

**Supporting Continuous Development for Teachers:  
Teacher Evaluation in Six High-Performing, High-Poverty, Urban Schools**

Stefanie K. Reinhorn

Susan Moore Johnson

Nicole S. Simon

Working Paper

The Project on the Next Generation of Teachers

Harvard Graduate School of Education

July 2015

**Abstract**

This qualitative, comparative case analysis of teacher supervision and evaluation is part of a larger, comparative case study, “Developing Human Capital Within Schools,” conducted by the Project on the Next Generation of Teachers. We conducted 142 semi-structured interviews with teachers and administrators in six high-poverty schools (three charter and three district), all of which had achieved the highest level in the state’s accountability rating system. Here, we analyze how teachers and administrators described and assessed teacher supervision and evaluation. Although teacher evaluation is often viewed as a marginal activity, intended primarily for removing failing teachers (Thomas, Wingert, Conant, & Register, 2010), in these schools it was consistently described as being central to teachers’ support and development. Teachers reported being observed frequently and given regular feedback. They appreciated their schools’ robust, ongoing supervision and evaluation processes and valued the ways it was integrated with other opportunities for professional learning.

This working paper is part of a larger study conducted by researchers at the Project on the Next Generation of Teachers. We are indebted to the Spencer Foundation and to the Harvard Graduate School of Education for funding this project, although all views presented here are our own. We appreciate the comments and recommendations of Judith Warren Little at the Annual Meeting of the American Educational Research Association, 2015. David Cohen, Andrés Alonso and Megin Charner-Laird also provided insightful feedback on earlier drafts of this paper. We are deeply grateful to the administrators and teachers who participated in this study.

## Supporting Continuous Development for Teachers: Teacher Evaluation

In recent years, the US federal government crafted several policies (Race to the Top, Teacher Incentive Fund Grants and No Child Left Behind waiver requirements) to motivate states, districts and schools to remake teacher evaluation processes. Most states responded to incentives and sanctions embedded in these policies with new legislation calling for massive changes in teacher evaluation systems (NCTQ, 2013). States, districts and schools are very early in the process of implementing these policies and many questions remain about their impact on teaching and teacher quality. More specifically, little is known about how teachers experience these reforms in their daily work.

In theory, teacher evaluation can improve teaching quality by two means: identifying and dismissing underperforming teachers and developing all teachers' professional practice. On paper, new teacher evaluation policies appear to attend to both the accountability and developmental aims of the process. Recent quantitative studies support the potential payoff of investing in teachers' learning over the course of the "career continuum" (Feiman-Nemser, 2001). For example, researchers find that teachers continue to improve for at least ten years into their career, especially when their schools provide supportive work environments (Kraft & Papay, 2014; Ladd, 2011). Currently, formal evaluation is one of the most widely adopted policies intended to improve teachers' effectiveness. Yet, we know very little about how these policies are being implemented locally and researchers have only begun to understand how evaluation can contribute to teachers' professional growth.

The goal of improving teaching quality is of paramount importance in urban schools, where low-income and minority students depend more on their teachers than their wealthier counterparts (Chetty, Friedman, & Rockoff, 2013; Downey, Hippel, & Hughes, 2008), yet are more likely to be taught by the least experienced, least effective teachers (Borman & Dowling,

2008; Carroll et al., 2000; Clotfelter et al., 2007; Hanushek et al., 2004; Ingersoll, 2004). Rivkin, Hanushek, & Kain (2005) report that having a “high-quality” elementary school teacher “can substantially off-set disadvantages associated with low socio-economic background” (p. 419). Thus, students in high-poverty schools stand to gain if evaluation improves teaching quality.

In this analysis, we sought to understand how teacher evaluation worked in six high-poverty schools (three charter and three district) that were demonstrating success with student achievement according to state accountability measures. All six schools, located in one Massachusetts city, were adhering to the state’s new teacher evaluation regulations, with most implementing its model evaluation system. Specifically, we examined how teachers and administrators described and assessed their experiences with teacher evaluation at their school.

In this sample, most school leaders had the authority to hire and dismiss or transfer teachers and so could exercise considerable control over who was teaching in their school. Therefore, they did not rely primarily on the evaluation process as a tool to “de-select” teachers (Hanushek, 2009) as one might expect if there were no flexibility in staffing policies. Instead, they were committed to using supervision and evaluation to develop their teachers. The large majority of teachers in this sample experienced ongoing, comprehensive supervision, which went well beyond what state and district policies required. Teachers described evaluation as contributing to their strong sense of responsibility for the quality of instructional practice and for their students’ learning.

In what follows, we begin by situating this paper in the current body of literature on teacher evaluation. After describing our research methodology, we present and discuss our findings, considering the particular histories and policy environments of the schools in the

sample. We conclude with a discussion of the implications of this study for future research, policy and practice.

### **Literature Review**

Until recently, teacher evaluation has had a disappointing track record. Several research papers, including the high-profile report, *The Widget Effect* of the New Teacher Project (Weisberg et al., 2009), document the failure of many U.S. school districts' evaluation practices to identify and respond to variations in teacher effectiveness (Donaldson, 2009; McLaughlin & Pfeifer, 1988; Toch & Rothman, 2008; Wechsler et al., 2007; Wise, Darling-Hammond, Tyson-Bernstein, & McLaughlin, 1984). Donaldson (2009) describes U.S. school districts as suffering from the “the Lake Wobegon effect” because almost all teachers receive positive ratings, even though teachers within schools are known to vary widely in their effectiveness (Rivkin et al., 2005; Rockoff, 2004). Researchers and practitioners generally agree with TNTP's (2010) assertion that teacher evaluation in the United States has been largely a “perfunctory compliance exercise that rates all teachers good or great and yields little useful information” (p.1). Furthermore, school districts rarely use evaluations to inform decisions about tenure or dismissal in charter or traditional schools (Donaldson, 2011; Tucker, 1997).

### **New Evaluation Policies**

In recent reforms across the nation, state legislators have significantly changed teacher evaluation policies. The number of states that differentiate teachers into multiple levels of effectiveness increased from 17 in 2011 to 42, plus Washington DC, in 2013(NCTQ, 2013). Almost every state requires classroom observations and almost half call for multiple observations yearly for new, and in some cases, all teachers. In 21 states, evaluators must provide feedback after observing a teacher. Data from 17 states and Washington DC, suggest that, as a result of

new regulations, school systems have made significant progress on differentiating between “poor, fair and great” teaching (Aldeman & Chuong, 2014). In addition, these schools are using higher-quality observation rubrics (Danielson, 2013) and providing teachers with more frequent and more detailed feedback on their practice.

Although many policy analysts see promise in the new regulation (Doyle & Han, 2012; NCTQ, 2013), the success of the policies ultimately depends on how they are implemented at the school level. Kimball and Milanowski (2009) studied implementation of a new standards-based evaluation system in one district and found that, despite a detailed rubric and district-level training of evaluators, there were idiosyncratic differences in how they were used existed across schools. In our prior research analyzing teachers’ experiences with evaluation in six high-poverty schools in one large, urban district, we found marked differences in implementation across schools (Reinhorn & Johnson, 2014). Although the district’s new policy was intended to help teachers improve and most teachers saw potential in the process, teachers in only one school said that they benefited from evaluation. In four schools, principals and teachers described evaluation largely as a set of bureaucratic procedures, and in the sixth, teachers said it was primarily a means for dismissing incompetent teachers. This variation in implementation supports the need for more detailed analyses of teacher evaluation practices within schools.

### **The Impact of Evaluation on Teaching Quality**

Very little research has focused on whether or how new teacher evaluation practices are supporting improvements in teaching quality. In a convincing study that links teacher evaluation and student assessments in Cincinnati, Taylor and Tyler (2012) found that, when experienced peer evaluators and administrators used “multiple, highly-structured classroom observations,” mid-career math teachers’ effectiveness, as measured by their students’ achievement, improved.

Importantly, the gains persisted and were even stronger several years after the evaluation cycle occurred. The authors suggested that the evaluation process provided teachers with detailed feedback and the incentive to improve. Although this research supports the potential of high quality evaluation processes, Taylor and Tyler's study does not explain how evaluation yielded these results. In an effort to understand teachers' views on the impact of evaluation practices, Donaldson (2012) interviewed 95 teachers in one mid-size urban district during their second year of implementing a new evaluation system. Evaluators assessed teachers based on classroom observations and students' growth on academic performance measures. Most teachers reported that the new evaluation process focused their attention more directly on students' test results, but did not lead them to make changes in their instructional practice. Teachers generally believed that evaluation reform was necessary and expressed positive views of the new evaluation system, although they provided mixed reports about the fairness and objectivity of the process. These contrasting findings raise significant questions about what the evaluation process should provide in order for it to help teachers improve their practice.

### **The Role of the Evaluator**

In a study of 121 new teachers in Houston, O'Pry and Schumacher (2012) found that teachers' perceptions were determined less by the evaluation tool itself than by how it was used. The authors surveyed teachers about their views of the district's standards-based performance appraisal system, and then interviewed those whose views of evaluation were most positive or negative. The factor that most consistently influenced teachers' perceptions of the process was the value that they thought their principal placed on the process. Sartain and colleagues (2011) also concluded that the impact of an evaluation system depends greatly on the evaluators'

implementation of the process. These researchers studied a pilot group of principals who were trained to use the *Charlotte Danielson Framework for Teaching* as an instrument for evaluating teachers and found that the ratings based on classroom observations were both valid and reliable. Also, students showed the greatest growth in classes where teachers consistently received the highest evaluation ratings, while students showed the least growth when their teachers received the lowest ratings. These researchers concluded that better evaluation tools could support school leaders in assessing teachers' instructional practices and engaging them in reflective conversations that support their improvement. However, this only occurred when principals had strong knowledge of the instructional framework, well-developed skills for instructional coaching, and they were highly engaged in the process. These findings suggest that school systems need to go beyond training principals to provide valid and reliable ratings and support them in using evaluation to improve teaching quality.

### **Evaluation in Charter Schools**

Because charter schools usually do not bargain collectively and are not obliged to comply with a teachers contract, many assume that they have considerable freedom in how they evaluate teachers. However, there is little research about how evaluation is used in charter schools and the degree to which their context influences the process. Donaldson and Peske (2010) interviewed administrators and teachers in five charter schools from three different Charter Management Organizations [CMOs]. Teachers in these charter schools described the evaluation process as more robust than in their former schools and said that it was grounded in weekly or biweekly observations and coaching sessions by their evaluator. They also described evaluation as being primarily a formative process used to support professional growth with less emphasis on the

summative assessment, which often gets more attention in many new evaluation systems (for example. Taylor, 2015).

Although they were not specifically studying evaluation, findings from Dobbie and Fryer's (2011) recent study of 35 New York City charter schools provide greater insight regarding the use of frequent formative and summative feedback to teachers. They found that five practices explained more than half of the variation in school effectiveness as measured by student test scores on state math and ELA assessments—frequent teacher feedback, data driven instruction, high-dosage tutoring, increased instructional time, and a relentless focus on academic achievement. In their analysis, they controlled for a wide range of educational strategies and philosophies and other explanatory variables. “Schools that gave formal or informal feedback ten or more times per semester have annual math gains that are  $0.075\sigma$  (0.021) higher and annual ELA gains that are  $0.054\sigma$  (0.017) higher than other schools” (p.16). This raises many questions for future research regarding the frequency and type of feedback provided to teachers and the relationship between this feedback and the formal teacher evaluation process.

Donaldson (2013) extended this line of research by analyzing principals' views of evaluation in 30 charter and conventional schools from two northeastern states. Hypothesizing that in their day-to-day experience, principals take a multi-faceted approach to improving teaching quality, Donaldson sought to understand how they viewed teacher evaluation as one of several approaches to human capital development. Over two-thirds of the principals she studied reported that they did not achieve the dual goals of development and accountability due to various factors—lack of time, inadequate observation tools, an unsupportive school culture and the formal nature of observations as limiting factors. Interestingly, in this study charter school principals identified time and culture as a barrier at roughly the same rates as their traditional

school counterparts. Notably, both charter and district principals were far less likely to cite teacher evaluation than they were hiring, as an important means of improving teachers' effectiveness in their school.

Questions remain about whether and how school leaders integrate evaluation with other strategies for developing human capital in schools. How, if at all, does the school's policy context influence implementation of evaluation processes? Donaldson's study suggests that we have much to learn about how schools, charter or traditional, establish a culture that is conducive to productive evaluation practices. To what extent is a positive organizational culture a prerequisite and to what degree can robust evaluation processes contribute to shaping a culture in which teachers appreciate feedback for contributing to continuous improvement?

This research base clearly establishes that creating and implementing an effective standards-based evaluation policy is challenging but has the potential to support schools in developing human capital. Donaldson and Peske (2010) suggest that "policies that focus on assessing and improving teacher practice over time may, in the end, be stronger mechanisms to improve teacher quality than tinkering with summative evaluation" (p.39). Currently many states, districts and schools are implementing new evaluation policies with little guidance from policymakers or researchers about how to achieve the dual goals of supporting improvements in teaching practice and holding teachers accountable for the quality of their instruction.

In this study, we seek to contribute to this line of research by closely examining the teacher evaluation practices at six urban schools, all of which serve students from low-income families and have been judged to be successful according to state accountability ratings. Here, we analyze how teachers and administrators experienced and assessed the evaluation processes at their schools.

## **Methods**

This paper is based on a qualitative, comparative case study embedded in a larger study, “Developing Human Capital Within Schools,” conducted by the Project on the Next Generation of Teachers. The larger study examines how six high-poverty, urban schools—all of which had received the state’s highest accountability rating—attract, develop, and retain teachers. In this paper, we focus on findings regarding teachers’ and administrators’ experiences and assessments of evaluation in their school.

### **Research Questions**

- 1) How do administrators implement teacher evaluation processes in six schools serving high-poverty students and judged to be successful according to state accountability ratings?
- 2) How, if at all, do the policies that regulate the school influence its implementation of teacher evaluation?
- 3) How do teachers describe and assess their experiences with evaluation? How, if at all, do their assessments of the evaluation processes vary from school to school?
- 4) What challenges do teachers and administrators identify in implementing evaluation policies?

### **Sample of Schools**

Our sample selection was guided by four principles. First, we sought a sample that included charter and district schools located in one city in Massachusetts. Second, we looked for schools that served high-poverty populations (where 70% or more of students were eligible for free or reduced-price lunch) and also enrolled high proportions of students of color. Third, we

sought schools that were considered to be high-performing, having achieved the highest rating in the state’s accountability system. Fourth, we sought schools that employed distinctive approaches to human capital development.

To attend to the first three principles, we examined publicly available demographic and student performance data. In seeking out schools that were having “success” with students from low-income families, we used the state’s accountability ratings as a proxy for students’ academic success. The Massachusetts Department of Elementary and Secondary Education [MA DESE] rates every school on a scale from 1 to 5, with 1 denoting the highest performing schools. The formula calculating a school’s rating relies heavily on results from the Massachusetts Comprehensive Assessment System [MCAS], the state’s high stakes standardized test. The formula accounts for growth in student performance and the school’s success in narrowing proficiency gaps among various subgroups of students, using a weighted average from the four most recent years of MCAS data. Although this definition of success is limited because it relies heavily on standardized test scores, it was the best proxy available for identifying schools that were positively affecting students’ academic outcomes. In addition, these ratings were used by the district and state to award and sanction schools and funders, school boards and the popular media monitored them carefully.

To attend to the fourth principle, we consulted our professional networks and considered available information about the approaches to human capital development used by specific schools and, in some cases, charter networks. Based on our initial inquiry, we drew up a sample of six schools—all geographically located within the boundaries of one large urban school district, Walker City School District [WCSD].<sup>1</sup> The sample included three district schools (one

---

<sup>1</sup> All names of schools, districts and individuals are pseudonyms.

traditional and two former turnaround) and three charter schools authorized by the state (including one restart of a chronically underperforming district school). All schools were elementary and/or middle schools, which facilitates cross-site comparisons. To recruit schools, we contacted school officials explaining our study and requesting their participation. All six schools that we approached agreed to participate in the study (For descriptive statistics for sample schools see below). The purposive nature of our sample has allowed us to conduct an in-depth, exploratory study of schools in a particular context. However, because the sample is small and deliberately selected, we cannot generalize our findings beyond our sample.

Supporting Continuous Development for Teachers: Teacher Evaluation

Table 1. *Selected Characteristics of Six Sample Schools\**

School Name	School Type	Grades	Estimated Enrollment	% Low-income students	% African American or Black Students	% Hispanic or Latino Students	% Other Non-white students	% White Students
Dickinson Elementary	Traditional District	PK-5	370	76	4	85	2	9
Fitzgerald Elementary	District - Former Turnaround	PK-5	390	85	70	25	3	2
Hurston K-8	District - Former Turnaround	PK-8	800	75	41	54	4	1
Kincaid Charter Middle	Charter – Restart of District School	6-8	475	88	50	30	10	10
Naylor Charter K-8	Charter	K-8	500	82	70	24	5	1
Rodriguez Charter K1-8	Charter	PK-8	420	72	55	20	7	18

\*Percentages are approximated for confidentiality purposes.

## Data Collection

**Interviews.** Between March and June 2014, we conducted semi-structured interviews with 142 teachers, administrators and other staff in the six schools. Administrator interviews lasted approximately 90 minutes and teacher interviews lasted approximately 45 minutes. At most schools, all members of the research team were present for interviews with the principal and CMO director. In addition all three researchers conducted interviews with teachers at each school, which facilitated cross-site comparisons, improved inter-rater reliability in coding data, and ensured that each research team member knew about each school's structures and culture.

We also purposively constructed our interview sample. At each school, we first interviewed school administrators in order to understand both what processes they used to select, develop, and retain teachers and why they used them. Then, at each school, we recruited a sample of teachers, varying in demographics, teaching experience, preparation, and teaching assignment. We solicited teachers' participation in various ways, including requests by email, flyers in teachers' mailboxes, and principals' recommendations. We also relied on recommendations from the teachers we interviewed about others in their school who might hold views different from their own. Teachers were promised confidentiality and anonymity as participants in the study. In addition, we interviewed other key staff members (e.g. curriculum coaches, program and family coordinators) when it became apparent that their work and views would inform our understanding of teachers' experiences.

In each school, the number of teachers we interviewed varied depending on the school size, the complexity of the organization and the practices used. We interviewed between 33% and 56% of the teachers at each school. (For descriptive statistics about the interviewees, see Appendix A). We used semi-structured protocols (Appendix B) to guide our interviews and

ensure that data would be comparable across sites and across interviewers (Maxwell, 1996). All interviews were recorded and transcribed verbatim.

The interview protocols included several questions that allowed interviewees to discuss systems and processes related to school-based professional learning, including teacher evaluation. We asked administrators what supports their school provided for new and experienced teachers, including supervision and evaluation, formal professional development [PD] and opportunities for teacher collaboration. With administrators and teachers, we used follow-up questions to explore each of these and to identify other sources of support. In order to learn about how teachers assessed these experiences, we asked them to reflect on which of the components of supervision and evaluation that they discussed worked well for them and which did not. By interviewing teachers and school leaders about a range of strategies to develop human capital in their schools, we sought to understand the connections among approaches within a school. In our visits to the schools, we observed a wide range of day-to-day practices and looked for evidence about the school's organizational culture.

Although this design, by definition, limits the depth of information we collected on any one topic, such as teacher evaluation, it situates that practice among other strategies for developing human capital, allowing us to consider relationships among them in our analysis.

**Document collection.** Although interviews were the main source of data for this study, we also gathered many documents that describe state, district and school policies and programs related to recruiting, developing and retaining teachers. The collected documents that informed analysis of interview data in this study included teacher evaluation rubrics, teacher handbooks, school policies, samples of feedback to teachers, and observation frameworks.

## **Data Analysis**

After each interview, we wrote detailed thematic summaries describing the participant and summarizing his or her views. First, we identified themes using *etic* codes drawn from the literature on the elements of developing human capital. Then, we used thematic summaries to analyze each site individually and to conduct cross-site comparisons, identifying common themes and differences. We used this preliminary analysis to supplement the *etic* codes with *emic* codes that emerged from the data (Miles & Huberman, 1994). For example, many interviewees reported working long hours in order to participate in the job-embedded professional learning opportunities and to prepare for their daily teaching responsibilities. Although we had not identified such “demands” as a theme from our review of the literature, preliminary analysis led us to include this code so that we could systematically attend to this issue in our analysis. We then used this preliminary list of codes to review a small sub-set of the transcripts, individually and together, in order to calibrate our understanding and use of the codes, as well as to refine the list and definitions. We repeated this process twice in order to finalize the list of codes and to improve inter-rater reliability. We then thematically coded each transcribed interview using the software, Dedoose (For a list of codes see Appendix C).

After coding all interviews, we engaged in an iterative analytic process, using data-analytic matrices (Miles & Huberman, 1994) to address our research questions about teacher evaluation. We relied on Dedoose’s function that allowed us to sort data by codes and by particular characteristics of interviewees to investigate our research questions. We analyzed the data for each school separately, completing a data analytic matrix (Miles & Huberman, 1994) which included the components of teacher evaluation at each school. We then reviewed school documents, such as teacher evaluation rubrics and school handbooks, to supplement this information. After establishing a clear understanding of the different elements of evaluation at

each school, we created a cross-school matrix and considered similarities and differences in teachers' and administrators' perspectives within and across schools. Finally, we wrote an analytic memo comparing teacher evaluation at the six schools in order to discern patterns about how teachers and administrators experienced and assessed these practices.

We used several strategies to address risks to validity. Throughout the process, we returned to the data to review our coding and check our emerging conclusions, seeking rival explanations or disconfirming data (Miles & Huberman, 1994). We also shared analytic memos, outlines, and drafts among ourselves and with colleagues who were familiar with this line of research but not involved in this research project so that they might offer alternative interpretations of the data.

### **Findings**

In US schools today, evaluation is often viewed as a marginal activity, intended primarily for removing underperforming teachers (Thomas, Wingert, Conant, & Register, 2010). The school leaders in this sample had a more complex view of evaluation, seeing it as central to building human capital within their organization. Teachers across the sample experienced evaluation as a robust, ongoing improvement process. Teachers and administrators in all schools also took the evaluation process seriously as a means of assessing teachers' effectiveness and determining future employment. However, teachers' sense of accountability seemed more driven by the day-to-day activities intended to support continuous improvement, such as frequent observational feedback, than by the formal elements of the evaluation process, including their rating on the rubric.

### **Contextual Factors Influenced Schools' Approach to Evaluation**

Two conditions enabled these schools, to varying degrees, to focus their evaluation efforts on developing teachers rather than on dismissing them. First, the schools in this sample had considerable staffing flexibility. All had substantial autonomy in hiring teachers and some had the added assurance of being able to transfer or dismiss teachers who did not do well. At the time of this study, none of the administrators in this sample would have described their school as having a significant proportion of underperforming teachers whom they were trying to “evaluate out.” Second, several of the schools had the ability to organize their administrative resources, so that principals and other supervisors could focus on what they saw as the primary function of evaluation—developing teachers.

Before examining the evaluation processes in these schools and the teachers’ perspectives on them, it is important to describe the range of schools in the sample and their particular histories and policy contexts, both of which influenced their approach to evaluation. Across the sample, school-based administrators who were responsible for supervising and evaluating teachers had different titles and roles, including principal, assistant principal, head of school, or dean of instruction, curriculum, or professional learning. For this paper, we refer to them all as supervisors or evaluators.

**The schools and their contexts.** *Naylor Charter School* and *Rodriguez Charter School* were well-established state-authorized charter schools that opened their doors ten and twenty years earlier, respectively, to serve elementary and middle school students from the area. Both were freestanding entities at their inception, although Naylor eventually became one of three schools in the Naylor Charter Network. Administrators at these schools had the authority to recruit, and hire teachers from the time their school opened.

*Kincaid Charter School* had been selected by the district and authorized by the state to “restart” a failing WCSD middle school in 2011, three years prior to this study. School officials promised, and Kincaid delivered, significant and rapid gains in student test scores. They successfully recruited more than 80% of the students who had been enrolled by the traditional school before the restart, more students than typically returned to the school each year under the prior administration. In accordance with the requirements for restarting a school, Kincaid had invited current teachers from the school to apply for positions in the new charter school; however, very few applied and no one who did was asked to return. When Kincaid opened, all teachers were new to the school.

*Dickinson Elementary School* was a century–old district school that served a largely immigrant population, most from the school’s surrounding community. In recent years, Dickinson had been recognized by the state and district for growth in students’ MCAS scores. In 2013-14, Dickinson began implementing WCSD’s new hiring system, which granted all district schools significant autonomy in selecting teachers from the district’s pool of candidates. Notably, Dickinson had a history and reputation of having very little teacher turnover and, therefore, experienced few opportunities to hire new teachers. In fact, when we conducted our study, over half of their teachers had worked at the school for more than twenty years. Many teachers in our sample talked about having waited years to apply for an opening at Dickinson before they received a job offer.

*Hurston K-8 School* and *Fitzgerald Elementary School*, also part of WCSD, had histories that differed substantially from Dickinson’s. The state had placed both in turnaround status as chronically underperforming schools four years prior to the study. At the beginning of the “turnaround” process, they were required by the state to replace at least 50% of the existing

faculty and the principal. After a process of reviewing the performance of current teachers, Hurston's new principal replaced about 80% of the school's teachers and Fitzgerald's replaced about 65%. In subsequent years, both schools demonstrated substantial growth on state standardized tests, allowing them to exit turnaround status, each in three years. Although school leaders and teachers at both schools were proud of this accomplishment, they were also forthcoming about the need for continuing to improve.

After exiting turnaround status, both Hurston and Fitzgerald (with support from their teachers) requested and received significant waivers from district policies in order to continue their reform efforts. Both schools were granted autonomy in hiring along with flexibility in many other areas ranging from budget to curriculum. Although Fitzgerald and Hurston K-8 were part of WCSD, both continued to have significant school-based control of their organizations.

*Selective hiring practices.* The autonomies granted by state and local policies allowed these school leaders to be very selective in hiring teachers. Administrators across schools similarly described the ideal candidates; one said, they were "...willing to constantly reassess, reinvent, and really be creative." They sought teachers who had the will to improve and were ready to, as one principal said, "take feedback and grow." A network administrator at Naylor Charter explained, "If we've hired well, they are people who are constantly looking to improve. If they are looking to coast, then it was probably just our fault in hiring." Four of the six schools required applicants to teach a demonstration lesson and then debrief it with an administrator or instructional coach, a process that helped the school identify teachers who were interested in improving and able to respond productively to critical feedback. Principal Forte at Fitzgerald explained,

What we really looked for was, how did they debrief with us? The reflection piece was the most important. I think it was more important even than the lesson. Were they able to

say what went well or didn't go well? Did they blame it on the kids? ...If they can't take any hard feedback, that's not going to go so well.

All of the school leaders viewed hiring as the first step in a developmental trajectory of instructional improvement for educators joining their organization.

**Ability to dismiss or transfer teachers.** Although school leaders exercised significant control in selecting the teachers in their schools, not all hires worked out. Teachers in four of these schools were not guaranteed a position in their school beyond the current year. At Naylor Charter and Rodriguez Charter schools, teachers could be dismissed at any time. Because Kincaid was an in-district charter, the WCSD teachers union represented the school's teachers. Kincaid had been granted exceptions from the WCSD teachers contract, including the ability to release teachers from their faculty. If a teacher had achieved tenure after three years, but was then released into the district's transfer pool by Kincaid Charter, that teacher would be guaranteed a job at another school in WCSD. Although Hurston K-8 was a WCSD school, it had a waiver from district policies allowing the principal to involuntarily transfer teachers out of his school, placing them back in the district pool to be reassigned to another WCSD school. In the other two schools, Fitzgerald and Dickinson, the principal could not unilaterally transfer teachers. If these principals believed a teacher was not a good match for their school, they had no formal recourse except, where applicable, to use the WCSD evaluation process to officially dismiss a failing teacher. Therefore, in four out of six schools, administrators had the authority to release or transfer teachers.

Although the relevant policies differed from school to school, all teachers across the sample seemed to believe that teachers who did not perform up to expectations would not remain at their school unless they improved. When a teacher was not a good match for the school or not able to meet its high expectations, he or she was counseled out or not offered a subsequent

contract. However, these administrators understood the negative consequences of high rates of teacher turnover, worried about the depth of their hiring pool, and therefore were cautious about using the authority they had under current policies to dismiss teachers. The administrators used the flexibility in staffing policies judiciously, and relied heavily on working with their current teachers to build instructional capacity in their school.

**Committed to developing teachers.** In all six schools, administrators saw the primary purpose of evaluation to be developing the teachers they had hired. Samantha Nelson, Naylor Charter’s network head, explained that this commitment to improvement called for observing teachers frequently:

We do believe that our whole mission is to be a human capital organization. We are here to develop our kids. We are here to develop our teachers. We are here to develop our administrators. This is what we do and what we’re all about. ... Therefore, we think that the most transformational thing is just being in people’s classrooms, talking with them afterwards.

Across schools, when we asked about teacher evaluation, administrators repeatedly responded by describing processes for ongoing supervision. Kincaid Charter’s Principal Kain explained, “We believe that teachers, or just people in general, grow with immediate feedback and real time instruction on how they are performing and giv[ing] them an opportunity to fix that in the moment.” He also reported that he expected some teachers would struggle more than others, and that those in the school were prepared to support them. “If we have teachers who are struggling, it’s often times ... rooted in a lack of skill. Our job as coaches is to help them with that.”

Administrators in this sample provided long explanations of how, as one Fitzgerald evaluator said, they “coach [teachers] or find them the help they need.” The fact that the evaluators in many of these schools referred to themselves as “coaches” indicated their focus on supporting teachers’ growth.

Teachers across schools concurred with their administrators' descriptions, saying that supervision and evaluation focused on promoting growth for all teachers. They reported that frequent feedback was the norm at their schools. A Naylor Charter teacher explained,

I think it's just the overall culture of our school. I think that in order to be an employee here, regardless of if you're an academic teacher, co-curricular teacher, even a staff member, you need to want feedback ... and actually try and change it the next day to make it better. You have to have that open mentality that this is going to make me better and I want to get better... [to] help your kids.

In all schools, teachers explained that the process of supervision and evaluation was embedded in a professional culture that promoted continuous growth for teachers. Many described their schools much as this Rodriguez instructional coach did hers.

I think there is a culture here that is about continually getting better...that means that every teacher, whether they're getting feedback from an administrator or not, is trying to get better in their own practice. There is no one who's doing the same thing from year to year.

One teacher explained that she appreciated that the administrators at her schools were "continuing to develop [her] as a professional." Another asserted that the processes for providing frequent and ongoing feedback contributed to an "extremely supportive environment." She added that she wanted to stay at the school in large part because of this emphasis. Teachers across the sample consistently expressed appreciation that their administrators focused on the developmental purpose of evaluation.

**Schools were organized to prioritize instructional leadership.** Where possible, principals organized their leadership teams to allow key administrators to focus intensively on supervision and evaluation, while others managed responsibilities, such as discipline or building maintenance. Many public schools had recently suffered cutbacks in administration. However, Naylor Charter, Kincaid Charter and Hurston K-8 had administrative positions designed to relieve those who supported and evaluated teachers of additional responsibilities. There were

deans tasked with responding to concerns about students' behavior, directors of operations overseeing the logistics of running a school, and administrators responsible for managing student data, recruiting teachers, and organizing professional development. The Director of Operations at Hurston K-8 explained, "My role has been to block and tackle so that [the evaluators] can spend their time in the classroom coaching teachers and at [teacher] team meetings." Across these three schools, we heard similar explanations. Evaluators reported that they spent most of their time supervising teachers and interacting with teacher teams. Notably, however, Hurston's administrative team included the same number of adults as Kincaid Charter's, yet they had twice as many students and teachers. Therefore, unlike Kincaid and Naylor, Hurston K-8 did not have enough evaluators to provide the same intensive level of supervision for all teachers, even though they did their best to organize administrators' responsibilities to make that possible.

The three remaining schools—Fitzgerald, Dickinson, and Rodriguez Charter—had smaller, more traditional administrative teams, making it impossible for evaluators to focus exclusively on instructional leadership. Having at most an assistant principal—and in the case of Dickinson having only a principal—they could not hand over their management responsibilities to others. Nonetheless, they made time to regularly observe teaching, review lesson plans, meet with teacher teams and provide teachers feedback on their work. In addition, school leaders at Fitzgerald and Rodriguez Charter allocated resources to hire full time instructional coaches to support teachers.

It is also notable that in this sample, all of the administrators who were responsible for supervising teachers had, themselves, been teachers, several with more than twenty years of teaching experience, which enriched their approach to supervision and evaluation; six of fourteen had taught in the school they were now leading.

### **Teacher Evaluation Policies**

All six schools were subject to the new Massachusetts teacher evaluation policies, which required districts and charter schools to either adopt or adapt the state’s new model evaluation system or revise their current system to meet regulations. WUSD (and therefore the three district schools in the sample) had adopted the state’s model system. One of the charters, Rodriguez, had also chosen to use that system. The other two charters, Kincaid and Naylor, revised their existing evaluation frameworks to meet the state’s new regulations. In accordance with state regulations, all of the schools in the sample used detailed, standards-based frameworks for observations and assessments. Five out of six of the schools were using rubrics with four rating levels for multiple elements under each standard of practice. The sixth school, Kincaid Charter, used an evaluation tool that included multiple indicators for each category, but only had two rating levels—“present” or “not present.” According to state regulations, evaluators were required to observe each teacher one to four times per year, depending on the teachers’ improvement plan and ratings from prior years, where applicable. In all schools, teachers were active participants in the evaluation process, setting goals for student performance and professional practice.

At all six schools, evaluators were required to provide teachers with mid-year formative and end-of-year summative assessments, which consisted of ratings on the major standards as well as an overall rating. Five of the schools used student achievement data as a component of their rating system and assessed teachers’ progress in reaching specified targets for student outcomes. In all of the schools, teachers were asked to assess themselves using the evaluation rubric prior to their formal evaluation meetings. Teachers who consistently received low ratings were placed on more directed evaluation plans, which were intended to support improvement but also might lead to dismissal.

### **Evaluation as a Continuous and Unusually Comprehensive Process**

Across schools, teachers described supervision and evaluation processes as continuous and unusually extensive. Evaluators had many ways to assess teachers' work throughout their schools and to support teachers' improvement. In the district schools, teachers were required to submit artifacts as evidence of their progress toward professional and student achievement goals. Artifacts could include samples of student work, lesson plans, copies of communication with parents, or photographs. One Dickinson teacher said that she was grateful to provide artifacts because they made her supervisor aware of aspects of her work that she otherwise "would probably never even know," such as correspondence with parents. Fitzgerald's Principal Forte went beyond the district's policy requirements and expected all teachers to submit lesson plans to her or the assistant principal weekly. Each month, teachers also submitted a packet including samples of student work and student assessment data, to which the principal and assistant principal responded with individualized, written feedback. At Naylor Charter, teachers were required to submit unit plans two weeks prior to using them so that their supervisor and a network administrator could provide feedback. Teachers and administrators also reported that their supervisors stayed informed about their professional practice by participating in team meetings, which focused on data analysis and curriculum planning. Interestingly, it was difficult to determine which practices were defined as part of the evaluation system and which were not, since teachers and administrators described them all as elements of an ongoing, integrated improvement process. Although teachers appreciated the varied ways that supervisors interacted

with their daily work, most described classroom observations as the dominant and most valuable component of supervision and evaluation.

**Frequent observations and feedback.** Classroom observations played a central role in evaluation in all six schools. In district schools, evaluators were obligated to formally observe and provide feedback to each teacher throughout the year; the frequency of observations depended on that teacher's rating on the summative evaluation in prior years. For example, a new teacher had to be observed—once for an announced visit and four times for unannounced visits—while a returning teacher with a history of proficient or exemplary summative ratings had to be observed only once, unannounced. The district encouraged evaluators to observe and provide feedback to teachers informally and frequently. In this sample, the three charter schools and one district school, Hurston K-8, expected their administrators to observe in classrooms far more frequently than the state required.

Most teachers described an intense cycle of observations followed soon after by written or oral critique from their supervisor. In this sample, about 40% of the 100 classroom teachers we interviewed reported that they were observed and received feedback from their supervisor at least twice per month. Approximately 20% estimated that they were observed and given feedback between 5 and 10 times per year. The final 40% of teachers estimated that they had been observed 1 – 4 times per year, as prescribed in the new state evaluation policy. Overall, therefore, these schools went well beyond the state requirements, although some teachers experienced that more than others.

Kincaid Charter and Naylor Charter expected that every teacher would be observed and provided face-to-face feedback at least twice per month. All teachers interviewed said that evaluators met these expectations and sometimes went beyond them. At Hurston K-8 and

Rodriguez Charter, administrators aspired to observe every teacher and provide feedback at least once per month. They also routinely conducted “walk-throughs” for quick observations.

Teachers’ accounts suggest that, although many teachers were observed frequently, these schools did not have the resources to maintain intense supervision for all teachers. At Dickinson and Fitzgerald, teachers said their principals spent a great deal of time in classrooms throughout their school, but most talked about receiving formal feedback only a few times per year. These administrators did not have sufficient time for more frequent observations and focused more on regularly reviewing lesson plans and student data as well as participating in team meetings. At these two schools, neither administrators nor teachers mentioned a target number of observations per year, but they reportedly met the district’s requirement to conduct between one and four unannounced visits per teacher per year.

*Teachers appreciated frequent observations.* Teachers valued frequent observations and detailed feedback, describing them with phrases such as “hugely helpful,” or by saying they made them feel “super supported.” One teacher, expressing a sentiment echoed by many others, explained that the reason she was so satisfied in her job was “the constant feedback, and I constantly feel like I’m getting better.”

Across various levels of experience, teachers explained why frequent observations were advantageous. Many asserted that evaluators had greater credibility if they observed in classrooms regularly. A teacher with 11 years of experience at Dickinson said, “She [the principal] knows what we're doing. ...It's not like that old school where you never saw the principal. ...I feel like she's always in our classes. The kids don't really even pay her any attention. ...It's more natural, more organic.” A Hurston K-8 teacher with 10 years of experience compared the quality of observational feedback from two different evaluators in her school. “His

feedback is a lot more valuable than hers, just because he is more consistently in there and I just feel like he gets more of the nuances because he must go to a million classrooms a day.”

Teachers also reported that their evaluators had a better understanding of their professional experience and their particular struggles if they observed them teaching regularly. An experienced teacher at Rodriguez Charter explained, “He knows my flaws. He knows what I need to work on. He knows me better than I know myself as a teacher.” Some teachers, such as this third-year teacher at Rodriguez Charter, suggested that administrators demonstrated their investment in developing teachers by observing them frequently: “Just the fact that my administrators are in my classroom on a weekly or bi-weekly basis, I think shows a lot. It means that they care, and they’re here to help us. ... A lot of times we will get feedback that things need to change. The way it is delivered is almost always in a very positive way.” In all six schools most teachers appreciated that their supervisors spent a lot of time in classrooms, even if in some cases they did not provide feedback after every visit.

***Observations as a way to get help.*** Given the regularity of the observation and feedback cycle and the cultural norms that encouraged continuous efforts to improve, many teachers viewed supervisory observations as providing an opportunity to get help in areas they found challenging. One teacher explained this by contrasting her current experience with that in her previous school:

You’re so honest with everybody here, and nobody’s this expert of, “I can do everything perfectly.” It’s just calming...whereas in my old school ...you’d find out they were coming in [to observe]. It was like you were ready for a performance. You had to do it perfectly and then they never came in again until three or four months later. [Here], they’re just always in and out of the room, so it’s nice. It’s a good way to just always keep getting better.

A colleague, who gave an almost identical report, added,

When I know something isn’t going well, I will ask to be observed so that I can get help on that. That’s totally the mentality here. I don’t like someone seeing me doing

something wrong or seeing something that's failing, but I'd prefer that ... [to] not getting any guidance on it.

Teachers also described times when they were praised for asking a supervisor for help, which reinforced the expectation that the focus of the supervisory practice was on improvement.

*Teachers appreciated detailed feedback.* Across schools, teachers valued high-quality feedback—detailed, relevant comments and suggestions that helped them to improve their instruction. Teachers appreciated receiving feedback about a range of topics including classroom management, general pedagogical strategies and what some called “content feedback.” Many teachers shared examples of supervisors’ suggestions that they found helpful. For example, a teacher at Naylor Charter said that her supervisor had helped her improve the questions she asked during read-alouds so that she could encourage various types of thinking by students in response to text. A Kincaid Charter teacher discussed the helpful feedback she received about the ratio of teacher talk to student talk during her lesson. Hurston K-8 evaluators often emailed teachers with post-observation feedback, which teachers repeatedly described as specific and relevant. Samples provided by teachers included one case where Hurston’s Principal Hinds wrote about the pacing of a lesson and another where he noted that the discussion consisted largely of the teacher responding to every student’s contribution before the next student spoke. In both cases, the principal offered suggestions about how the teacher might adjust his/her practice.

Across schools, feedback went beyond support for classroom management. Although teachers frequently praised the feedback as detailed and insightful, many of the examples they shared with us focused on general pedagogical strategies applicable across content areas. It was unclear how often teachers received subject-specific pedagogical feedback, although several teachers said that they would like more.

**Day-to-day supervision contributed to teachers' sense of accountability.** Teachers reported that the intense levels of day-to-day supervision contributed to a sense of high expectations and professional responsibility within the school, which seemed to matter more to them than the ratings they received in the formal evaluation process. Many teachers talked about feeling accountable for their work because they knew that their supervisor or a peer might come in to observe at any time. One teacher explained, "I can't slack...knowing that [my supervisor is] in my class every two weeks. Knowing that other teachers are teaching my [lesson plans], there is no day that I can come in... not well-prepared.... I just can't lower my bar." Teachers at Fitzgerald also described feeling that they were held to high expectations for submitting and responding to feedback from their supervisor on weekly lesson plans and monthly data packets. One said, "I think it is where all schools should be, but it is demanding. I mean people are expected to turn in lesson plans, they're expected to get feedback on it and implement the feedback in their classroom every week."

**Formal evaluation matters, too.** The school leaders' developmental goals for evaluation were not at odds with the formal aspects of the summative process, which included mid- and end-of-year meetings to discuss ratings on the evaluation rubric. Teachers and administrators took these parts of the process seriously as supports for teachers' development and as a means of assessing teachers' effectiveness, which might inform current and future employment decisions.

Teachers generally viewed the formal evaluation process as providing an accurate assessment of their professional practice. A teacher at Rodriguez Charter described how his evaluator might compare the teacher's self-assessments with his rating: "He'll say,... 'I see you [put yourself] at a three; or you think you're at a two and a half, and I see you at a three. This is

why. ... I see you doing X, Y, Z.” Some teachers said that they respected the fact that even the rating process emphasized the goal of helping teachers to improve. A Naylor teacher explained,

You’re graded in a number of categories in a rubric from 1 to 4, just like the students are. It’s like a standards-based employee evaluation. I got mostly like 1.5s and some 2s. I think I had one 3. ...In another context, I would have felt like they were starting a paper trail to fire me. It was clear that actually that was right where they expected me to be. They expect their first-year, maybe even second-year teachers to be working hard, but not really mastering all the things they want you to master. It was a very positive conversation. I was offered a job the next year. At the same time, I felt like I had just walked into my teacher’s office and they had given me an F on the paper. It’s just a different scale. It’s like I went to school in Britain. I was not going to get a hundred percent on my essay anymore. They just don’t do that.

As this and many similar comments suggest, teachers trusted that the process was intended to support their growth and therefore they could appreciate tough assessments of their practice.

They did not expect or want a rubber-stamp process; nor did they think that falling short of the highest rating would be the first step out the door.

Nonetheless, across the sample, teachers and administrators recognized that evaluation could be used to hold teachers accountable for meeting professional expectations. Teachers believed that evaluators had given and would continue to give teachers low ratings on evaluation if they were justified and that those teachers could be dismissed. This contributed to a sense of accountability and made the evaluation process a serious one, but did not seem to generate fear or widespread anxiety among teachers. One administrator offered a perspective that distinguished between purposes of supervision and evaluation, which was consistent with her counterparts’ accounts from other schools, “I think that supervision is the way to get change in teaching...evaluation is really important when somebody doesn’t change.” At several of the schools, teachers and administrators spoke of teachers who were on official improvement plans with goals that needed to be met in order to keep their position. Administrators also told of

teachers who were not offered a position the following year and in some cases were asked to leave mid-year.

*“No surprises” formal evaluation processes.* Many teachers in the sample described formal evaluation processes as an outgrowth of day-to-day supervisory practice. Teachers variously described the formal evaluation process as “just a tiny piece of what we already do on a daily basis,” or “connected to the informal.” One teacher expressed the sentiments of many in saying that the formal process “shouldn’t be a big deal. It really hasn’t [been].” Another clarified why that would be the case: “I know exactly what my goals are and what I’m doing, so it wasn’t surprising how she graded me. I graded myself really hard, but I knew what I was working on so it made sense to me.” The administrators’ intentions matched the teachers’ experiences as explained by Principal Hinds.

I think evaluation without ongoing supervision is meaningless. It becomes only the way that you terminate employment. And so my belief is that I and every member of my administrative team needs to be in classrooms all the time, giving feedback, asking questions, pushing people. And then all of that just gets rolled into an evaluation. No surprises.

### **Supervision and Evaluation Embedded in Professional Learning Structures**

Ratings and feedback, which generally include a supervisor’s diagnosis and sometimes suggestions for improvement, do not necessarily lead to changes in practice. Donaldson (2012), who studied a school system in its second year of implementing a new teacher evaluation system, found that, although teachers had generally positive views of the process, they did not say that it had led to changes in their pedagogy. The reports of teachers in this sample stand in stark contrast to those findings.

**Teachers believed their practice was improving.** Strikingly, across schools teachers frequently reported that supervision and evaluation processes led to them to make changes in their pedagogy. A Kincaid Charter teacher with six years of teaching experience said,

I think I've become a drastically better teacher in the three years that I've worked here because it's been this really close cycle of being observed and then feedback on what to work on and then observed again and then feedback again. ... and also in ... being more self-aware.

An early elementary teacher at Rodriguez Charter who had ten years of experience described in more specific terms how his principal provided him with observational feedback, over time, which supported him in dramatically shifting his instructional approach.

She kind of said, "Why don't you think about doing this, that and the other thing?" I said, "Okay" and that first two, three, four weeks of changing my entire teaching style was a disaster. ...I started tweaking it and figuring it out and she would come in and observe and critique and give good positive comments and negative ones. ...Looking back I can't even imagine how much of a disservice I was doing to kids back then in the way that I was teaching.

Supervisors in these schools realized that their frequent feedback and suggestions for improvement were necessary, but not sufficient, for teachers to change their instructional practice. Therefore, they often connected the supervision and evaluation process to the school's other learning opportunities so that teachers could have additional sources of ongoing support for improvement. When interviewees spoke about supervision and evaluation, their comments often led them to discuss other collaborative learning processes such as coaching, practices for reviewing data, lesson planning structures, whole school PD sessions, and peer observations.

**Teachers' goals created alignment.** In all six schools, teachers were asked to set student performance goals and professional practice goals to guide their supervision and evaluation, and in several schools administrators and teachers explained how these goals were intentionally connected to school-wide and team-based goals. For example, at Hurston K-8, the teacher leaders for instructional teams set school-wide goals based on in-depth analysis of data. They

then facilitated their content and grade level teams in setting goals to guide the collaborative work of their team. Team goals also could be adopted or adapted by individuals for their evaluation process. An administrator at Hurston K-8 discussed the advantages of explicitly connecting these processes. “So there’s... an alignment from the individual to the team to the school that makes sense to people, and it doesn’t feel like they’re pulling [evaluation] goals out of the hat.”

Teachers across schools appreciated the fact that connections among individual, team and organizational learning goals focused rather than fragmented their efforts. Teachers in various schools described how this played out. One teacher told about how she and colleagues on her team were working “on open responses and the quality of our students’ writing in relation to what they read,” which had “driven a lot of the work [they] did as a team this year”; this was her goal for personal evaluation as well. A fourth grade teacher explained how she and her colleagues had developed a goal related to “reflection for fourth-grade students,” which elaborated on the school-wide PD focus on social-emotional learning. Notably, in all of the schools, teachers helped to define the connections among goals. None were mandated by administrators and most varied from one team to the next within a school.

**Supervision connected to one-on-one coaching.** Supervisors provided teachers with one-on-one coaching to help them understand and enact the suggestions provided in their supervisory feedback. Teachers praised instances when supervisors used observation debriefs to go beyond critique and help them implement suggestions. For example, a teacher with eight years of experience reflected on an observational debrief with his evaluator in which they planned questions that he might ask individual students during a subsequent lesson.

It was helpful to do because the meeting itself produced the question series. It’s good to have a product come out of those meetings, rather than just ... “Here’s what I saw, fix it”

type of things. Also because it was sort of teaching me how to do it myself. ...I found myself doing it more and more ... having that idea in mind.

Teachers also provided examples of their supervisors modeling instructional strategies with students in their classroom. In other situations, supervisors met with teachers to help plan lessons based on student data or feedback from an observation. A teacher described a series of one-on-one coaching sessions she had with her supervisor in conjunction with observational feedback.

Every Monday [she] and I meet after school and look up the lessons for the rest of the week... In that meeting for 20 minutes we debrief over very specific things that she knows I'm already working on—aligned exit tickets, and also what line of questioning is going to reach peak thinking the quickest. That's very quick cogent points. She looks at [lesson plans] and gives me feedback. ... I send her the new lesson plans that I've made, so she's constantly hearing from me.

At Rodriguez Charter and Fitzgerald Elementary, where the faculty included instructional coaches, evaluators often directed the coaches' work with individual teachers, based on information that they had gleaned from supervisory visits. A Fitzgerald coach explained,

Sometimes, [Principal Forte] is very directive with my time, like "I want you to spend three hours a week with [Teacher X] and help him write math lesson plans and go into his classroom to give him feedback on his teaching, specifically working on implementing lesson plans, because he writes great lesson plans, but then he doesn't implement them." And sometimes it is a lot more loose... she'll [share with] me about how a teacher is struggling and then say, "Could you work with them?" and then it's up to me to kind of put together a plan.

Many teachers and the administrators at Rodriguez Charter valued the ways that more general pedagogical feedback from a supervisor was complemented by content-specific, curricular support from a coach in the form of lesson planning, data analysis and in-class modeling of practices. Supervisors across the sample strategically coordinated one-on-one coaching to complement supervisory feedback, making it more likely that teachers could implement suggestions.

**Supervision also connected to other collaborative professional learning.** Evaluators also integrated topics from supervision into teacher team meetings and whole-school

professional development (PD) sessions to support teachers in their learning. Although the schools had varying amounts of team and whole-school learning time, administrators in the schools strategically planned how this time was used, creating coherent learning experiences for teachers. For example, at Kincaid Charter an evaluator explained that she used her classroom observations to identify “bigger-picture trends,” that is, needs that were apparent across the school, which might inform the topics for weekly whole-school PD sessions. She further explained that in response to a “trend” in observations she might also guide a teacher leader to work with her team on that particular topic during their planning block. Teachers appreciated this multi-faceted approach to supporting improvement and many provided examples like this teacher’s: “Student discussion was an area of growth...they gave me some strategies [to use] in [feedback]. We’ve also had PD around what teacher talk should sound like...how much time I’m talking versus how much time they’re thinking.”

At Fitzgerald, administrators also integrated professional learning opportunities, but rather than making PD responsive to supervision, they focused individual supervisory feedback on the established long-term PD topic—fostering higher-order thinking among students. They created an observation feedback form, which included twelve indicators from the teacher evaluation rubric that they believed relate to fostering higher order thinking. A Fitzgerald teacher with six years of experience appreciated the efforts to create connections among professional learning opportunities.

[I was] really impressed with the amount of consistency and thoroughness with which the whole topic of higher order thinking was integrated and just drove the instructional plan for the school. ... Therefore, that was the theme that went all the way through a lot of our meetings; higher order of thinking in writing, higher order of thinking questions in math, ...all the topics are related to that. That was a good cohesive theme.

**Supervision connected to targeted peer observations.** In some cases, teachers reported that their supervisor suggested and orchestrated an opportunity for them to observe a peer (from

their school or another) as follow-up to feedback from a supervisory observation. A teacher at Fitzgerald recalled,

A couple of weeks ago...one of the teachers in the second grade was having a problem teaching... three digit addition with and without regrouping. The instructional coach ... came in with her to observe the way that I did it. How did I execute my lesson? The vocabulary. What did I allow the kids to do? The turn and talk. The group work.

A Rodriguez teacher said that it “ helped a lot, just to see” a colleague teaching. Teachers and administrators in these schools explained that peer observations were used as a way to help teachers better understand a teaching approach or a particular aspect of instructional practice that they were trying to adopt or improve.

Across schools, administrators strategically coordinated the many professional learning activities in their school in order to optimize the possibility of improving teaching quality. Supervision and evaluation processes were intertwined with the work of instructional coaching, teacher teams, whole school PD, and peer observation.

### **Limitations in Implementing Supervision and Evaluation Processes**

Despite the overwhelmingly positive views of supervision and evaluation, teachers and administrators experienced challenges as they implemented the process. Understanding these not only explains what more would be required to improve evaluation, but also suggests what it takes for any school to design and implement an evaluation process that teachers experience as supportive of their learning.

**Mismatches between evaluators and teachers.** Across schools, some teachers were disappointed with the type of feedback they received and often attributed the shortcomings to the fact that their supervisors did not have knowledge and experience that matched the content they taught. These teachers explained that their supervisors were able to provide feedback about

classroom management and general pedagogy, but not content-specific pedagogy or curriculum.

A music teacher in one school expressed the views of other teachers in similar situations:

I think that, being in the arts, those supports over the past two years have developed me into being a person who has strong management. ... I haven't necessarily had that person really driving my curriculum. ... I've had the autonomy to do what I want and figure it out. ... I've been less developed in that regard, but as far as management goes, it's very supportive.

An early-career science teacher at a charter school described the feedback she received from her supervisor, a former math teacher.

My feedback has mostly only been on ... talking less and having my kids talk more. Near the end of the school year I felt like she was just trying to find little things that she could tell me to improve on. ...—I don't know if she just didn't know what I needed improve on ... I assume I have lots to improve on because I'm definitely not perfect. ... [S]he had no feedback for me on content.

A middle school math and science teacher at Naylor Charter described a similar experience. “She often gives me good feedback more about management, but as far as content, I don't get much.”

A history teacher who was supervised by a former English teacher said that his feedback often focused on how to teach English skills in history class, but neglected the “nitty gritty of history.”

Teachers of students with special needs expressed concern because their supervisors did not have experience or knowledge specific to special education. “I don't feel that my [evaluator] can coach me in special education.” Her colleague agreed and elaborated further, saying that he found the feedback to be “very standard...cookie-cutter” not addressing the particular challenges he faced due to his students' learning disabilities. Notably, these teachers consistently expressed a desire to improve. As one teacher explained, “I think there's still huge amounts of growth I could make, but it's hard accessing that growth when the people up ahead of you don't know what you're doing.” Across the sample, teachers who thought that their supervisor lacked experience and knowledge specific to their teaching assignment said that this significantly limited the benefits of supervision and evaluation.

Several administrators acknowledged the challenge of mismatched expertise, and we saw evidence of the ways in which they tried to address it. An administrator at Kincaid Charter expressed her concerns about supervising teachers in a subject she had never taught.

I think it's really detrimental to the quality of instruction that students receive. ...I think I'm doing moderately well, but I don't have a [subject] background and I'm doing my darnedest to support my teachers. I would love more support in that arena. ...I think it's actually really hard to find an expert in any area in our organization. People who are moderately qualified are considered experts. I think that's a concern.

At Fitzgerald Elementary and Rodriguez Charter, the administrators relied on instructional coaches to support teachers in curriculum planning and subject-specific pedagogical strategies. Teachers and administrators appreciated the intentional pairing of support that was possible when they had specialized coaches on the faculty. However, they only had coaches in mathematics and literacy, so this did not fully address the challenge, and some schools had no coaches at all.

Across schools, many administrators relied on teachers within their faculty as sources of content knowledge and subject specific-pedagogical feedback. At Dickinson Elementary, Principal Davila was the only administrator and therefore had to evaluate all teachers in her school. She explained, "I don't have all the answers...I taught kindergarten for years. Fifth-grade math? I do my observations...I say [to teachers] 'Learn from one another.'" Davila and other school leaders created many opportunities for teachers to learn from each other through peer observation, collaborative data analysis and other team discussions. Principal North at Naylor Charter periodically invited teachers to observe colleagues along with her to provide content-specific feedback. She explained that she might ask a veteran math teacher to participate in a collaborative debrief, explaining to him, "Content-wise I think I know the feedback I need to give, but would love your support on it." Other principals allocated time for veteran teachers to meet with colleagues who taught the same content but weren't on the same teacher teams.

Teachers valued these opportunities to work on curriculum and content-specific pedagogy. A middle school math teacher at Hurston K-8 described observations by administrators as “affirming,” but said she found observations by her colleague who also taught math to be more helpful. “He just knows more about the content. He can tell if students are understanding or not a little bit more than they can because not everybody’s an expert in everything.” Although teachers proved to be valuable resources for each other, the veteran teachers who were expected to support others often said they wished for better feedback, themselves.

Administrators did not always succeed in finding expert support for every teacher’s ongoing development. Nonetheless, teachers endorsed the overall process of supervision and evaluation, while wishing that it could better meet their subject-specific teaching needs.

**Insufficient Time.** Second, participants said that evaluators had insufficient time to provide comprehensive supervision to all teachers. Although most teachers in the sample praised frequent observations and feedback from supervisors, some teachers—ranging from novices to veterans—at Rodriguez Charter, Hurston K-8, Fitzgerald and Dickinson, wished for more.

Administrators at these four schools spoke of the daunting amount of time and resources required to conduct frequent observations and provide detailed feedback for all teachers. Most administrators in the sample had between 15 and 20 supervisees but some had more, with the highest number being 39 for one of the principals. The principal at Fitzgerald Elementary said, “We just can’t keep up. We’re lucky to have two of us.” Similarly, an administrator at Rodriguez Charter said, “I have 20 people I evaluate and supervise, and it feels like too many to me. I’m always thinking, ‘Oh, I haven’t been there for so long!’”

Teachers generally believed that administrators spent more time supervising new and struggling teachers than competent veteran teachers, who still might improve. Experienced

teachers understood why novices' learning needs were a higher priority, but nevertheless wished for more intensive support so that they might further develop their practice. For example, a Hurston teacher with ten years of experience said that her supervisor's periodic observations were not enough: "I don't want oversight but I would like more feedback—someone who knows my classroom, has seen Student A in October and now can tell me how Student A progressed in March." A colleague with nine years of experience wished to be observed more frequently so that she could have in-depth discussions about her "delivery of instruction," such as "Did it make sense to do that activity ... in groups?" Although the Hurston evaluators were providing this level of feedback to other teachers in the school, they did not have sufficient time to invest similarly in teachers they deemed effective. At each of these four schools, a small number of teachers expressed interest in having more frequent observations and feedback from their evaluators.

### **Discussion and Conclusion**

Teachers in this sample of six successful, high-poverty schools appreciated receiving frequent feedback on their instructional practice and believed that the supervision and evaluation process helped them to improve. Two factors enabled these school leaders to use evaluation principally as a process to support teachers' professional growth. First, administrators in these schools had considerable flexibility in hiring, transferring and dismissing teachers, so that they did not need to rely entirely on evaluation to dismiss underperforming teachers. Second, some of the school administrators had the resources and flexibility to organize their administrative team's roles and responsibilities so that evaluators could spend their time almost exclusively on instructional leadership.

In all six schools, teachers and administrators consistently described the primary purpose of the process as supporting improvements in teaching quality although they reported that it also played a role in holding teachers accountable for their professional work. By closely connecting supervision and evaluation to other professional learning, administrators not only increased the relevance of the feedback but also were able to provide substantial support to aid teachers in responding to suggestions made by supervisors. Notably, teachers from novices to veterans expressed interest in this process for supporting their growth. There are lessons learned and questions that emerged from the similarities and differences across schools as well as the implementation challenges.

### **Implications for Policy and Practice**

In many ways, the policy environment in which these schools operated supported their approach to supervision and evaluation. The Massachusetts Model System for Educator Evaluation and the state's new regulations emphasized a cycle of continuous improvement. The evaluation policies allowed these schools to continue and build upon their existing focus on developing teachers. In addition, to support implementation of the policies, Massachusetts DESE and Walker City School District provided schools with detailed resources including models from practice. For example, they disseminated samples of how individual teachers might align their evaluation goals with improvement goals for their teacher teams and their schools. The administrators and teachers in this study appreciated working under teacher evaluation policies that aligned with their approach to the process.

This study reinforces previous research showing that doing supervision and evaluation well requires evaluators to possess content and pedagogical knowledge, coaching skills, and a readiness to attend to both development and accountability (Kane & Cantrell, 2013; O'Pry &

Schumacher, 2012; Sartain et al., 2011). This is a substantial challenge for most schools and school systems. All of the supervisors in this sample were reported to attend to developmental goals for teachers, which we found remarkable in contrast to our previous study where evaluators in only one of six schools were reported to have this orientation (Reinhorn & Johnson, 2014). However, by their own account, they struggled to find sufficient time to complete a high level of supervision for all teachers. In this sample, even with a cohort of experienced former teachers serving as evaluators, the schools struggled to match evaluators' expertise with teachers' assignments, leaving some teachers wishing for more from the process.

**A time-intensive process.** Teachers in this sample appreciated frequent feedback and support when it was provided. However, the comprehensive process described in these schools is extremely time-intensive and would require a school or school system to commit significant resources to support it, both initially and over time. Schools, school systems and state departments of education must consider how much time they are willing to devote to developing their teaching force, particularly in high-needs schools; what role supervision and evaluation is going to play in that process; and how to provide schools the flexibility, resources, and support to accommodate their efforts. At Naylor Charter and Kincaid Charter, policy flexibility and sufficient resources existed so that administrators could focus almost exclusively on instructional leadership, while others took responsibility for discipline or building operations. Hurston K-8 and Fitzgerald Elementary had the freedom to reorganize their administrative teams during the turnaround process and they both added positions, permitting the principals and other key administrators to devote more time to instructional leadership. But when Fitzgerald and Hurston exited turnaround status, reductions in their funding obliged them to decrease the size of their administrative teams making the reforms difficult to sustain. At the time of this study, Fitzgerald

had returned to having a traditional elementary administrative team with a principal and assistant principal, while Hurston, despite retaining non-traditional roles, was administratively understaffed relative to their goals. Rodriguez Charter and Dickinson also would have benefitted from more funding to increase the size of their leadership teams. Across schools, additional evaluators and additional administrators focusing on non-instructional management tasks would have created more opportunities for greater investment in supervision. Some of these schools—both charter and district—were involved in raising funds or creating partnerships with non-profit organizations in order to support their efforts, although even these were not sufficient to meet their needs.

In order to effectively implement new teacher evaluation policies, schools serving high-poverty communities may, over time, need more financial support than others where students have fewer needs and teachers require less support. Increased funding during turnaround efforts may allow schools to add administrators and coaches to their faculties, helping to solve the problems of time and content expertise. Yet when this funding subsequently is withdrawn after the turnaround process, the schools may be forced to revert to traditional administrative teams, which cannot sustain a robust supervision and evaluation process.

**Matching content expertise for supervision and evaluation.** In this study, when there was a good match between the evaluator’s knowledge and the teachers’ assignment, teachers benefitted. However, when the evaluator’s content knowledge did not match the teacher’s assignment, feedback focused on general pedagogy and classroom management. Although helpful, this did not support teachers’ growth in subject-specific pedagogy.

These schools went to great lengths to create productive matches and supplement those with support from coaches and peers where appropriate. And yet, this challenge persisted.

Almost any school could benefit from identifying and relying on effective teachers in different content areas, whether as formal or informal coaches, as these schools did. Ultimately, though the person considered most knowledgeable in the school would be left seeking support as well. In addition, teachers in some content areas had no peer in the school. Therefore, the solutions to the challenge of content match in supervision are unlikely to exist completely within the confines of a single school. Schools (particularly small schools), CMOs and districts should consider how to work across schools to capitalize on expert knowledge in the field to better inform the supervision and evaluation process. As demonstrated in this sample, teachers are likely to be part of the solution.

**Integrating professional learning opportunities.** As schools and school systems seek to solve these challenges, it is important to recognize the power of integrated strategies for developing human capital in schools. The school leaders in the sample viewed supervision and evaluation as part of their broader strategy to develop their teachers. Teachers benefitted when their supervisors connected one-on-one supervisory experiences with their professional learning in teacher teams and in whole school PD. Other school leaders can learn from their successes in integrating teachers' learning opportunities for teachers.

**Preparing principals to lead school-based human capital development.** New evaluation policies continue to have great promise as a way to support improvements in teaching quality for all teachers—not just underperforming teachers. This is of vital importance as schools and school systems venture into the uncharted territory of the Common Core State Standards. However, it is evident that this promise can only be achieved with significant investments in preparing administrators who lead and implement the evaluation processes with a focus on learning and improvement.

Principals' development programs in universities, school systems and independent organizations, can assist in these efforts by preparing school leaders who value and understand how to invest in developing teachers in their schools. Unless principals view this as a core function of their job, it will always be superseded by the managerial tasks that can easily dominate a school leader's time. In these schools, when school leaders were able to organize so that they could focus on instructional leadership, teachers viewed it as a productive investment.

School systems will need to hire school leaders—ideally experienced, expert teachers themselves—who are committed to supporting continuous improvement and can create or sustain a culture that prizes growth for all its community members, including them. At Naylor and Kincaid Charter Schools, organizational structures allowed for the evaluators to be supervised and evaluated by CMO administrators regularly, much like what is happening with teachers in these schools. Toward this end, the school, school system, or CMO will need to invest in continuing to develop their evaluators after they are hired, preparing them to go well beyond achieving validity and reliability by calibrating their judgments and ratings. If supervision and evaluation are to truly improve teaching throughout a school, supervisors must understand teachers' daily work, support improvement, and integrate their frequent classroom visits and feedback with professional learning throughout the school.

### **Implications for Research**

We were especially struck by the degree to which evaluation was integrated with other school improvement efforts in these six schools. As researchers seek to understand whether and how teacher evaluation practices lead to improvements in teaching quality, they will likely need to broaden their focus and consider the array of practices that may be interacting with evaluation. Future research on evaluation could examine how effective evaluation practices interact with

staffing policies and other professional learning opportunities to support development of teaching capacity within schools.

Given that the policy context played a significant role in how these schools approached evaluation, further research is warranted. This sample was small and contained schools with particular histories and flexibilities in how they organized and operated, all made possible by various state and local policies. Many questions remain about how schools are enabled and constrained by the systems and the policy environments in which they operate. In particular, there is more to learn about the similarities and differences across charter schools, district schools, and within-district charter schools.

Many questions remain about how, if at all, teacher evaluation processes support improvements in teaching quality. In this study, teachers appreciated frequent, detailed feedback, but in some cases wished for more content-specific critiques of their work. Teachers also reported on various types of supervision ranging from observational feedback to commentary on lesson plans. Future research on teacher evaluation could investigate the effects these different types of feedback have on instruction. What value do different types of feedback have, given the investment of time required? Future research is also warranted about the benefits of different levels of frequency of observations and feedback to teachers. Is there an optimal frequency to support teachers' growth and does it vary depending on the teachers' experience? How, if at all, does an evaluation process change over time if a faculty experiences frequent or infrequent turnover?

This study reveals the value of integrating evaluation processes into a school's larger program for instructional improvement. However, implementing evaluation processes in this way

## Supporting Continuous Development for Teachers: Teacher Evaluation

is a challenging undertaking requiring skillful school leaders, ample resources and policies that afford these leaders significant school-based control. Across the U.S., states and school systems have invested heavily in reforming their evaluation policies; now they must take the next important step to invest in schools and support school leaders as they implement these policies.



### References

- Aldeman, C., & Chuong, C. (2014). *Teacher evaluations in an era of rapid change: From “unsatisfactory” to “needs improvement.”* Bellwether Education Partners. Retrieved from [http://bellwethereducation.org/sites/default/files/Bellwether\\_TeacherEval\\_Final\\_Web.pdf](http://bellwethereducation.org/sites/default/files/Bellwether_TeacherEval_Final_Web.pdf)
- Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research*, 78(3), 367–409.
- Carroll, S., Reichardt, R., Guarino, C., & Mejia, A. (2000). *The distribution of teachers among California’s school districts and schools*. Santa Monica, CA: RAND Corporation. Retrieved from [http://www.rand.org/pubs/monograph\\_reports/MR1298z0](http://www.rand.org/pubs/monograph_reports/MR1298z0)
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2013). *Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood* (Working Paper No. 19424). National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w19424>
- Clotfelter, C. T., Ladd, H. F., Vigdor, J. L., & Wheeler, J. (2007). High-poverty schools and the distribution of teachers and principals. *North Carolina Law Review*, 85(1348).
- Danielson, C. (2013). *The Framework for Teaching Evaluation Instrument, 2013 Edition: The Newest Rubric Enhancing the Links to the Common Core State Standards, with Clarity of Language for Ease of Use and Scoring*. The Danielson Group.
- Dobbie, W., & Roland G. Fryer, J. (2011). *Getting beneath the veil of effective schools: Evidence from New York City* (Working Paper No. 17632). National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w17632>
- Donaldson, M. L. (2009). *So long, Lake Wobegon? Using teacher evaluation to raise teacher quality*.
- Donaldson, M. L. (2011). *Principals’ approaches to developing teacher quality: Constraints and opportunities in hiring, assigning, evaluating, and developing teachers*. Center for American Progress.
- Donaldson, M. L. (2012). *Teachers’ perspectives on evaluation reform*. Center for American Progress.
- Donaldson, M. L. (2013). Principals’ approaches to cultivating teacher effectiveness constraints and opportunities in hiring, assigning, evaluating, and developing teachers. *Educational Administration Quarterly*, 49(5), 838–882.
- Donaldson, M. L., & Peske, H. G. (2010). *Supporting effective teaching through teacher evaluation: A study of teacher evaluation in five charter schools*. Center for American Progress.
- Downey, D. B., Hippel, P. T. von, & Hughes, M. (2008). Are “failing” schools really failing? Using seasonal comparison to evaluate school effectiveness. *Sociology of Education*, 81(3), 242–270.
- Doyle, D., & Han, J. G. (2012). *Measuring teacher effectiveness: A look “under the hood” of teacher evaluation in 10 sites* (No. 50Can). Connecticut Coalition for Achievement Now (ConnCAN).
- Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record*, 103(6), 1013–1055. <http://doi.org/10.1111/0161-4681.00141>
- Hanushek, E. A. (2009). Teacher deselection. In D. Goldhaber & J. Hannaway (Eds.), *Creating a new teaching profession* (pp. 109–118). Washington DC: The Urban Institute.

- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources*, 39(2), 326–354.
- Ingersoll, R. M. (2004). Why do high-poverty schools have difficulty staffing their classrooms with qualified teachers. *Washington DC: Center for American Progress and Institute for American's Future*.
- Kane, T. J., & Cantrell, S. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study*. Retrieved from [http://www.metproject.org/downloads/MET\\_Ensuring\\_Fair\\_and\\_Reliable\\_Measures\\_Practitioner\\_Brief.pdf](http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf).
- Kimball, S. M., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, 45(1), 34–70. <http://doi.org/10.1177/0013161X08327549>
- Kraft, M. A., & Papay, J. P. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*, 0162373713519496. <http://doi.org/10.3102/0162373713519496>
- Ladd, H. F. (2011). Teachers' perceptions of their working conditions: How predictive of planned and actual teacher movement? *Educational Evaluation and Policy Analysis*, 33(2), 235–261. <http://doi.org/10.3102/0162373711398128>
- Maxwell, J. A. (1996). *Qualitative research design*. Thousand Oaks, CA: Sage Publications.
- McLaughlin, M. W., & Pfeifer, R. S. (1988). *Teacher evaluation: improvement, accountability, and effective learning*. Teachers College Press.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks: Sage Publications.
- National Council on Teacher Quality. (2013). *2013 State Teacher Policy Yearbook*.
- New Teacher Project. (2010). *Teacher Evaluation 2.0*. New Teacher Project.
- O'Pry, S. C., & Schumacher, G. (2012). New teachers' perceptions of a standards-based performance appraisal system. *Educational Assessment, Evaluation and Accountability*, 24(4), 325–350. <http://doi.org/10.1007/s11092-012-9148-4>
- Reinhorn, S. K., & Johnson, S. M. (2014). Can evaluation provide both accountability and development for teachers? Evidence from six high-poverty schools.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417–458. <http://doi.org/10.1111/j.1468-0262.2005.00584.x>
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review*, 94(2), 247–252.
- Sartain, L., Stoelinga, S. R., & Brown, E. R. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation*. Consortium on Chicago School Research.
- Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *The American Economic Review*, 102(7), 3628–3651.
- Taylor, K. (2015, March 22). Cuomo fights rating system in which few teachers are bad. *The New York Times*. Retrieved from <http://www.nytimes.com/2015/03/23/nyregion/cuomo-fights-rating-system-in-which-few-teachers-are-bad.html>
- Thomas, E., Wingert, P., Conant, E., & Register, S. (2010). Why we can't get rid of failing teachers. *Newsweek*, 155(11), 24–27.

- Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education*. Education Sector.
- Tucker, P. D. (1997). Lake Wobegon: Where all the teachers are competent (or, have we come to terms with the problem of incompetent teachers?). *Journal of Personnel Evaluation in Education*, 11(2), 103–26.
- Wechsler, M., Tiffany-Morales, J., Campbell, A., Humphrey, D., Kim, D., Shields, P., & Wang, H. (2007). The status of the teaching profession 2007. *Santa Cruz, CA: The Center for the Future of Teaching and Learning*.
- Weisberg, D., Sexton, S., Mulhern, J., Keeling, D., Schunck, J., Palcisco, A., & Morgan, K. (2009). *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness. Second Edition*. New Teacher Project.
- Wise, A. E., Darling-Hammond, L., Tyson-Bernstein, H., & McLaughlin, M. W. (1984). *Teacher evaluation: A study of effective practices* (No. R-3139-NIE). RAND Corporation. Retrieved from <http://www.rand.org/pubs/reports/R3139.html>

## Appendix A

Table 2. *Number of Interviewees at Each School*

School Name	Administrators*	Non-Teaching Staff **	Teachers in Training	Teachers	% of Total Teachers in the School Interviewed
Dickinson Elementary	1	2	n/a	15	56
Fitzgerald Elementary	2	2	n/a	14	47
Hurston K-8	4	5	n/a	21	31
Kincaid Charter Middle	5	4	2	16	38
Naylor Charter K-8	2	3	2	17	46
Rodriguez Charter K1-8	3	3	3	16	36

\* Administrators include directors of CMOs and school based administrators who directly supervise teachers.

\*\* Non-teaching Staff includes instructional coaches, parent coordinators, data leaders, recruitment officers, deans of discipline and other administrators who do not teach students and do not supervise teachers

Table 3. *Total Teachers Interviewed at Each School & Years of Experience*

School	Novice* (1- 3 years)	2nd Stage (4 - 10 years)	Veteran (11+ years)
Dickinson Elementary	3	5	7
Fitzgerald Elementary	1	11	2
Hurston K-8	6	11	4
Kincaid Charter Middle	4	11	1
Naylor Charter K-8	8	7	2
Rodriguez Charter K1-8	1	9	6
<b>Totals</b>	<b>23</b>	<b>54</b>	<b>22</b>

\*Does not include Teachers in Training

**Appendix B: Interview Protocols**  
**Teacher Interview Protocol**

Intro: Study Explanation emphasizing that we really want to learn about your experience at this school.

1. Background:
  - a. How did you come to be in your current position at this school?
  - b. Starting with college, can you tell us what you've done?
    - i. Probe for: training and employment
2. Current Teaching Assignment:
  - a. What do you teach here?
  - b. How did you wind up in this position?
3. Overall view of school:
  - a. If another teacher would ask you, "What's it like to teach at \_\_\_\_\_?" How might you respond?
  - b. What are the advantages and disadvantages of being a teacher here?
4. Hiring:
  - a. How were you hired at this school? Step-by-step.
  - b. Do teachers play a role in hiring other teachers? If so, how?
  - c. Has the hiring process changed at this school? If so, how and why?
5. Induction:
  - a. Did you have some kind of induction as a new teacher at this school? What worked and what didn't?
  - b. How are new teachers inducted now? How have things changed since you got here?
6. Support:
  - a. What kinds of supports are available here for teachers to improve their instruction?
  - b. What works well for you? What doesn't? (Probe: PD, Coaching, Collaboration, Evaluation)
7. Evaluation:
  - a. How is your teaching evaluated? Describe the process.
  - b. Was it helpful? How?
8. Administration:
  - a. Who do you go to for support? For what?
9. Social & Psychological Supports:
  - a. What sorts of social and psychological supports does your school offer for students?
  - b. What support do you get for interacting with parents and families?
10. Career goals:
  - a. How long do you expect to stay at this school? In what roles?
    - i. If yes: What keeps you at this school?
    - ii. If no: Why do you think you might leave?
11. Union:
  - a. What role does the union or the contract play in this school?
12. More: Do you have any additional comments?

# Supporting Continuous Development for Teachers: Teacher Evaluation

## Principal Interview Protocol

Overview of Study: 6 Schools, All high-poverty, high-minority. All Level 1.

1. Background:
  - a. How long have you been at this school? Prior experience in education? Anything else we should know about how you got here?
2. School Overview:
  - a. Could you first provide an overview of its structure and programs?
  - b. (Where applicable) What does it mean for your school to be a pilot/turnaround/charter school?
  - c. (Where applicable) How did you go about selecting teachers when ---- was placed in turnaround?
  - d. How would you describe it to a teacher or parent who might be interested in it—both its strengths and weaknesses?
3. Teachers: We'd like to get a sense of who your teachers are.
  - a. Where do they come from?
  - b. What formal or informal preparation do they have?
  - c. What attracts them to the school?
  - d. Approximately, what proportion has fewer than 10 years of experience? 5 years of experience? 0-5 years of experience? (Has that changed or remained steady?)
4. Recruitment and Hiring:
  - a. Could you describe the process you use to recruit and hire teachers? (Applicants per position? Teaching demonstration? Who decides?)
  - b. What challenges do you face in recruiting teachers?
  - c. Are there specific demographics or subject areas that you have trouble finding/attracting? If so, how have you addressed those challenges?
5. Assignment:
  - a. How do you assign teachers to a particular grade or subject?
  - b. Could you describe the teachers' responsibilities, both during school hours and outside of school hours? Scheduled and unscheduled time?
6. Compensation:
  - a. Please tell us about the pay scale for teachers. Are there additional stipends? If so, can you describe these opportunities?
7. Collaboration:
  - a. Are the teachers organized by teams, grade-levels, subjects? If so, what does that mean for how they do their work? What is the work of those teams?
8. Supports:
  - a. What supports can a new teacher count on in getting started? And for more experienced teachers?
9. Role:
  - a. Are there specialized roles for some teachers? (Teach Plus, team leaders, etc.) If so, please describe these roles.
10. Curriculum:
  - a. Does the school provide a curriculum for the teachers? If so, please tell us about it.
11. Professional Learning:

## Supporting Continuous Development for Teachers: Teacher Evaluation

- a. Do you have formal professional development? Instructional coaches? If so, please tell us about them.
12. Supervision and Evaluation:
  - a. How do you supervise teachers? How do you evaluate teachers? Are these separate processes? Do students' test scores play a role in evaluating teachers?
13. Dismissal:
  - a. How frequently do you dismiss or decide not to rehire a teacher? Reasons?
14. Retention:
  - a. How long do teachers stay? Why do they stay? Why do they leave? Is there a type of teacher who stays or leaves? Is turnover a challenge?
15. Policy Context:
  - a. Does state or local policy play a role in how you approach building your teaching capacity?
16. Union:
  - a. What role if any does a teachers' union play at your school?
17. Have we missed anything?

**Appendix C: List of Codes and Descriptors**

<b>Codes</b>	<b>Description</b>
Assignment	Teacher Assignment: What do you teach/ your job at the school, views of your assignment
Background	Background: Past work history, education
WhyTeach	Why teach? Personal sense of purpose can include changes in views over time.
SchoolOverview	Facts about the school (the facts but not mission or culture), might include specific school goals
HistorySchool	History of School
FacultyComposition	Descriptions of the composition of the faculty
Equity	Interactions, policies or dynamics described in relation to race, ethnicity, social class
Hiring	Related to teacher recruitment, hiring, including teacher's experience of being recruited / hired. –timing, demo lessons, debriefs, meetings with current teachers, written applications, Who the school seeks and how they find candidates
WhySchool	Why chose school - why teach at this particular school? May reflect changes over time
Mission	Descriptions of what the school aspires to accomplish (if explicitly talking about mission do not double code with culture)
AdminAdmin	Interactions / relationships among administrators (including non-teaching positions such as coaches and guidance, deans and other non-teaching roles)
AdminLeadership	Descriptions of administrators' style, vision, agenda, priorities, purposes, etc. (includes self-descriptions)
AdminRole	Specific responsibilities and job descriptions of non-teaching faculty
AdminStudent	Relationship between administrators and students (include coaches, guidance, deans and other non-teaching roles)
AdminTeach	Interactions between administrators and teachers (include coaches, guidance, deans and other non-teaching roles in this code)
Demands	Teachers professional responsibilities and expectations, work hours, teachers views on demands
Resources	Material and human resources (money, buildings, positions, --if it is about admin roles will be double coded in Adminroles) - Facilities
Accountability	Related to external accountability (state accountability status and state testing, turnaround status) - what the state does and then what is done as a result
DistrictNetworkState	Formal relationships / governance from State, District or CMO, includes school boards and trustees
Testing	References to standardized tests, state tests, network tests and interim assessments and how used in the school
MonitoringStudents	Teachers' use of assessments and instructional strategies to monitor achievement
Student	Student Characteristics: Descriptions of students and their community
Neighborhood	Descriptions of the local surroundings of the school
Family	Ways of connecting families and community to school, Perceptions of parents/ families + Teacher and admin connections to parents / families
CurriculumPedagogy	What and how you teach - including instructional planning
SpEdELL	Descriptions of programs or approaches for educating students with special needs and /or ELLs
SpecialSubj	Referring to non-core academic classes (art, music, library, dance, etc) and extra-curricular or co-curricular programs or activities

## Supporting Continuous Development for Teachers: Teacher Evaluation

SchoolCulture	Expressions of school-wide norms & values including kids, teachers and parents (not explicit statements of mission), big picture that everyone from school would understand
Colleagues	Commentary on colleagues and their characteristics (what I think about the people I work with) - big picture impressions of colleagues
ProfCulture	Professional Culture- the norms of being a teacher or admin in this school. Big picture expectations for how we work together as professionals
Eval	Related to teacher supervision and evaluation: observations, feedback, meetings between supervisors and teachers, how work with teachers on instruction
Coaching	Formal instructional coaches, but NOT induction mentoring
Induction	Programs and supports (formal and informal) for new teachers: prior to day 1 and after day 1
FormalCollab	Deliberate, structured groups working together-organized by the school- including whole school sessions – including approach to lesson planning and who is included and who is not - JUST TEACHERS
InformalCollab	Specific work with colleagues that is not organized by the school, informal collegial interactions -JUST TEACHERS
RolesTeach	Formal roles and opportunities for career advancement (Teach Plus etc.) may have double coding when example of influence through a formal role including leadership teams, Teachers in Training, etc.
InfluenceTeach	Teacher opportunities as brokers of influence (teachers generally in their work having influence), including committees where you can voice concerns - Admin change view because of a teacher
CareerGrowth	Individual professional growth for career progression
OrderDiscipline	Safety, systems, expectations and rules for students, and enforcement
StudentSupports	Social and emotional and academic supports for students and behavioral - outside of classroom structure
StudentTeach	Interactions among teachers and students inside and outside the classroom.
Turnover	Why other people stay or leave; both causes and frequencies, personal plans to stay or leave, also about satisfaction and dissatisfaction, might be stuck in job
Pay	Payscale, stipends and other things related to compensation
Partners	Partners including City Year, Teach Plus, Ed Schools, etc.
Union	Related to the union and the contract
Gem Quote	This is a great quote.

**Descriptor Categories for Characterizing Interviewees**

<b>Descriptor</b>	<b>Definition</b>	<b>Response options</b>
School	School Name	
Grade Level	Grade Level Presently Taught	Non-teacher, pre-k, k,1, 2, 2 or more grades, all grades, 3, 4, 5, 6, 7, 8
Gender	Gender	Male, female
Position	Position in the school or system	Recruitment officer, CMO administrator, principal, non-teaching faculty or administrator, teacher, assistant teacher / resident teacher, split role: teacher and other non teaching job
# years at this school	# of years working at this school counting this year	
# years at charter	# of years working at charter schools in total	
# years at district schools	# of years working at district schools in total	
# years at private schools	# of years working at private schools in total	
# years teaching	total of charter, district, private years	
Race / Ethnicity	Self-identified race / ethnicity	Black, White, Caribbean, Cape Verdean, Latino/a, multi-racial, other, Asian
Classroom type	Type of classroom in which interviewee teaches	Self-contained elementary multi subjects, departmentalized core subject, specific subject non-core, ELL/ Special Ed only
Age	Age of interviewee	