capturing stable teacher attributes. To explore differentiation in parent ratings, I examine descriptive statistics for each of the individual survey questions, as well as the distribution of a standardized overall score for parent ratings of teachers. I then conduct an exploratory factor analysis across the teachers' individual item ratings in the parent survey, in order to determine whether they correspond to one or more than one underlying dimension of teachers' performance. Finally, in order to test

whether parent ratings vary across individual teachers, I conduct an analysis of the variance in parent ratings of teachers that occurs between teachers rather than within teachers, in comparison to the same variance decomposition analysis for student surveys and value added measures.

I examine the stability of parent ratings of teachers by predicting individual teachers' overall and within-school ratings in school year 2011-12 as a function of their own ratings in school year 2010-11, using an ordinary least squares regression model. In this way, I examine how well prior-year parent ratings predict future-year parent ratings, as well as how well prior-year parent ratings predict other future-year teacher evaluation measures, including student ratings and lesson observation ratings. In all models using within-district teacher ratings, I also include indicator variable controls for school type (i.e. elementary, middle and high school). These controls help to explain significant variation in teacher ratings for some of the evaluation measures considered in this study, improving the precision of my estimates of interest.<sup>4</sup>

*Alignment with Alternative Measures.* I address my remaining research questions by examining the alignment between parent ratings of teachers and alternative teacher evaluation measures in several ways. First, I again employ an ordinary least squares regression models to assess how well parent perspectives about teachers in school year 2011-12 predict same-year teacher ratings from students, from formal lesson observations, or from value added measures of teachers' students' achievement gains in math or language arts. In each case, my sample consists of teachers rated both by parents and by each of the other measures. In practice, this means that teachers at all grade levels (K-12) are included when predicting lesson observation ratings with parent ratings, but only those in grades 3-12 are included when assessing alignment with student rating measures, and only a subset of teachers in grades 4-8 in math and language arts are available for assessing alignment with value added measures. In preliminary analysis, I

investigated but did not find that the associations between parent ratings and these alternative measures varied significantly across samples, grade levels or school type. As in my investigation of parent rating stability, I examine separate models using either within-school or within-district measures, and include controls for school type in the latter.

In order to better identify the specific areas of alignment between parent perspectives and alternative measures, I also consider both student surveys and lesson observation ratings in greater detail. Specifically, I use a separate model to predict each of the individual questions on the student survey, and each of the individual dimension ratings from the lesson observations, using parent perspectives in the same or prior school year.

Finally, I examine the extent to which parent perspectives about teachers add distinct value when used alongside other teacher performance measures to predict teachers' same-year value added performance in either math or language arts. To do this, I predict teacher value added measures as a function of either parent ratings, student ratings, or lesson observations (and include no other covariates in the models), and report for each the resulting adjusted R-squared. Then, I explore whether the R-squared for predicting value added measures improves when including parent measures alongside alternative measures in the same model. I also compare this to a prediction using just student survey and lesson observation measures, without parent ratings.

## **Results**

I start by providing some descriptive statistics about parent ratings of teachers, both overall and by survey item. These results are detailed in Table 1. Mean teacher scores are relatively high on a five point scale, consistent with previous research that suggests parents' ratings of teachers tend to be quite positive. The lowest rated survey items relate to parents' reports about

how informed they are about their child's progress, and whether the teacher notifies parents about the positive work their child has done in class. The highest rated items relate to whether parents feel the teacher motivates their child, and parents' overall rating of the teacher on a graded scale (A+ to D). Across the full range of survey questions, parent responses are highly correlated with the same parents' overall rating of a teacher. Not surprisingly given this pattern, an exploratory factor analysis suggests that there is only one common factor across teacher ratings on the individual survey items.<sup>5</sup>

Table 1 also includes statistics for a single principal component factor across items. That factor explains 88.4 percent of the total variance in teacher scores, with an alpha coefficient across all items of 0.955. The degree of uniformity of parent responses across a diverse set of questions suggests that parents may provide feedback about teacher performance that is fairly one-dimensional.

#### *Differentiation and Stability of Parent Survey Measures*

Descriptively, parent measures appear to meaningfully differentiate between teachers, both overall and within a particular school. Figures 1 and 2 show the distribution of the overall average of teachers' parent rating scores, and the within-school centered parent rating scores, respectively. The distribution pattern for overall ratings indicates possible ceiling effects, while within-school scores more closely approximate a normal curve distribution. In addition to these descriptive trends, a formal test indicates that parent ratings of teachers do differentiate between individual teachers. Table 2 details the proportion of variance in parent, student, and value added measures that reflects between-teacher, rather than within-teacher variance. A statistically

significant portion (0.146) of parent survey rating variance occurs between teachers, a result that is comparable to that of student survey and value added measures.<sup>6</sup>

In addition to differentiating teacher performance, parent ratings appear to capture stable attributes of teachers across different school years and classrooms. Table 3 includes a series of models predicting various teacher performance measures from school year 2011-12, using either parent perspectives from school year 2011-12, or parent perspectives from the prior school year 2010-11. Parent ratings of a teacher in the prior school year are highly predictive of that teacher's parent rating in the next school year. A one standard deviation increase in prior-year parent ratings predicts a 0.569 standard deviation higher parent rating in the subsequent year, even though the teacher is working with a different classroom of students and families. In line with this result, the correlation between parent ratings of a teacher from one year to the next is substantial, at 0.691.

The stability of parent ratings of teachers is not simply the result of school-wide stability in parent ratings. Table 4 shows the same models as Table 3, but uses within-school teacher rating measures rather than within-district measures. In this specification, prior year within-school parent ratings remain highly predictive of subsequent parent ratings. An increase of one standard deviation in prior-year within-school parent ratings of a teacher is associated with a 0.710 standard deviation higher within-school parent rating in the subsequent school year. The correlation between within-school parent ratings of a teacher from one year to the next is 0.545.

#### *Alignment with Other Teacher Evaluation Measures*

As expected, both overall and within-school parent perspectives about a teacher are highly correlated with student ratings of that teacher in the same school year. For example, a one

standard deviation increase in parent ratings of a teacher corresponds to a highly significant 0.644 standard deviation higher student rating of that teacher in the same school year, using within-district teacher measures. In addition to their association with student ratings, parent ratings are also significantly associated with both lesson observation ratings of teachers and with math teachers' value added performance ratings in the same school year. A one standard deviation increase in parent ratings of a teacher corresponds to a 0.333 standard deviation higher lesson observation rating, and a 0.365 standard deviation higher math value added rating of the teacher. The same associations using within-school teacher performance measures are similar, but in the case of math value added are notably attenuated. It may be that parent perspectives identify a mix of both individual teacher and school effects on student achievement gains, resulting in a somewhat smaller association with within-school teacher value added measures that do not include any school effects.

Parent ratings from the prior school year also predict student ratings of teachers in the current school year to some extent, at least when using within-school measures. A one standard deviation increase in prior year within-school parent ratings of a teacher corresponds to a 0.353 standard deviation higher within-school student rating of the teacher in the next year. This suggests that students and parents may both be picking up on some of the same stable attributes of teachers' practice. However, prior year parent ratings do not predict future lesson observation ratings. Parent perspectives may be better at picking up on aspects of teachers' instructional practice that relate to the success of their interactions with particular students in the same school year, rather than a more stable measures of their instructional skill-set and behaviors.

*Alignment with Specific Components of Other Evaluation Measures*

The previous results offer some insight into the types of information provided by parent surveys of teachers. I extend this investigation further by examining the associations between parent ratings of teachers in the same and prior school years and either specific student survey questions or individual lesson observation dimensions. Table 5 shows results from separate models that predict teachers' rating on each individual student survey question from school year 2011-12 with either same-year or prior-year parent ratings of the same teachers. While all student survey questions are significantly associated with parent ratings in the same school year, those that explain the most variance in parent ratings relate to students' perception of their teacher as caring, encouraging, and creating a welcoming classroom environment.<sup>7</sup> Student ratings of teachers' support for student achievement and clarity of explanations come next. Relatively less predictive questions include questions related to students' reports of how the class uses time and behaves, and whether the teacher keeps students informed of their progress. Models testing the relationship between student questions and prior-year parent perspectives reveal a similar pattern, but with attenuated effect sizes. Student reports of the teacher being caring and encouraging stand out as the strongest predictor of parent ratings across different years and classrooms, with effect sizes that are the most consistent with the same-year models. This suggests that parents may be best able to detect a stable measure of individual teacher ability in these two areas of teachers' performance.

Table 6 details the results of a similar investigation predicting individual dimensions from teachers' lesson observation ratings with either same-year or prior-year parent perspectives. While the overall associations between lesson observation dimensions and parent perspectives are weaker, the evidence suggests some similar trends in terms of what parents can identify. Lesson rating dimensions related to teachers' success in creating a positive classroom

environment, as well as successfully encouraging student engagement, are the two most strongly associated with parent perspectives, across the separate models. In addition, as shown in Table 7, when controlling for those two dimensions from the lesson observation scores, there is no significant association between parent ratings and any other lesson observation dimension. These results suggest that parent perspectives offer a fairly narrow view of teachers' instructional practice. In addition, even the associations with these two dimensions are not statistically significant in models using prior-year parent ratings as independent variables, which may indicate that parent ratings provide a weak signal regarding stable attributes of teachers' instructional practice.

#### *Predicting Teacher Value Added Performance*

While parent ratings of teachers are modestly predictive of math teacher value added in the same school year, they would be particularly useful as proxies for teacher value added performance if they could provide additional distinct information about teachers' value added performance when considered in tandem with other evaluation measures. I examine this issue and detail my findings in Table 8. Parent ratings of teachers predict less variance in teacher math value added than either student ratings or lesson observation ratings, with an adjusted R-squared of 0.065, compared to 0.208 and 0.143, respectively. In addition, parent ratings of teachers do not appear to provide much distinct information about teachers' value added performance when considered in tandem with student measures. The inclusion of parent ratings alongside student ratings in the same model does little to improve the models' explanatory power. While parent measures do help to explain more variance in value added when used in tandem with lesson observation ratings, this combination still explains less variance than a combination of student



ratings and lesson observations. Overall, these results suggest that parents provide information that is similar to, but less accurate than, the information students themselves can provide about teachers' impacts on student learning.

While researchers typically identify smaller teacher effects on students' reading achievement gains than their math gains (e.g. Gates, 2012), the lack of any significant association across student, parent, and lesson observation measures with teachers' language arts value added is unexpected. While it is possible that none of these measures relate to teachers' effects on student achievement gains in language arts, teachers' value added in this subject area may also be imprecisely estimated. It is possible that either the high volume of co-teaching assignment in the district in language arts or the impact of targeted school-level interventions and supports provided by the district introduced additional noise or bias into the language arts teacher value added estimates. These or other factors may have made it more difficult to reliably assess individual teacher contributions in this subject area.

## **Conclusions and Discussion**

This study provides a variety of new evidence regarding the characteristics and potential contributions of parent ratings of teacher quality for use in teacher evaluation. First and foremost, I find that parent perspectives provide information about a distinct and stable attribute of teachers' performance. The perception that one classroom of students' parents has towards a teacher is highly consistent with that of a future classroom's parents in the following school year. This stability is apparent even when considering relative teacher ratings of parents within the same school environment. In addition to capturing a consistent quality about teachers, the distribution of parent ratings allows for meaningful differentiation of teachers' relative

performance, both within and across schools. These findings align with and extend existing research on parent evaluation of teachers. They strongly suggest that educators and policymakers can gain meaningful information about teacher quality – as understood by parents – by leveraging survey measures of the type examined in this study. However, parent responses also appear fairly one-dimensional. If this finding holds true across other contexts and surveys, then teachers may not gain particularly nuanced formative feedback from parent surveys.

In addition to clarifying those foundational characteristics of parent surveys measures used for teacher evaluation, this study also offers new insight into the type of information that parents may provide about teachers by assessing their alignment with other evaluation measures. Consistent with prior research (Peterson et. al, 2003), I find that parent ratings of teachers are strongly associated with student ratings, likely because parents draw from their child's reports to inform their own opinion. I also find that parent ratings can predict teachers' math value added performance ratings, and are modestly correlated with some areas of formal lesson observation ratings. These results suggest that educators and policymakers can expect parent perspectives to be reasonably aligned with other valued measures of teacher quality, while also providing additional, distinct information.

In light of these results, parent perspectives may be useful as a partial proxy in cases where other evaluation measures are unavailable. Given their strong alignment with student ratings, they may be most valuable as proxies for student perspectives, particularly in elementary grades or other contexts where student evaluations are deemed infeasible to collect or use in high stakes evaluations. They appear to also provide some information about students' achievement gains in teachers' classrooms, at least in math. However, it is notable that using student surveys and parent surveys in tandem to predict teachers' value added performance explains little additional

variance compared to using student surveys alone. This, combined with the relatively modest association between parent perspectives and teacher value added performance, suggest that school districts that rely on non-value added measures as substitute for assessing teachers' impacts on student achievement may be better served by leveraging student perspectives and lesson observation data about teachers, rather than parent perspectives as a proxy of choice.

The areas of alignment between parent evaluations of teachers and specific elements from either student or lesson observation measures provide additional clarity regarding the particular teacher characteristics and behaviors that parents are able to report on. While parent ratings align with a range of different student survey questions, they are most strongly and consistently aligned in their assessment of how caring and encouraging the teacher is towards the student. In contrast, parents appear less likely to pick up on students' perceptions about the teachers' overall management of their classrooms' behavior or the efficient use of time during instruction. Overall, these results suggest that both parents and students are able to identify teachers' interpersonal characteristics in their separate interactions with teachers.

In the area of lesson observations, I observe that parent measures are primarily aligned with teachers' success in generating a positive classroom environment, and in actively engaging student participation. This is consistent with the idea that parents are best able to assess aspects of teacher quality that hinge upon strong interpersonal skills. When controlling for teachers' ability to engage students and create a positive classroom environment, no other dimensions of instructional practice are meaningfully associated with parent ratings. The lack of alignment with other lesson observation dimensions, along with the generally weak agreement between prior-year parent ratings and future teacher observation ratings, suggest that administrators may gain only a limited view of teachers' instructional skills from parent

perspectives, and one that does not effectively predict their instructional practice across different classrooms and school years.

When interpreting these findings, it is also worthwhile to consider the particular characteristics of the study sample. The district in this study is a charter school district that families attend by choice, and it has made substantial efforts to establish frequent parent-teacher interactions. Both of these characteristics of the sample may inform our interpretation of the results. First, although the positive skew that I observe in parent ratings of teachers is consistent with prior research, it is possible that the charter-selecting parents in this sample may be more positively disposed towards their school's teachers than is typical in some other urban public schools. Second, the degree to which parents and teachers interact may be an important mediating factor that influences the accuracy and variety of information gained from assessing parent perspectives about teacher quality. The emphasis that this district places on parent-teacher relations is likely common to many public schools. However, schools that do not emphasize teacher-parent relations may see different results from parent surveys. In particular, such schools may include more parents who report little or no interaction with teachers, and who may be less informed about teacher quality as a result. Thus, my results may be representative of what relatively informed and engaged parents are able to report about their children's teachers.

The findings from this study can help to guide educators and policymakers who are considering whether and how much they may rely on parent measures to assess teacher quality. However, additional research is needed to fully gauge both the unique information that parents can provide, and the potential benefits of leveraging their input as part of comprehensive evaluation systems. First, while this study includes a range of parent survey questions and finds that parents held a fairly one-dimensional parent view of teachers, it is possible that the use of

more varied or specific questions will identify additional distinct aspects of teaching that parents can provide feedback about. For instance, parent survey measures may be better suited to identifying explicit teacher practices related to building relationships with parents, rather than more generic assessments of teacher quality. Second, research that assesses the characteristics of parent ratings across more heterogeneous district contexts would be better able to identify any important differences or biases in parent ratings that may correspond to teachers' or families' differing demographic and socioeconomic characteristics. This type of investigation can help to assess the fairness and reliability of parent measures for use in teacher evaluation.

Finally, and perhaps most importantly, we know little at this point about whether using parent perspectives in teacher evaluations will create desirable incentives or provide useful feedback to teachers or schools that may improve their effectiveness. More research is needed that examines whether and how the use of parent perspectives in teacher evaluation systems can positively influence teacher-parent interactions and ultimately impact valued student outcomes. Future investigations can also clarify the extent to which teachers, parents, and other key stakeholders value the inclusion of parent perspectives in evaluation systems, both for high stakes and low stakes purposes. This study provides foundational information about the characteristics of parent ratings that can inform their use and adoption in schools. Future work can address the critical questions of implementation that will determine their usefulness in practice.

## Notes

<sup>1</sup>Interested parties may contact me with clarifying questions regarding additional district context pertinent to the study.

<sup>2</sup>Because the district did not track student assignment to individual classrooms within teachers, classroom average characteristics are instead calculated within the sample of all peer students taught by the same teacher, in the same subject and grade level.

<sup>3</sup>Note that because I use teacher value added measures as a dependent variable in my analysis, I do not apply any Empirical Bayes shrinkage to my teacher value added estimates. Results using measures that include an Empirical Bayes adjustment are very similar to those presented here.

<sup>4</sup>Teacher characteristics, including individuals' age, gender, years of experience, and race were also available for most of the teachers in my sample. In preliminary analysis, I found that my findings were robust to – and virtually unchanged by – the inclusion of controls for teacher characteristics.

<sup>5</sup>The first two Eigen values in this exploratory principal components factor analysis were 9.722 and 0.291.

<sup>6</sup>An alternative specification testing within-school parent rating measures also indicates jointly significant between-teacher effects.

<sup>7</sup>A Steiger's Z-test indicates that the simple correlation between survey questions (#'s 2 and 7) and parent ratings of teachers is significantly different from the same correlation between parent ratings and item #'s 3, 10, 11, or 12 on the student survey.

## Figures

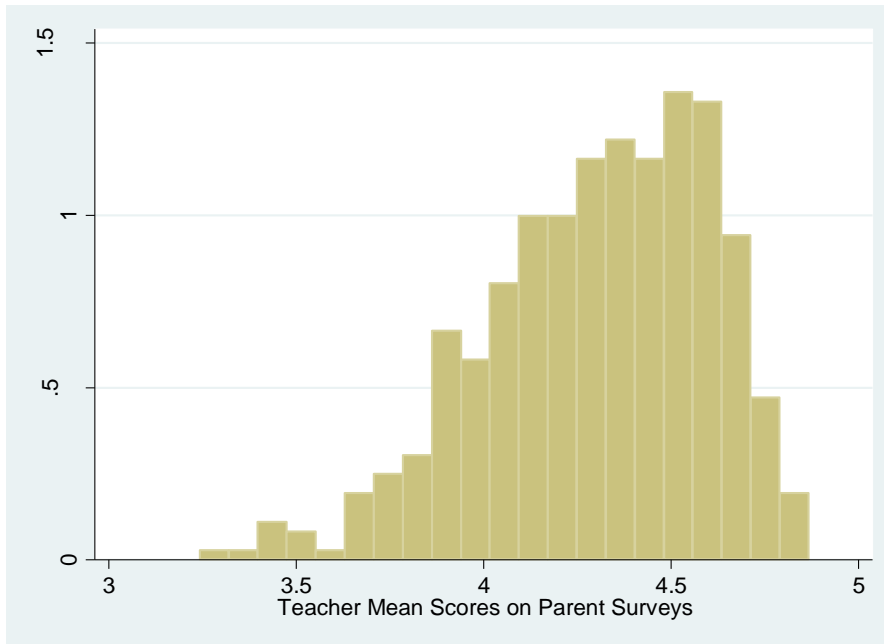


FIGURE 1. *Distribution of the Average of Individual Teachers' Overall Parent Survey Ratings*

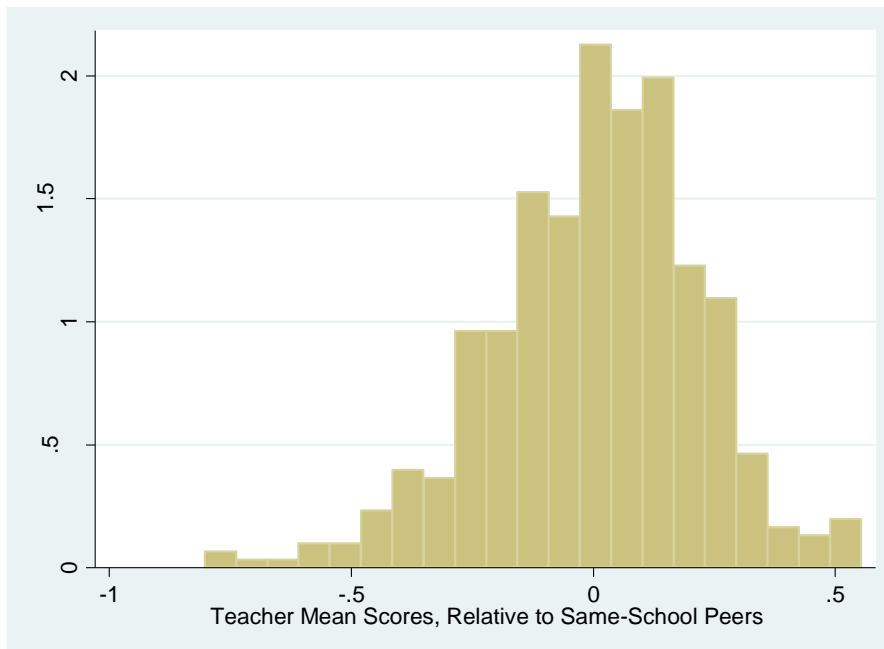


FIGURE 2. *Distribution of the Average of Individual Teachers' Within-School Parent Survey Ratings*

## References

- Advocates for Children of New York, June 2012. "Essential Voices: Including Student and Parent Input in Teacher Evaluation." Report retrieved March 14, 2013, from [http://www.advocatesforchildren.org/policy\\_and\\_initiatives/policy\\_reports](http://www.advocatesforchildren.org/policy_and_initiatives/policy_reports)
- Bill & Melinda Gates Foundation. 2012. "Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains." MET Project Research Paper. Retrieved March 14, 2013, from <http://www.metproject.org/reports.php>
- Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J (2010, July 20). "Teacher layoffs: An empirical illustration of seniority v. measures of effectiveness." CALDER working paper.
- Bursztyn, L., & Coffman, L.C. (2010). *The schooling decision: Family preferences, intergenerational conflict, and moral hazard in the Brazilian Favelas*. Harvard University, Cambridge, MA.
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Domina, T. (2005). Leveling the home advantage: Assessing the effectiveness of parental involvement in elementary school. *Sociology of Education*, 78, 233-249.
- Ehlert, M., Kodel, C., Parsons, E., and Podgursky, M. (2012). Selecting Growth Measures for School and Teacher Evaluations.
- Epstein, J. L. (1985). A question of merit: Principals' and parents' evaluations of teachers. *Educational Researcher*, 14(7), 3-8.
- Fan, W. and C.M. Williams. (2010). The effects of parental involvement on students' academic self-efficacy, engagement and intrinsic motivation. *Educational Psychology*, 30(1), 53-74.



- Goldhaber, D., and Theobald, R. (2011) "Managing the teacher workforce in austere times: The implications of teacher layoffs." Seattle, WA: University of Washington Center for Education Data and Research. Retrieved on March 14, 2013 from <http://www.cedr.us/publications.html>
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? *Child Development, 76*, 949–967.
- Kane, T. J., Taylor E. S., Tyler J. H., and Wooten, A. L. (2011). Identifying effective classroom practices using student achievement data, *Journal of Human Resources, 587-613*.
- Kraft MA, Dougherty SM. (Forthcoming). The Effect of Teacher-Family Communication on Student Engagement: Evidence from a Randomized Field Experiment. *Journal of Research on Educational Effectiveness*.
- Ostrander, L. P. (1995). Multiple judges of teacher effectiveness: Comparing teacher self-assessments with the perceptions of principals, students, and parents. [Doctoral Dissertation] VA: University of Virginia.
- Powell-Smith, K.A., Stoner, G., Shinn, M.R., & Good, R.H. (2000). Parent tutoring in reading using literature and curriculum materials: Impact on student reading achievement. *School Psychology Review, 29*, 5–27.
- Peterson, K. D. (1987). Teacher evaluation with multiple and variable lines of evidence. *American Educational Research Journal, 24*, 311–317.
- Peterson, K. D. (1989). Parent surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education, 2(3)*, 239–249.

- Peterson, K., Wahlquist, C., Brown, J., & Mukhopadhyay, S. (2003). Parent surveys for teacher evaluation. *Journal of Personnel Evaluation in Education*, 17(4), 317-330.
- Rockoff, J. E., Staiger D. O., Kane, T. K., and Taylor, E. S. (2011). Information and employee evaluation: evidence from a randomized intervention in public schools. *American Economic Review*.
- Rothstein, R., Ladd, H. F., Ravitch, D., Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Linn, R. L., Shavelson, R. J., and Shepard, L. A. (2010). *Problems with the Use of Student Test Scores to Evaluate Teachers*. Economic Policy Institute. Retrieved March 14, 2013, from <http://www.epi.org/publication/bp278>
- Sheldon, S.B., & Epstein, J.L. (2005). Involvement counts: Family and community partnerships and mathematics achievement. *Journal of Educational Research*, 98, 196–206.
- Simon, B.S. (2001). Family involvement in high school: Predictors and effects. *NASSP Bulletin*, 85, 8–19.
- Sirvani, H. (2007a). Effects of teacher communication on parents' attitudes and their children's behaviors at schools. *Education*, 128(1), 34-47.
- Sirvani, H. (2007b). The effect of teacher communication with parents on students' mathematics achievement. *American Secondary Education*, 36(1), 31-46.
- Taylor, E.S., & Tyler, J.H. (2011). "The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers." National Bureau of Economic Research working paper no. 16877.

## Tables

TABLE 1

*Parent Survey Sample Statistics and Descriptive Statistics for SY 2011-12 Teacher Ratings (All responses on a 5-point likert scale)*

	N of Teachers	Parent Response Rate	Mean # of Responses per Teacher
Parent survey statistics, district-wide implementation SY 2011-12	466	83%	25.21
Parent survey statistics, pilot implementation SY 2010-11	133	69%	20.30
SY 2011-12 Survey Items and Statistics at the Teacher Level			
	Teacher Mean Score	Teacher Std. Dev.	Teacher-level Correlation with Question #11
1. How comfortable do you feel talking with this teacher about your child?	4.29	0.297	0.830
2. How well informed are you about the progress your child is making in this teacher's class?	4.10	0.381	0.822
3. How clearly has this teacher communicated the learning goals he/she has for your child in this class?	4.14	0.362	0.832
4. How professional is this teacher in your interactions with him or her?	4.37	0.249	0.818
5. How often does this teacher return your phone calls or emails within 24 hours (not including weekends)?	4.25	0.343	0.765
6. How respectful is this teacher towards your child?	4.40	0.262	0.820
7. This teacher notifies me of the positive work my child has done in his/her class.	4.10	0.426	0.845
8. How helpful has this teacher been in improving your child's academic performance?	4.31	0.305	0.854
9. This teacher has a positive impact on my child's character.	4.28	0.314	0.861
10. This teacher motivates my child to work hard and to do his/her best work for this class.	4.57	0.271	0.863
11. Overall, how would you grade this teacher (A+, A, B, C, D)?	4.50	0.298	1.00
	Cronbach's Alpha	Proportion of Variance Accounted For	
Descriptive statistics for a single factor across teacher-level individual item scores	0.955	0.884	

Note: Surveys administered between March 1st and April 30th of each school year. All survey questions other than question #2 were asked in both school year 2011-12 and in the pilot survey in school year 2010-11. Correlations with question 11 are shown for a sample of parent surveys with no missing item-level responses, in order to avoid any potential bias resulting from item imputation.

TABLE 2

*Between-and-Within Teacher Variance Decomposition of Student, Parent, and Value Added Measures*

	Within-teacher Variance Proportion	Between-teacher Variance Proportion
Parent Survey Questions, Parent-level Average Score	0.854	0.146***
Student Survey Questions, Student-level Average Score	0.787	0.213***
Student Adjusted Achievement Gains in Mathematics	0.829	0.171***
Student Adjusted Achievement Gains in Language Arts	0.901	0.099***

TABLE 3

*Predicting same-year and future-year teacher evaluation measure scores with teachers' parent survey scores*

	School Year 2011-12 Teacher Evaluation Measures			Math Value Added, Gr. 4-8	Language Arts Value Added, Gr. 4-8
	Student Perspectives	Lesson Observations	Parent Perspectives		
Same-Year, SY 2011-12 Parent Perspectives	0.644*** (0.062)	0.333*** (0.059)		0.365** (0.132)	-0.101 (0.156)
Controls for School Type	Yes	Yes	n/a	Yes	Yes
Adjusted R-Squared	0.289	0.078		0.078	<0
N	292	443		67	86
Prior Year, SY 2010-11 Parent Perspectives	0.195 (0.170)	0.104 (0.146)	0.569*** (0.100)	(insufficient sample)	
Controls for School Type	Yes	Yes	Yes		
Adjusted R-Squared	0.095	<0	0.508		
N of teachers	74	96	86		

Note: Teacher samples vary by measure: value added data are available in grades 4-8; student survey data in grades 3-8; lesson observations and parent surveys in grades K-12. Parent and student ratings are represented as an average overall survey score for each teacher, while lesson observations scores are a standardized principal component factor of all observation dimensions. Additional controls include indicator variables for school type (i.e. Elementary, Middle, or High School). ~p<.1 \*p < .05, \*\*p < .01, \*\*\*p < .001.

TABLE 4

*Predicting same-year and future-year within-school teacher evaluation measure scores with teachers' within-school parent survey scores*

	School Year 2011-12 Teacher Evaluation Within-School Measures			Math Value Added, Gr. 4-8	Language Arts Value Added, Gr. 4-8
	Student Perspectives	Lesson Observations	Parent Perspectives		
Same-Year, SY 2011-12 Within-School Parent Perspectives	0.485*** (0.045)	0.232*** (0.045)	n/a	0.150 (0.092)	-0.016 (0.094)
Adjusted R-Squared	0.283	0.054		0.025	<0
N	293	444		67	86
Prior Year, SY 2010-11 Within-School Parent Perspectives	0.353* (0.143)	0.109 (0.124)	0.710*** (0.119)	(insufficient sample)	
Adjusted R-Squared	0.065	<0	0.289		
N of teachers	74	96	86		

Note: Teacher samples vary by measure: value added data are available in grades 4-8; student survey data in grades 3-8; lesson observations and parent surveys in grades K-12. Parent and student ratings are represented as an average overall survey score for each teacher, while lesson observations scores are a standardized principal component factor of all observation dimensions. ~p<.1 \*p < .05, \*\*p < .01, \*\*\*p < .001.

TABLE 5

*Models predicting individual student survey question ratings of teachers with parent perspectives from the same or a prior school year*

Dependent variables: Teachers' SY 2011-12 Standardized Ratings for Individual Questions from Student Surveys	Same-Year Parent Perspectives			Prior-Year Parent Perspectives		
	Coefficient	Standard Error	Adjusted R-squared	Coefficient	Standard Error	Adjusted R-squared
This teacher is a strong role model for our school. ( <i>item</i> #13)	0.635***	(0.061)	0.312	0.244	(0.166)	0.080
This teacher makes me feel that he/she really cares about me. (2)	0.610***	(0.063)	0.272	0.324*	(0.159)	0.140
The teacher in this class encourages me to do my best. (7)	0.598***	(0.064)	0.246	0.424*	(0.173)	0.172
In this teacher's class we learn a lot almost every day. (6)	0.583***	(0.063)	0.239	0.165	(0.160)	0.043
I look forward to coming to this teacher's class. (1)	0.581***	(0.061)	0.298	0.181	(0.154)	0.109
How often does this teacher recognize your good work? (5)	0.577***	(0.064)	0.224	0.167	(0.169)	0.108
When you don't understand something, does this teacher work with you until you do understand it? (9)	0.577***	(0.065)	0.223	0.222	(0.170)	0.075
How clearly does this teacher explain things? (8)	0.574***	(0.065)	0.220	0.136	(0.170)	0.048
This teacher believes that each and every student in my class will succeed. (4)	0.566***	(0.066)	0.206	0.22	(0.186)	0.061
How well respected do you feel in this class? (3)	0.485***	(0.066)	0.174	0.136	(0.164)	0.062
I know how I am doing in this teacher's class. (10)	0.439***	(0.068)	0.121	0.001	(0.163)	-0.027
How often do your classmates behave the way this teacher wants them to behave? (11)	0.406***	(0.065)	0.228	-0.06	(0.140)	0.192
Our class stays busy and doesn't waste time. (12)	0.398***	(0.063)	0.250	0.023	(0.153)	0.188
N of teachers		293			74	

Note: Parent perspectives represented as a teacher average survey score across all items. Item numbers reflect the order of questions in the survey instrument. Sample includes teachers in grades 3-12, where student survey data is available. All models include indicator variable controls for school type (i.e. Elementary, Middle, High School). ~p<.1 \*p < .05, \*\*p < .01, \*\*\*p < .001.

TABLE 6

*Models predicting lesson observation ratings in specific areas with parent perspectives from the same or a prior school year*

Dependent variables: Teachers' SY 2011-12 Standardized Scores on Individual Lesson Observation Dimensions	Same-Year Parent Perspectives			Prior-Year Parent Perspectives		
	Coefficient	Standard Error	Adjusted R-squared	Coefficient	Standard Error	Adjusted R-squared
Positive classroom climate and teacher messaging (D3)	0.391***	(0.061)	0.085	0.186	(0.135)	<0
Student engagement strategies and results (D2)	0.377***	(0.061)	0.084	0.159	(0.143)	<0
Efficient pacing and seamless classroom routines (D2)	0.300***	(0.061)	0.068	0.024	(0.135)	0.013
Student character development and results (D3)	0.291***	(0.058)	0.105	0.085	(0.148)	<0
Differentiated instruction (D4)	0.263***	(0.061)	0.048	0.158	(0.155)	<0
Frequent and efficient review of prior lesson material (D4)	0.254***	(0.060)	0.071	0.182	(0.131)	0.027
Evidence of student mastery of lesson goals (D4)	0.248***	(0.062)	0.033	-0.004	(0.140)	<0
Effective checks for student understanding (D1)	0.242***	(0.061)	0.028	0.026	(0.141)	<0
Effective and efficient explanations of content (D2)	0.241***	(0.061)	0.035	0.069	(0.149)	<0
Evidence of effective lesson planning (D2)	0.237***	(0.061)	0.036	0.049	(0.136)	0.018
High behavioral expectations for students (D3)	0.224***	(0.060)	0.062	0.048	(0.136)	<0
High standards for student responses (D1)	0.221***	(0.060)	0.039	0.078	(0.145)	<0
Quality student independent work (D2)	0.221***	(0.061)	0.028	0.038	(0.144)	<0
Effective use of daily assessments (D1)	0.203***	(0.061)	0.020	0.102	(0.139)	0.006
Planned, rigorous questioning (D1)	0.176**	(0.060)	0.023	0.103	(0.146)	0.009
Clarity, rigor, and standards-alignment of lesson goals (D1)	0.139*	(0.062)	0.022	-0.020	(0.115)	0.015
N of teachers		444			96	

Note: Parent perspectives represented as a teacher average survey score across all items. Sample of teachers across grades K-12. All models include indicator variable controls for school type (i.e. Elementary, Middle, High School). Lesson observation dimension descriptions paraphrased for brevity, and categorized into the observation rubric's four conceptual domains: D1) academic rigor and clarity, D2) design and delivery of core lesson elements, D3) development of a positive classroom culture, and D4) ensuring student learning.

~p<.1 \*p < .05, \*\*p < .01, \*\*\*p < .001.

TABLE 7

*Models predicting lesson observation ratings in specific areas with parent perspectives, with controls for key lesson dimensions*

Dependent variables: Teachers' SY 2011-12 Standardized Scores on Individual Lesson Observation Dimensions	Same-Year Parent Perspectives		
	Coefficient	Standard Error	Controls for positive classroom environment and student engagement
Efficient pacing and seamless classroom routines (D2)	0.011	(0.046)	Yes
Student character development and results (D3)	0.027	(0.046)	Yes
Differentiated instruction (D4)	0.066	(0.056)	Yes
Frequent and efficient review of prior lesson material (D4)	0.077	(0.057)	Yes
Evidence of student mastery of lesson goals (D4)	0.035	(0.056)	Yes
Effective checks for student understanding (D1)	-0.006	(0.050)	Yes
Effective and efficient explanations of content (D2)	-0.008	(0.051)	Yes
Evidence of effective lesson planning (D2)	0.007	(0.054)	Yes
High behavioral/social expectations for students (D3)	-0.079~	(0.042)	Yes
High standards for student responses (D1)	-0.028	(0.050)	Yes
Quality student independent work (D2)	0.007	(0.055)	Yes
Effective use of daily assessments (D1)	0.031	(0.058)	Yes
Planned, rigorous questioning (D1)	-0.066	(0.051)	Yes
Clarity, rigor, and standards-alignment of lesson goals (D1)	-0.046	(0.059)	Yes
N of teachers		444	

Note: Parent perspectives represented as a teacher average survey score across all items. Sample of teachers across grades K-12. All models include indicator variable controls for school type (i.e. Elementary, Middle, High School). Lesson observation dimension descriptions paraphrased for brevity, and categorized into the observation rubric's four conceptual domains: D1) academic rigor and clarity, D2) design and delivery of core lesson elements, D3) development of a positive classroom culture, and D4) ensuring student learning. ~ $p < .1$  \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .



TABLE 8

*Comparing Adjusted R-squared in Models Including Parent Survey Ratings with Other Measures to Predict Same-year Teacher Value Added*

Alternate Teacher Evaluation Measures	SY 2011-12 Math Value Added		SY 2011-12 Language Arts Value Added	
	Base Prediction Adjusted R-squared	Including Parent Perspectives	Base Prediction Adjusted R-squared	Including Parent Perspectives
Parent survey (overall score)	0.060	n/a	<0	n/a
Parent survey (all individual questions)	0.065	n/a	<0	n/a
Student survey (overall score)	0.116	0.104	<0	<0
Student survey (all individual questions)	0.208	0.230	0.006	<0
Lesson observations (single factor)	0.109	0.148	<0	<0
Lesson observations (all individual dimensions)	0.143	0.133	<0	<0
Lesson observations and student surveys (overall)	0.175	0.165	<0	<0
Lesson observations and student surveys (itemized)	0.170	0.154	<0	<0
N of teachers	67	67	86	86

Notes: "Overall" models of student surveys and lesson observations include as a single covariate the overall average teacher score or a single factor score across all survey questions or lesson observation dimensions, respectively. Alternate models include all individual teacher scores on each survey question or dimension as multiple individual covariates.