National Louis University

Digital Commons@NLU

4-2020

Increasing Trust and Teacher Voice in Order to Improve Teacher Evaluation in Chicago

Tim Riff

Recommended Citation

Riff, Tim, "Increasing Trust and Teacher Voice in Order to Improve Teacher Evaluation in Chicago" (2020). *Dissertations*. 433.

https://digitalcommons.nl.edu/diss/433

INCREASING TRUST AND TEACHER VOICE IN ORDER TO IMPROVE TEACHER

EVALUATION IN CHICAGO

Tim Riff

Educational Leadership Doctoral Program

Submitted in partial fulfillment

of the requirements of

Doctor of Education

in the National College of Education at National Louis University

National College of Education

National Louis University

February, 2020

INCREASING TRUST AND TEACHER VOICE IN ORDER TO IMPROVE TEACHER

EVALUATION IN CHICAGO

Tim Riff

Educational Leadership Doctoral Program

Approved:

Maria McDaniel-Hall _Harrington Gibson_
Chair, Dissertation Committee   Program Director

Harrington Gibson _Elizabeth Minor_
Member, Dissertation Committee   Dean's Representative

_____ 2 - 14 - 20
Dean, National College of Education   Date Approved

ABSTRACT

Chicago Public Schools' teacher evaluation program includes cycles of observation and feedback as well as student test scores to determine teacher evaluation scores. Teachers have expressed support for and trust in this teacher evaluation program, yet they have misgivings about the inclusion of standardized student test scores. An examination of best practices in teacher evaluation and its implementation in Chicago, through the voices of the teachers and evaluators, identifies opportunities for improvement. By supporting teachers and administrators to engage in professional conversations, offering teachers a voice in how teacher evaluation is implemented, and eliminating the use of standardized student test scores, specifically the Value Added Model, CPS could increase trust and validity in the REACH Students teacher evaluation program and support teachers to engage in a process of continuous improvement.

PREFACE

I have served as an administrator in Chicago Public Schools for the past ten years, the last seven as a school principal. One of the primary responsibilities of the school principal is to hire and develop teachers, empowering them to be as effective as possible for the benefit of the students. Each year, I conduct over 100 teacher observations, which comprise the primary component of each teacher's evaluation score. Through this process, I have observed the stress that the teacher evaluation process causes and have witnessed firsthand the inherent inequities of the current teacher evaluation program.

I am fortunate to work in a high-performing school, enjoying the benefits that accompany working in such a school: teachers are highly rated, teaching positions are desirable and sought after, and we possess a stable and consistently growing student population as well as the significant autonomy which the district gives to high performing schools.

Before becoming a principal, I taught and worked with leaders of schools in areas of the city which did not perform nearly as well as my current school. Reflecting on the teachers that I observed in these schools, I was struck that the overall quality of instruction was comparable in these lower performing schools to that of my current, high performing school. I observed a middle school math teacher at a school in an impoverished neighborhood who employed creative engagement strategies, whose room was dynamic, colorful, and text rich, who related well to his students, and they scored consistently higher than other middle school students in the area. In my school, I have observed teachers who employ traditional teaching methods, using basal readers and rote memorization, and their students' scores are well above the norm despite the fact that

they largely ignore many of the distinguished teacher practices from the CPS Framework for Teaching.

Understanding that the development of teachers and the quality of instruction may not be wholly correlated to student results motivated me to research this issue of equity. While teacher evaluation is necessary, both for the benefit of the teacher and to comply with current state and federal laws, removing the use of standardized test scores from the teacher evaluation score addresses the inequity that exists as a result of the myriad factors which affect student test scores. Additionally, better supporting the implementation of teacher evaluation and giving teachers a voice in its implementation will build trust and ownership, subsequently increasing active teacher participation, and improving the impact of teacher evaluation on teacher practice.

TABLE OF CONTENTS

LIST OF FIGURES

Figure		Page
Figure		**Page**

LIST OF TABLES

CHAPTER ONE: INTRODUCTION

Chicago Public Schools (CPS) is the nation's fourth largest school district with 355,156 students and 21,355 teachers. A diverse and varied school district spanning the boundaries of the city, CPS is comprised of 642 schools. These schools are divided into 514 traditional schools, 118 charter schools, nine contract schools, and one SAFE school, which is a school for students who are removed from their schools for repeated or drastic behavior violations. All CPS schools are governed by the Chicago Board of Education, which, in turn, is subject to the Illinois School Code (Chicago Public Schools [CPS], 2020).

The era of accountability in education began in earnest when President George W. Bush signed the No Child Left Behind (NCLB) update to the Elementary and Secondary Education Act (ESSA) of 1965, on January 8, 2002. This update set yearly benchmarks for student achievement, called Adequate Yearly Progress (AYP). Facing sanctions if they did not meet these yearly benchmarks, schools began to place an increased emphasis on the core subjects of reading and math and test scores. Cochran-Smith and Lytle (2006) noted that the presence of this accountability measure "narrows the curriculum" and "oversimplifies the process of teacher learning and practice" (p. 670). For a time, at my current school, the teachers only taught reading and math, in addition to art and music. To this day, there remains a bias towards these core, tested subjects, and I hear from teachers that they occasionally do not have time for science or social studies, while they also complain that there is not enough time for reading and math instruction. Our students now benefit from a variety of subjects, yet teachers place an emphasis and invest the majority of their time and effort into reading and math instruction. NCLB did not address teacher evaluation specifically, but it increased the pressure for teachers and schools to continuously attain higher test scores in order to meet the AYP benchmarks. It was not until

2010 that Illinois passed a law which required the use of student growth scores in teacher evaluation.

The Performance Evaluation Reform Act (PERA), which the Illinois General Assembly passed into law on January 1, 2010 to strengthen the case for the State of Illinois to receive additional funding through the federal Race to the Top grant, requires that all schools in Illinois use student growth scores as a component of teacher evaluation as of September 1, 2013. PERA identifies three types of assessments that school districts can use to measure student growth for teacher evaluation. Assessment Type I is a standardized assessment used by multiple districts, created and scored by an independent party. Type II assessments are district-wide assessments, created and scored by the district, and used in multiple schools. The third type of assessment, Type III is aligned to the teacher's specific curriculum and can be selected or created by the teacher. PERA recommends that school districts use at least one Type I or Type II assessment and one Type III assessment; however, it does allow school districts to use two Type III assessments in lieu of using a Type I or Type II assessment. This new law changed the landscape of teacher evaluation dramatically, requiring student growth metrics to be included in teacher evaluation. In contrast to the NCLB accountability system, which used student proficiency on reading and math assessments, this new teacher evaluation system uses student growth on reading and math assessments. In order to fulfill the requirements of PERA, CPS redesigned their teacher evaluation program, which requires that student growth comprises 30% of every teacher's evaluation as of 2016. This growth is measured using a Value Added Model (VAM), which purports to control for student environmental and demographic factors such as socio-economic status, race, location, parent level of education and more.

In CPS, the teacher evaluation program is called Recognizing Educators Advancing Chicago (REACH) Students. Evaluators conduct clinical classroom observations of teachers and rate the teacher practice based on the CPS Framework for Teaching. These ratings comprise the teacher practice score. In addition to the teacher practice score, each teacher's summative evaluation includes student growth measures. All teachers who do not teach reading or math to students in grades 3-8 receive scores based on two performance tasks administered at the beginning and the end of each year. These two assessments are intended to measure growth of student mastery of specific standards, and for these teachers the growth that their students demonstrate constitutes their student growth score. All teachers who teach reading and/or math to students in grades 3-8 also receive a VAM score based on their students' growth on the Northwest Evaluation Association (NWEA) reading and/or math assessments. The district creates individual goals for each student based on their proficiency on the NWEA the previous spring, and teachers are evaluated by how much growth each student demonstrates the following spring and if they meet their growth goal. This growth, for each student, is used to calculate a VAM score, which then comprises 30% of the teacher's summative evaluation score.

Teachers have mixed feelings about this new evaluation system, in particular the use of student growth on test scores (Jiang, Sporte, & Luppescu, 2014). Teachers resisted wholesale the use of student assessments to measure their effectiveness. Jiang et al. (2014) report that "Only half of teachers said the assessments used to measure student growth are fair assessments of their students' learning, regardless of whether those assessments are individual value-added, school-wide value-added, or performance tasks" (p. 7). This resistance demonstrates a fundamental distrust of the use of student test scores in teacher evaluation, regardless of which scores are used and how these scores are included. In particular, special education teachers voiced deep concern

that this was an unfair way to measure their effectiveness with students who require additional

support and generally score lower on assessments than their non-disabled peers

(Jiang, Sporte, & Luppescu, 2014).

The teacher practice score consists of the results of a combination of formal and informal

observations. The following are general guidelines, but exceptions do exist. Evaluators observe

and rate tenured teachers through one informal and two formal observations every two years.

Evaluators observe and rate probationary teachers who are in their first three years, through one

informal observation and two formal observations each year. The informal observations are 10-

15 minutes in length and do not require a pre-observation or post-observation conference. The

formal observation cycle consists of a pre-observation conference, the observation, and a post

observation conference. The observer rates the teacher on individual components of the CPS

Framework for Teaching (Included in Appendix A), which is based on the Danielson

Framework.

Purpose

The purpose of this analysis is to assess teacher evaluation in Chicago in order to identify

opportunities for improvement, both through structural changes and implementation adjustments.

This analysis will reveal whether and to what extent inequities exist in the current REACH

Students teacher evaluation program as well as opportunities for improving the current system in

order to provide more support and improved opportunities for teachers and outcomes for

students. Marzano (2012) noted the three primary components of a teacher evaluation system

that focuses on teacher development.

> 1. The System Is Comprehensive and Specific
> 2. The System Includes a Developmental Scale

3.The System Acknowledges and Rewards Growth (pp. 16-18)

This analysis will show that the design of the REACH Students teacher evaluation system meets these criteria, and was designed to support teacher development in order to improve overall instructional impact. The manner in which a teacher evaluation program is implemented, however, can have unintended effects. In a study of public schools in Connecticut, Donaldson et al. (2016) found that some evaluation reforms "disproportionately benefitted higher-performing and lower poverty schools" (p. 196), and they emphasized the importance of "thinking carefully about how to implement new evaluation policies so that schools enrolling low-income students, students of color, and lower-performing students fully benefit from such policies" (p. 196). Careful thought about implementation is a central theme to this research. School leaders in higher poverty, lower scoring schools cited the additional compliance requirements as a barrier to effective implementation of the program. Sporte, Stevens, Healey, Jiang, & Hart (2013) found this inequity to be present in the pilot implementation of the new teacher evaluation system in Chicago, where "schools that were already high-performing at the start of the pilot were better able to use the observation process to improve student outcomes than were schools that were struggling before the project began" (p. 31).

Using existing research in order to identify best practices in teacher evaluation and adult learning theory, this analysis will identify specific recommendations for improved implementation of the current teacher evaluation program in order to increase trust, foster ownership, and engage teachers in a process of continuous improvement.

Rationale

I am the principal at Oriole Park School, a neighborhood school on the northwest side of Chicago. The neighborhood is heavily populated by city workers: police, firefighters, teachers, and city workers. Oriole Park School enjoys a high level of parent engagement, and parents hold the school accountable for the educational experience and outcomes that we provide.

I evaluate teachers using REACH Students framework and am accountable for my school's growth on the NWEA math and reading assessments. Each year, I conduct the required teacher observations, as well as frequent informal classroom walkthroughs, and I work with teachers to analyze their students growth on the NWEA reading and math assessments twice per year, during the winter and spring administrations of the assessments. Our students consistently perform well on the assessments, and about 60% of students meet their growth goals every year, which ranks the school above the 90th percentile for growth. There is, of course, some variance from grade to grade and class to class. Certain teachers have consistently high results, and other teachers have consistently lower results. I have observed a notable correlation between these results and the teacher practice scores on the REACH framework, but this correlation is not present in all cases.

This potential disconnect between teacher practice scores and student assessment results has led me to wonder if the CPS Framework for Teaching is a valid instrument to measure quality of instruction in the manner that it is currently implemented and what inequities might exist in the current evaluation program. With additional support, administrators could be both more consistent and aligned in the manner in which they rate teachers and more effective in professional conversations to build trust and engage teachers in continuous improvement. By participating in the decision-making process, teachers could become more invested in, recognize the value of, and participate fully in the REACH Students process to facilitate self-reflection and

continuous improvement. The REACH Students teacher evaluation program is designed for teacher development, but the inclusion of standardized scores, specifically the VAM, undermines teacher trust in the system in addition to incentivizing working in the schools with the lowest poverty level, pushing more experienced and desirable teachers away from the students with the greatest need. Perhaps the removal of the VAM from teacher evaluation, which is a barrier to trust, would increase equity of access to high quality teachers and outcomes for all students.

Goals

Jiang et al. (2014) found that a more comprehensive teacher evaluation program, one which focuses on teacher development (similar to the REACH framework), resulted in higher student achievement in reading. The implementation of the REACH Students teacher evaluation program in Chicago was successful in promoting more professional conversations between teachers and administrators and creating a shared vision for high quality teaching. However, inadequate support for professional learning, reflection, and self-assessment, as well as the continued use of standardized test scores in teacher evaluation undermine teacher trust in the REACH Students teacher evaluation program. This lack of trust and investment in the program limits the effectiveness of the program itself. I have experienced this distrust firsthand from teachers who feel that student test scores are not a fair assessment of their instructional impact. These teachers approach the evaluation process with skepticism and do not fully engage in the observation cycle with the goal of development; they view the process as a compliance exercise which must be tolerated. This disengagement is exacerbated by the lack of high-quality professional development opportunities, deepening teachers' unwillingness or inability to engage in reflection and self-assessment.

This analysis seeks to leverage teacher and administrator perceptions of the benefits, challenges, and effectiveness of the REACH Students teacher evaluation program in order to make suggestions for improvement. Furthermore, this analysis will examine what opportunities exist to build more ownership of teachers in the program. Finally, the research will examine the impact of the inclusion of the VAM metric and its effect on students' equitable access to high quality teachers.

Research Questions

If teachers become more invested in the process, will participation in it improve their teaching practice and outcomes for their students? Understanding the impact of valid, normed observations, embedded professional development, teacher investment in the process, and the removal of the VAM metric from teacher evaluation will provide a rationale for making adjustments to the manner in which teacher evaluation is implemented in CPS. This understanding will also provide additional context when evaluating the quality and validity of teacher evaluation systems both in and outside of Chicago. By examining past and current research and best practices in teacher evaluation, opportunities for improvement of CPS's REACH Students teacher evaluation program emerge. Specifically, this analysis seeks to examine the following questions:

1. Whether and to what extent do opportunities exist to increase teacher participation and ownership in the implementation of teacher evaluation in Chicago?
2. Whether and to what extent do teacher and administrator perceptions of the REACH Students teacher evaluation program inform opportunities for improvement?

3. Whether and to what extent does the use of the Value Added Model in teacher evaluation create inequity for Chicago Public Schools' students and teachers?

By answering these questions using current research and feedback from teachers and administrators, this analysis will provide clear steps to improve the REACH Students teacher evaluation program and its impact on instruction and student results in Chicago. The existing research on teacher evaluation provides useful context and illuminates opportunities for improvement.

CHAPTER TWO: REVIEW OF LITERATURE

As a result of the 2009 Race to the Top Initiative, and specifically the Performance Evaluation Reform Act (PERA) in Illinois, which mandated teacher evaluation reform to include student growth as well as teacher practice in teacher evaluation, school districts were required to redesign their teacher evaluation systems. Chicago Public Schools (CPS) implemented their current teacher evaluation program, REACH Students, in 2012, which requires that the teacher evaluation score includes a combination of teacher observation as well as student growth measures. In the ensuing eight years, researchers have studied this new type of multiple measure teacher evaluation system and its implementation in school districts across the country (Darling-Hammond, 2012, Fullan, 2011, Kane, Taylor, Tyler, & Wooten, 2011, Marzano, 2012, Sartain et al., 2011, White, Cowhy, Stevens, & Sporte, 2012). With the goal of improving teacher practice, and by extension student learning outcomes, this body of research can identify opportunities for improvement of teacher evaluation programs.

In addition to examining best practice opportunities for improvement, increasing teacher trust in the evaluation program will provide teachers with a more comprehensive and responsive evaluation program. Removing the VAM metric as a component of teacher evaluation will eliminate inherent inequities, strengthen the overall system, and increase teacher trust and buy-in. While the use of student growth measures is required in the current law, the use of local and curriculum-based assessments, in lieu of the VAM which may not be aligned to the curriculum (Marzano & Toth, 2013), in teacher evaluation will remove this barrier to teacher trust in the program. Leveraging adult learning theory as well as implementing these best practices will further increase teacher trust and active engagement in the teacher evaluation and its continuous improvement process. The literature reviewed for this program evaluation will examine four

areas of focus regarding teacher evaluation: the evolution of high-stakes teacher evaluation and the use of the Value Added Model, evolving teacher evaluation in Chicago, fostering teacher trust throughout the evaluation process, and using adult learning theory to maximize teacher instructional and pedagogical capacity.

The Evolution of High Stakes Teacher Evaluation and the Use of the Value Added Model

In the era of No Child Left Behind, schools were held accountable to measures of student proficiency and whether these metrics met the benchmarks of AYP targets of increased proficiency year after year. In 2009, the national movement for teacher accountability and revision of teacher evaluation began with the Race to the Top initiative which mandated the use of student growth scores in teacher evaluation. As of 2009, only four states required teacher evaluations, and no states used evaluations to make tenure or dismissal decisions (National Council on Teacher Quality [NCTQ], 2019). At that time, no states used student growth scores to measure teacher effectiveness.

The 2009 Race to the Top prompted Illinois' PERA law and caused states across the country to modify the manner in which they evaluate teachers in order to include a student growth metric, which brought the proliferation of the VAM. The VAM was originally created to compare teachers to other teachers and rate them hierarchically against one another based on the calculated effect the teacher had on their students' growth. Paige et al. (2019) note that the primary function of the VAM is to "statistically control for outside variables, including students' prior test performance, and student level background variables (e.g., whether students are eligible for free-and-reduced lunches)" (p. 531). William Sanders created the VAM model in order to fulfill Tennessee Governor Lamar Alexander's desire to create a merit pay system for teachers

(Paige et al., 2019). Conditioning federal funding to use of the VAM, Race to the Top effectively forced districts across the country to employ this model.

The use of the VAM for teacher evaluation has prompted much debate, with proponents (Hanushek, 2011, Corcoran, 2010, Corcoran & Goldhaber, 2013, Fulmer, 2012), and opponents (Darling-Hammond et al., 2012, Rothstein, 2010, Briggs & Domingue, 2011) voicing their support and concern. In the ten years since Race to the Top, the use of VAM has prompted legal action by teachers and teacher unions. So far, the courts have only ruled against VAM on procedural grounds, while some courts have expressed concern about the wisdom of employing this model (Paige et al., 2019).

The ability to distinguish between an effective teacher and an ineffective teacher is in the public interest. Certainly, everyone benefits when schools have as many effective teachers as possible. Fulmer (2012) contends that while the VAM may not be a perfect measure, using the VAM makes it "possible to distinguish very low-performing from very high-performing teachers or schools" (p. 365). This suggests that there is some value to the VAM, to sort effective schools and teachers from ineffective ones, yet it may lack the precision required for accurate evaluation of a specific teacher in a specific year.

Validity is a primary complaint about the use of VAMs. Darling-Hammond (2012) noted:

> …of teachers who scored in the bottom 20% of rankings in one year, only 20% to 30% had similar ratings the next year, while 25% to 45% of these teachers moved to the top part of the distribution, scoring well above average. (p. 9)

A majority of the teachers receive different ratings from one year to the next under the VAM. This lack of reliability exposes doubts about validity of the VAM. These doubts are even more troubling when considering that decisions about teacher tenure and retention are based on a teacher's VAM score. One would expect that a teacher's impact is largely consistent year over

year, yet their VAM score can vary dramatically. There are other examples that call into question

the validity of the VAM. Darling-Hammond (2012) notes that "Teacher effectiveness also varies

significantly when different statistical methods are used" (p. 9). For example, Briggs &

Domingue (2011) recalculated the teachers' scores that were made public in Los Angeles and

found that 40%-55% of them would receive different scores using a different model.

To underscore the potential inequity in the VAM, Darling-Hammond (2012) noted that

"teachers are advantaged or disadvantaged based on the students they teach" (p. 10). Rothstein

(2010) found that VAM scores were weakly related to long term education outcomes. Darling-

Hammond (2012) found that even after controlling for environmental factors, there was a

correlation between demographic factors and student outcomes.

> After controlling for prior student test scores and student
> characteristics, the study still found significant correlations
> between teacher ratings and students' race/ethnicity, income,
> language background, and parent education. (p. 12)

With the presence of myriad factors - which were supposed to be controlled - continuing to

influence final VAM scores, confidence that a teacher's VAM score is an accurate measure of

their effectiveness is difficult to achieve. This doubt about the validity of the VAM undermines

the validity of teacher evaluation scores that include VAM as the student growth metric. In fact,

in 2015, the American Educational Research Association released a statement that cautioned

against the use of VAM for high-stakes decisions regarding educators (American Educational

Research Association [AERA], 2015, November 11).

President Obama signed the Every Student Succeeds Act (ESSA) into law on December

10, 2015. This new law was meant to replace NCLB, and it focused on high standards, preschool

for all students, and local control to encourage innovation. To this end, ESSA allows for local

control of teacher evaluation, yet some districts continue to employ the VAM - although fewer

and fewer of them do. According to Close et al. (2019) "Our findings suggest that while the use of VAMs to hold teachers accountable for their levels of effectiveness is still transpiring, VAMs are losing traction among states" (p. 23). In fact, only 15 states currently use the Value Added Model, while 23 states that formerly employed this model either no longer require it or allow for local control. Illinois allows for local control, yet some districts, including CPS, continue to employ the VAM (National Council on Teacher Quality [NCTQ], 2019).

A growing number of states encourage the use of student learning outcomes which are, "objectives set by teachers, sometimes in conjunction with teachers' supervisors and/or students, to measure students' growth" (Close et al., 2019, p. 25). This local control over measures of student growth is replacing the VAM in many states, and the use of these "standards have become the basis for assessments of teaching that produce ratings that are much more stable than value-added measures" (Darling-Hammond, 2013, p. 13). Additionally, Darling-Hammond (2013) notes that these measures are "predictive of student learning gains and productive for teacher learning" (p. 14). A growing trend away from the VAM and towards standards-based assessments could lead to a more accurate student growth measure for teacher evaluation.

Evolving Teacher Evaluation in Chicago

According to Marzano (2012) and Darling-Hammond (2012), evaluation systems must be comprehensive and focused on teacher growth in order to be effective. While student achievement is a necessary component of teacher evaluation, engaging in collegial inquiry and a continuous improvement process is the necessary primary objective. Linda Darling-Hammond (2012) stated that effective teacher evaluation systems include: multiple classroom observations, trained administrators, rubrics, mentors, collaboration with other teachers, and professional

development. Subsequently, Darling-Hammond (2013) provided a framework for effective

teacher evaluation systems, including the following seven criteria:

1. Based on professional learning standards
2. Include multifaceted evidence of teacher practice
3. Knowledgeable and trained evaluators
4. Use of feedback, tied to professional learning opportunities
5. Encourage teacher collaboration
6. Expert teachers are a part of the review and assistance process
7. Panels of teachers and administrators should oversee the process (p. 8)

These seven criteria constitute a framework to evaluate the implementation of the teacher

evaluation program in Chicago Public Schools. The structure of the REACH Students framework

includes evaluation based on professional standards, multiple measures of teacher practice,

trained evaluators, and the use of feedback, but there is little opportunity for teacher

collaboration or participation in implementing the evaluation system. The use of professional

standards fosters an important common vision of high quality instruction, and the inclusion of

multiple measures in the evaluation system can help teachers and administrators to identify

specific skills for improvement (Kane, Taylor, Tyler, & Wooten, 2011).

Before the implementation of the current REACH Students framework for teacher

evaluation, principals simply used a checklist of teacher behaviors at the end of the year to give

feedback to teachers on their performance in three areas: Instruction, School Environment, and

Professional Standards. With regard to instruction, there were nine criteria ranging from writing

lesson plans, creating bulletin boards to applying learning theory. The school environment

section had seven criteria, such as taking attendance, communicating expectations, and

encouraging growth. The Professional Standards section referred only to three criteria: teacher

appearance, diction, and judgement. This checklist lacked depth and specificity, and according to

Sartain et al. (2011), both teachers and principals reported that the use of the checklist provided

little guidance to teachers as to how to improve their practice. The checklist had been in place for 30 years, and in 2008, CPS CEO Arne Duncan began the transformation of teacher evaluation in Chicago by piloting an evidence-based teacher evaluation program in 44 schools, scaled up to 100 schools the following year (Sartain et al., 2011). Chicago was ahead of most other school districts in this effort to improve teacher evaluation by focusing on teacher reflection and improvement. Furthermore, Sartain et al. (2011) report that overall teacher evaluation scores appear to have been inflated using this checklists system, wherein a small percentage of teachers were designated as "Unsatisfactory" (0.3%), and a majority of teachers received the highest rating, "meaning 93 percent of the district's teachers were Excellent or Superior according to the checklist evaluation system" (p. 4). (The full CSP Teacher Evaluation Checklist can be found in Appendix B.)

In 2009, the $4.35 billion in federal Race to the Top initiative pushed this nascent teacher accountability movement farther with the intent "to remove any existing legal barriers to linking student achievement data to teacher evaluations" (Sartain et al., 2011, p. 4). As a result, Chicago implemented the REACH Students teacher evaluation program in order to comply with the state of Illinois' PERA law requiring the use of student growth scores in teacher evaluation. Additionally, the CPS Framework for Teaching offered a more comprehensive evaluation process by requiring professional conversations before and after all observations, which created opportunities for teachers and evaluators to engage in collegial inquiry, discuss observation feedback, and identify specific opportunities for growth.

In 2012, CPS created the Framework for Teaching (Appendix A), based on the Danielson Framework. This rubric provides a shared vision for effective teaching practice as well as the basis for professional coaching conversations between the teacher and evaluator. The Framework

for teaching is significantly more comprehensive than the one page checklist, identifying four

Domains of teacher practice: Planning and Preparation, Classroom Environment, Instruction, and

Professional Responsibilities. Each of these Domains contains 4-5 components or teacher

behaviors, and in contrast to the prior checklist, the Framework for Teaching offers a description

of what each component practice looks like at each of the four performance levels:

Unsatisfactory, Basic, Proficient, and Distinguished. With these descriptions as a guide, a teacher

and administrator can engage in a discussion which compares the current teacher practices with

the descriptions in the rubric. As opposed to simply receiving a rating, as they did with the

checklist, teachers are required to engage in professional conversation about their practice and

identify opportunities for improvement.

Implementing teacher evaluation effectively not only requires educating the stakeholders

on the evaluation process and content of the framework, it also requires addressing mindset and

preconceived notions about high quality teaching. Sartain et al. (2011) noted that this new

evidence-based system in which administrators collect evidence and engage in productive and

effective professional conversations regarding continuous improvement of teacher practice will

require additional training and support for teachers and administrators in order to be effective

(Sartain et el., 2011).

Charlotte Danielson, whose framework for teaching was used as the basis for Chicago's

REACH Students teacher evaluation program, recognized the importance of using such a tool to

develop a shared vision of high quality instruction. Danielson (2012) stated that a teacher

evaluation program which "engages teachers in reflection and self-assessment—yields benefits

far beyond the important goal of quality assurance" (p. 27), stating that "It's all about the

conversation" (Danielson, 2012, p. 25). Danielson (2016) expressed misgivings about the use of

her framework for teacher accountability, stating that she was "deeply troubled by the transformation of teaching from a complex profession requiring nuanced judgment to the performance of certain behaviors that can be ticked off on a checklist" (p. 22). While her framework identified specific teacher practices which were deemed effective, she did not support the use of the rubric as the final word on effective instructional practice or as a roadmap to follow. Recognizing that professional learning is the goal of any teacher evaluation program, Danielson (2016) emphasized the importance of trust in the process, allowing teachers to be vulnerable and take risks. When a teacher is observed a mere handful of times in a year, and each observation comprises a portion of their final evaluation score, there is little room for teachers to take risks or try something new and innovative. Every moment and action counts.

In order to engage in a reflective process, one that teachers trust, they must have an understanding that their instructional practice is fluid, and that the overall quality of their instruction will benefit from engagement in a continuous improvement process (Sartain et el., 2011).

Fostering Teacher Trust throughout the Evaluation Process

Implementation of any teacher evaluation program without significant buy-in from the teachers will not result in improved educational outcomes (Fullan, 2011, White, Cowhy, Stevens, & Sporte, 2012). In order to build trust, teachers must have input, and their feedback must be incorporated into the implementation process (White et al., 2012). The manner in which a school district implements a teacher evaluation program can have an effect on teacher engagement. Moskal et al. (2016) found that "improving practical implementation can increase staff engagement" (p. 298). Teachers had a voice in the original creation and implementation of the PERA law, as teachers, administrators, union leaders and other stakeholders comprised the

Performance Evaluation Advisory Council (PEAC), which was charged with providing input from educators to the Illinois State Board of Education (ISBE) and monitoring PERA development and implementation. On a local level, increasing teacher participation in the implementation process and incorporating their feedback when making decisions about implementation, can increase the level of teacher engagement in the REACH Students teacher evaluation program.

In a longitudinal study of schools in Chicago, Byrk and Schneider (2003) found that in schools with high levels of trust, teachers were more likely to take risks, try new approaches and collaborate with parents. These schools were more likely to attain increased academic gains, "8 percent in reading and 20 percent in mathematics in a five-year period" (Byrk & Schneider, 2002, p. 43). On the other hand, schools with lower levels of trust did not attain academic gains. With high levels of trust, teachers are more comfortable engaging in honest conversations with one another about their challenges and concerns, as the risks associated with struggles are diminished in a trusting environment. The relationship between the teacher and the administrator is essential for creating and preserving this trust. A sample of teachers from the top quarter of Chicago elementary schools reported that they were "encouraged to stretch and grow" and are "continually learning" (Byrk & Schneider, 2002, p. 34). This growth mindset must be fostered and encouraged both with teachers and administrators. Administrators who cultivate trust and create the conditions for innovation and risk taking allows teachers to more fully engage in the continuous improvement process (Danielson, 2016).

Trust is essential when teachers are evaluated by the same administrator who is responsible for supporting their professional growth. (Sporte, Stevens, Healey, Jiang, & Hart, 2013). In Chicago, the first major barrier to trust was the inclusion of student growth scores in

teacher evaluation scores, against the will of the teachers, undermining teacher trust in the evaluation program (Sporte, Stevens, Healey, Jiang, & Hart, 2013). Darling-Hammond (2013) suggests that appropriate use of student work samples and portfolios, related to the curriculum, is a better indicator of teacher effectiveness than student assessment results, and it is problematic for districts to utilize value-added methods to draw conclusions about individual teacher effectiveness, based on student test scores.

Hazi (2014) pointed out that the use of student test data and evaluation programs designed by non-educators, such as elected officials and consultants, could result in legal challenges. For instance, "minority teachers may challenge a dismissal based on discrimination" and tenured teachers who suffer "adverse employment decisions, may have to challenge local evaluation policy and state statutes as unconstitutional" (p. 138). While there were no legal challenges to the implementation of the REACH Students evaluation program in Chicago, the use of student test scores in teacher evaluation was a major point of contention in the teacher contract negotiations with CPS in 2012, which resulted in a teacher strike. Surprisingly, one year later, teachers reported that they generally believed their administrator was fair and accurate with regard to their evaluation (Sporte et al., 2013).

`In a further challenge to teacher trust of the evaluation program in Chicago, schools with differing demographics also experience marked differences in teacher evaluation scores, where, on average, teachers in high-poverty schools receive lower observation scores than teachers in low-poverty schools (Jiang, & Sporte, 2016). Teachers with the lowest VAM and observation scores are overrepresented in schools that serve the most disadvantaged students, while teachers with the highest ratings on observations are vastly underrepresented in highest-poverty schools (Jiang, & Sporte, 2016). This uneven distribution of highly rated teachers exposes inequities in

the broader school system. In addition to trusting the validity of evaluation scores, stakeholders must feel that the program is equitable in order to fully place their trust in it.

I have witnessed this issue from both sides. Working with principals in schools with high levels of poverty, I have witnessed schools struggling to meet students' basic needs, a necessary prerequisite to addressing the work of raising scores. I have worked with many dedicated, caring, and capable teachers in these schools whose evaluation scores suffer due to the environmental barriers that their students face (Darling-Hammond et al., 2012, Rothstein, 2010). On the other hand, as a principal of a high-performing, low poverty school, I have worked with teachers who enjoy an automatic boost to their evaluation score every year, based on the overall performance of the school. This inequity is a result of the inclusion of the VAM scores in evaluation, and it would need to be addressed in order to earn the complete trust of teachers in the evaluation program.

To combat these barriers to trust, district and school leaders must find ways to build trust within schools and across the district and ensure that all schools experience the positive benefits of the REACH Students program (Sporte, et al., 2013). Within the schools, the onus rests on the administrator to build trust and a collective purpose. Bryk (1996) recognizes this responsibility yet notes that the school district bears the responsibility to create the conditions for such trust, saying that the school district "can create conditions more conducive to its development and nurture its sustenance" (p. 34). One way to foster such trust is through comprehensive professional learning and support, which will be examined in the next section. A teacher evaluation program which is valid and well supported will be more effective. Teachers and students who enjoy the benefits of a comprehensive and effective teacher evaluation program will see its value and place more trust in it.

One essential strategy for increasing teacher trust in the evaluation program would be to give teachers a "degree of choice in how and by whom they will be evaluated" (Phillips, Balan, & Manko, 2014, p. 12). This can be accomplished on the school level, through administrator collaboration and flexibility around teacher observations, and it can be accomplished at the district level by including teachers in policy decisions about evaluation implementation. In addition to increasing teacher trust in evaluation, empowering teachers with increased influence over school policy will improve overall school culture and improve student achievement (Sebastian, Allensworth, & Huang, 2016).

Creating the conditions for trust in teacher evaluation is merely the first step, and while this trust is essential to create teacher engagement, the ultimate benefit of teacher evaluation is teacher development. This aspect of teacher evaluation is often taken for granted. CPS gives little guidance or support for collegial inquiry and adult learning with regard to teacher evaluation. Adult learning is the primary goal of teacher development, and it must be undertaken properly in order to maximize the effectiveness of the teacher evaluation program.

Utilizing Adult Learning Theory to Build Teacher Instructional and Pedagogical Capacity

Districts must ensure that all evaluators and teachers have access to high quality district-wide professional development in order to combat differences and potential inequities between schools in their capacity to support professional growth (Sporte, Stevens, Healey, Jiang, & Hart, 2013). In order to provide high quality professional development for all evaluators and teachers, school leaders must understand that adults possess varied learning styles and developmental needs (Drago-Severson, 2006). Understanding that all adults approach learning in their own distinct manner will help school districts to design learning experiences that meet the needs of all evaluators and teachers (Drago-Severson, 2009). Creating professional learning opportunities

which include a measure of choice will address this need of adult learners. Warren Little (2012)

identified the following common characteristics of successful, learning-centered schools.

> …close involvement with students and their work; shared
> responsibility for student progress; access to new knowledge about
> learning and teaching; sensibly organized time; access to the
> expertise of colleagues inside and outside the school; focused and
> timely feedback on individual performance and on aspects of
> classroom or school practice; and an overall ethos in which teacher
> learning is valued and professional community cultivated. (p. 22)

These practices of successful schools provide a framework for creating effective adult learning

and teacher development opportunities. Of particular importance to teacher evaluation is

feedback on practice and access to the expertise of colleagues inside and outside of the school.

The quality and manner of the feedback is related specifically to professional support with

collegial conversations, and access to the expertise of colleagues must be considered when

structuring the implementation of a teacher evaluation program. The components of shared

responsibility and ethos of valuing teacher learning refer back to the underlying trust and

engagement of teachers in the evaluation program.

In addition to employing best practices in adult learning theory, districts must provide

ongoing professional development. Reid (2018) studied how public school districts train their

staff to employ their teacher evaluation program and found that in order to best support

evaluators, they should "receive initial and ongoing training that goes beyond training principals

to navigate the logistics of the teacher evaluation process" (p. 252). This support for

understanding the content of the evaluation program is merely a first step. The ongoing training

and access to the expertise of others is essential for the continued effectiveness of the evaluation

program.

Desimone (2017) identifies five main features of effective professional learning.

1. Content focus
2. Active learning
3. Coherence
4. Duration
5. Collective participation (pp. 4-5)

Content focus refers to the professional development focusing on what the students learn and how they learn it. This focus is directly tied to the CPS Framework for Teaching, Domain 1: Planning and Preparation. As a part of the required pre-observation conference, the administrator and teacher discuss the content and pedagogy planned for the lesson. Active learning refers to opportunities for teachers to get involved by observing each other, analyzing lessons, and planning implementation. Active Learning is an opportunity for growth in the current program. There is a lack of active learning in the current REACH Students model. Coherence refers to the overall plan for professional learning and how the pieces fit together. Duration refers to the number and frequency of opportunities for professional growth. Collective participation is akin to active learning, but again, it refers to opportunities for teachers to learn in collaboration with their peers, which is not the current practice in CPS.

Instilling collaboration as a core component of the teacher evaluation process will support high quality and consistent adult learning, including: teacher teaming, teacher leadership, collegial inquiry and mentoring relationships (Drago-Severson, 2006, Drago-Severson, 2009). Accompanied by professional training and an explicit, common purpose, peer coaching can be an effective strategy to improve student outcomes (Skinner, & Welch, 1996). Burgess, Rawal, and Taylor (2019) found that students at schools in which teachers engage in peer observation scored higher than their counterparts in schools where the teachers did not engage in peer observation. These student results, as a measure of teacher effectiveness, suggest that the act of engaging in peer observation improved the effectiveness of the teacher's instruction. Peer observation offers

an opportunity for school districts to embed professional learning into the teachers' daily work. In contrast to when it is an additional component of a teacher's responsibilities, professional learning is most effective when it is ongoing and embedded in the teacher's practice (Darling-Hammond, 2013).

Conclusion

Considering existing research and recommendations of experts, CPS's REACH Students teacher evaluation program meets several criteria for an effective teacher evaluation program, including a set of professional standards and multiple measures of student learning. However, significant opportunities exist in the implementation of this teacher evaluation program to increase teacher trust by eliminating the use of VAM as a student growth metric, by including teacher voice, and by leveraging adult learning theory in order to create a comprehensive program of invested teachers and expert evaluators.

In addition to considering best practices and current research in teacher evaluation, listening to the voices of the practitioners, teachers, and their evaluators, provides context and additional guidance for recommendations for improvement of the REACH Students teacher evaluation program.

CHAPTER THREE: METHODOLOGY

Research Design Overview

Program evaluation as described by Patton (2008) includes an evaluation of a program's

intended outcomes, unintended outcomes, and implications (p. 5). Simply stated, Patton (2008)

describes evaluations as processes to answer three questions.

> What?
> So What?
> Now What? (p. 5)

This research is a document analysis of the Chicago Consortium for School Research

(CCSR) published reports regarding teacher evaluation, including an analysis of responses of

teachers and administrators as well as a correlation of teacher evaluation scores and other school

factors. According to Bowen (2009), "Document analysis is a systematic procedure for

reviewing or evaluating documents" (p. 27), which has several advantages, including: efficiency,

availability, cost-effectiveness, and coverage, yet it may be undermined by insufficient detail or

biased subjectivity (Bowen, 2009). By applying both a quantitative and qualitative component,

mixed methods, to the analysis, this report benefits from the advantages of both methods. Patton

(2008) notes that, "A balanced approach to methods has become commonplace with increasing

emphasis on using mixed methods whenever possible to overcome the inherent and inevitable

weaknesses and limitations of any single method" (p. 461).

The CCSR has researched and published a series of studies of teacher evaluation in

Chicago. With access to both teacher evaluation data and student assessment results, the CCSR

has been able to identify trends and correlations between these two data points. Additionally, the

Consortium has analyzed specific questions on its My Voice, My School survey to lend more

context to this analysis through teacher and administrator responses to questions about the

REACH Students teacher evaluation program. An examination of teacher and administrator responses on the annual My Voice, My School survey, to questions regarding teacher evaluation in Chicago, provides context and insight into its implementation. Analyzing these responses suggests further actions to improve the REACH Students teacher evaluation program.

By extracting specific data points contained in the CCSR reports from 2012-2015, trends in results, attitudes, and conditions emerge that demonstrate both areas of success and opportunities for improvement. While the purpose of the Consortium research was not limited strictly to the components of teacher evaluation, I have selectively included those data sets and survey responses that related directly to perceptions and conditions of teacher evaluation in Chicago. The Consortium has not yet completed its analysis and publication of the subsequent years of data, so the results below are limited to the years 2012-2015.

Participants

There are no participants in this study. All of the data contained herein has been published and is currently publicly available in the CCSR reports, which are available for download from the CCSR website: www.consortium.uchicago.edu. I selected these specific reports for examination due to their singular access to CPS teachers, principals, demographic information, and test results. As a result of the partnership between CPS and the University of Chicago these reports contain data which is not publicly available elsewhere. The specificity and depth of the data set provides a compelling insight into the methods and mindsets involved in teacher evaluation in Chicago.

The Consortium has published several works that examine teacher evaluation implementation in Chicago. I have extracted the relevant data from the CCSR reports that assesses teacher evaluation, how stakeholders think about it, and the effect that it has on teacher

practice as well as the intended and unintended effects of the design of the teacher evaluation program and its implementation. Using these data, I will identify trends and opportunities to achieve improved outcomes for teachers and students with adjustments to the implementation of the REACH Students program.

Data Gathering Techniques

Beginning in 2010, the CCSR published a report each year on the state of teacher evaluation in Chicago. These reports use existing data regarding best practices in teacher evaluation as well as aggregate stakeholder responses about their perception of the evaluation program, its implementation, and its effectiveness. Gathering the data contained in these reports and combining the data sets across years creates a longitudinal data set which tracks implementation and perception year over year. These data comprise the data in this analysis.

In addition to providing insight into the perceptions of teacher evaluation, these data identify the differences between schools with regard to school and teacher characteristics as well as provide information specifically related to equity between schools with regard to access to high quality teachers and instruction.

Ethical Considerations

All of the data contained herein consists of publicly available and aggregate data. As such, there are no ethical considerations with regard to the publication of the data set. The participants responded to questions on surveys with the understanding that the responses would be published in the aggregate form, and by responding they agreed to such publication.

The aggregate student scores have already been published, and in no way do they indicate a specific school, teacher, or student. While the data indicate members of a subgroup of the data set by using some demographic information, such as poverty level of the students in a school,

there is no specific information about which schools are included in this group, and the subgroups themselves contain hundreds of schools.

Data Analysis Techniques

After examining the data contained in all of the CCSR reports, I was able to identify specific trends, which in turn suggested opportunities for improvement of the REACH Students program. If teachers report, year over year, that they feel that their administrator is capable of assessing their instructional practice, that indicates a strength in the program. On the other hand, if teachers express concern over the use of standardized test scores, that indicates an opportunity for improvement. By examining responses to the same question, year after year, one can see if strengths and opportunities persist or if they change.

By aggregating the data in this manner, a comprehensive picture of the implementation of teacher evaluation in Chicago emerges. I chose to include those responses and trends which referred directly to the content of teacher evaluation as well as the process of implementation and how it has evolved since the implementation of the REACH Students teacher evaluation program.

Conclusion

The CCSR has provided a wealth of information regarding teacher evaluation in Chicago. This analysis employs the data from 2012-2015, identifying trends and opportunities for improvement, both with the structure of the teacher evaluation program and the manner it is implemented. With a focus on strengthening the program through trust and engagement as well as implementing structural changes, this analysis will provide specific strategies and actions which will lead to a more comprehensive and effective teacher evaluation program in Chicago. Following is a picture of the current state of the REACH Students teacher evaluation program.

CHAPTER FOUR: RESULTS

Assessing the 4C's (As Is)

Wagner's 4C's provide a framework to create a holistic picture of the current state of teacher evaluation in CPS by examining the existing Context, Competencies, Conditions, and Culture. Wagner's 4C's compel leaders to "think systematically about the challenges and goals of change" (Wagner et al., 2006, p, 98). Employing this framework brings to light the existing strengths and opportunities for growth and provides a basis upon which to envision an ideal future state. Wagner (2006) asserts that systems are designed to "produce the results you're getting" (p. 106), so being deliberate about planning for these four components will result in the desired outcome. To better understand this framework, one must start by defining each of the four components.

Wagner (2006) defines Context as the "skill demands that all students must meet to succeed as provider, learners, and citizens, as well as the aspirations, needs, and concerns of the families and communities a school serves" (p.104). The broad context in which a school operates can "help inform and shape the work we do to transform the culture, conditions, and competencies" of the system (Wagner et al., 2012, p. 104). As such, context is paramount, yet it is difficult to control as it involves so many stakeholders with varied desires, motivations, and expectations.

Competencies, according to Wagner (2006), are the "repertoire of skills and knowledge that influence student learning" (p. 99). In the case of a school system, this refers to the competencies of the district, school leaders, and teachers to influence student learning. These competencies must be carefully cultivated in order to maximize the impact of the school program.

Wagner defines Conditions as "the external architecture surrounding student learning, and tangible arrangements of time, space, and resources" (p. 101). Conditions are very much in the locus of control for schools and school districts. When considering a change in a program, the conditions create opportunities for change through organization of time and allocation of resources.

Finally, Culture is defined by Wagner (2006) as the "shared values, beliefs, assumptions, expectations, and behaviors related to students and learning" (p. 102). Culture encompasses both the individual mindsets of the stakeholders, as well as the collective mindset of each subgroup. Each school system has an overarching culture, each school has a culture, and teachers and administrators have their own culture. Without a culture that supports the desired outcome, achieving said outcome is extremely challenging.

Figure 1. *As Is Diagnostic Tool*

“As Is” 4 C’s Analysis for Improving Teacher Evaluation Implementation

Context
- The school is a part of a large district, in which there is an adversarial relationship between teachers and the district.
- Teachers object to the use of student test scores in their evaluation.
- There have been many leadership changes, resulting in a lack of focus.
- There has been inconsistent and shifting support for teacher evaluation.

Conditions

+
- Principals have access to manuals, rubrics, and training at least once per year
- Principals must be certified, which includes norming with external evaluators
- Principals are motivated by having effective teachers
- Principals evaluate all teachers at least once per year

−
- The highest rated teachers work in schools with the lowest poverty levels
- Teachers in the schools with the highest poverty levels receive lower ratings
- Observations are evaluative
- One size fits all rubric does not work for all practitioners
- There is a wide range of vision of the different performance levels

Culture

+

−
- Lack of trust between teachers and the district
- Lack of trust between teachers and administration

- Teachers believe that their evaluators are capable, and they recognize the value in the REACH Framework, but observations cause stress and teachers feel that the potential benefits are not outweighed by the costs
- Observations are seen as evaluative
- Teachers are protective of their practice

The teacher evaluation process has the support of teachers and administrators, but its effectiveness is undermined by a lack of trust, consistency, and commitment.

Competencies

+
- Principals are trained on the evaluation process and Framework for Teaching
- Some principals are trained on adult learning and coaching
- Experienced Instructional Effectiveness Specialists support principal evaluation in a limited manner

−
- Principals receive scant ongoing training and support
- Learning opportunities for evaluators are not consistent across the district
- Principals tend to rate their teachers significantly higher than external evaluators
- Little changes year to year, resulting in continuation of habits and practices from year to year

Context. Chicago Public Schools (CPS) is the nation's fourth largest school district with 355,156 students and 21,355 teachers. There are a total of 514 principals and 1,446 central office and network staff. A diverse and varied school district spanning the boundaries of the city, CPS is comprised of 516 schools. The demographic makeup of CPS students is: 46.6% Hispanic, 35.9% African American, 10.8% white, and the remaining 6.7% of students are a mix of races (Chicago Public Schools [CPS], 2020).

Relations between the CPS administration and the Chicago Teacher's Union (CTU) have historically been tempestuous, particularly in the last seven years, including a seven day strike in 2012, a one day strike in 2016, and another fifteen day strike in 2019. Mayor Rahm Emanuel and his appointed school board unilaterally lengthened the school day without increasing teachers' salaries, providing the primary rationale for the 2012 strike. During the period of 2012-2018 CPS was led by four different Chief Executive Officers (CEO). Of the four, Barbara Byrd-Bennet served the longest term as CEO and was convicted in a kickback scheme with a professional development provider. This instability and lack of credibility further undermined the relationship between CPS administration and the CTU. As the teachers' contract that was negotiated to end the 2012 strike elapsed in April 2016, teachers engaged in a one day strike, called a "Day of Action", and subsequently negotiated a new contract that reduced the number of observations that each teacher was required to receive. Tenured teachers now receive three observations over two years, instead of four, and non-tenured teachers receive two formal and one informal observation each year, instead of three formal and one informal observation. After the strike of 2019, teacher evaluation remained unchanged in CPS.

The CTU has consistently opposed the use of student test scores in teacher evaluation; however, the Performance Evaluation Reform Act (PERA), which the Illinois General Assembly

passed into law on January 1, 2010, requires that all schools in Illinois use student growth scores as a part of teacher evaluation as of September 1, 2013. PERA requires the use of either a Type I Assessment, a standardized assessment used by multiple districts which is created and scored by an independent party, or a Type II assessment, created and scored by the district itself, and used in multiple schools district-wide. Additionally, PERA requires that each district uses a Type III assessment, which is aligned to the teacher's specific curriculum and can be selected or created by the teacher.

Initially unwilling to exclude the VAM scores (Type I Assessments) from any teacher's evaluation, CPS assigned a school wide VAM score to teachers who did not teach reading or math in the tested grades, 3rd -8th. CPS has since modified its evaluation program so that teachers who do not teach reading or math in the tested grades use two Type III performance task assessments and neither a Type I nor a Type II assessment. For all other teachers, CPS employs a VAM, which is applied to standardized student test scores, as the primary component of each teacher's student growth score. The use of the VAM in teacher evaluation is controversial, and according to Darling-Hammond et al. (2012), the VAM "should not be used for high-stakes, individual-level decisions, or comparisons across highly dissimilar schools or student populations" (p. 8), "because they can't control or disentangle influences on student progress; they inconsistently rate teachers, and they don't account for students assigned to teachers in a particular year" (Darling-Hammond et al., 2012, p. 6). Many researchers caution against the use of the VAM in teacher evaluation. Braun (2005) concluded that VAM results should "not serve as the sole or principal basis for making consequential decisions about teachers" (p. 15). McCaffrey et al. (2003) note that for many school districts, the results may only serve as "a starting point for administrators (such as principals or superintendents) to target teachers for

more thorough review" (p. 120). Marzano (Quinn, 2014) notes the limitation of student test scores in evaluating teachers, claiming that while he supports the use of student test scores in teacher evaluation, it is "one piece only of the evidence for student learning" (p. 16). Cochran-Smith and Lytle (2006) note that the presence of this accountability measure "narrows the curriculum" and "oversimplifies the process of teacher learning and practice" (p. 670).

In addition to narrowing the curriculum, some teachers modify their approach based on prior student outcomes. Smith et al. (2001) found that teachers in classrooms with higher levels of prior achievement were more likely to use interactive instruction and less likely to use didactic instruction. Interactive instruction promotes participation and active involvement of students. Students interact with each other and the instructor. Conversely, didactic instruction is a one-way street, knowledge delivered from the teacher to the student. Didactic instruction looks like a lecture rather than a classroom of active students. Using the Illinois Test of Basic Skills, the former required Illinois standardized assessment, mathematics and reading tests for 1996 and 1997, and based on teacher responses to the 1997 Chicago Consortium Survey, in schools where teachers employed didactic instruction often, student outcomes were 3.9% lower in math and 3.4% lower in reading than the city average. Conversely, the student outcomes for teachers who used less didactic instruction were 4.4% higher in math and 3.7% higher in reading than the city averages (Smith et al., 2001). In schools where teachers employed interactive instruction less often, student outcomes were 4.5% lower in math and 4.5% lower in reading than the city average. Conversely, the student outcomes for teachers who used more interactive instruction were 5.1% higher in math and 5.2% higher in reading than the city averages (Smith et al., 2001).

While they must include a measure of student progress, CPS is not required to use a VAM in teacher evaluation. Teachers expressed concern about the use of student growth on test

scores, despite the legal requirement of using them (Jiang, Sporte, & Luppescu, 2015). Inclusion of the VAM in teacher evaluation scores in CPS undermines teacher trust in the entire evaluation program (Sporte et al., 2013). The conflict regarding teacher evaluation resulted in adversarial relationship between the teachers and CPS and a lack of trust in the teacher evaluation program.

When designing the REACH Students teacher evaluation program, CPS modified the Danielson Framework slightly in order to create their CPS Framework for Teaching, which is the rubric that administrators use for rating teacher and student behaviors during teacher observations. CPS has subsequently created additional frameworks with which to assess non-core subject teachers and staff, including special education, library, counselor, case manager, and related service providers: social work, psychology, nursing, and speech language pathology.

Competencies. Support for administrators in teacher evaluation has diminished, from CPS requiring several professional development sessions per year in 2013, to none in 2019. CPS creates a premade template of a presentation for administrators to present to teachers each fall. A team of Instruction Support Specialists (ISS), who are certified evaluators, collaborate with principals to support them in observation practice and also observe and contribute to the evaluation of a small number of teachers, as needed or requested by principals.

The ISSs provide a useful service to principals by conducting norming observations with them. In this process the principal and the ISS observe a teacher simultaneously and compare notes and ratings after the observation. Sartain et al. (2013) found that there was a marked difference between the scores that a principal gave a teacher when compared with those of an external observer of the same lesson. Principals expressed that they were more likely to give a teacher a higher rating in order to avoid conflict and preserve a positive relationship (Sartain et al., 2011, p. 16). Another explanation for this disparity could be that the principal took into

account previous observation scores, as some of the variation disappeared when accounting for

prior observation scores (Sartain et al., 2011, p. 16).

Table 1		
Comparison of Principal REACH Ratings with External Observer Ratings		
	Odds Ratio	
REACH Performance Level	Principal	External Observer
Distinguished	6.18	1
Proficient	1.5	1
Basic	1	1
Unsatisfactory	0.74	1

(Sartain et al., 2011, p. 16)

Table 1 shows that principals were more than six times as likely to give teachers the

highest rating, Distinguished, one and a half times as likely to give teachers the second highest

rating, Proficient, and 26% less likely to give teachers the lowest rating, while the principal and

observer were equally as likely to give a Basic rating.

This inherent subjectivity further undermines trust in the evaluation program. While in

this case, the scores for teachers appear to be artificially high when rated by a principal, any

appearance of subjectivity could potentially undermine the results in teachers' minds, especially

if they do not earn the score that they believe they deserve. Sartain et al. (2011) identify this trust

between the principal and teacher is central to the purpose of teacher development, as it is the

"foundation to have critical conversations about instructional practice" (p. 58). Without adequate

trust, conversations about instructional practice are less likely to have the desired impact.

Principals are further challenged by the fact that they work in isolation. Some principals

are the sole evaluator in their school and others have one or a handful of colleagues with which

to share the workload, collaborate, and norm ratings. With minimal professional development

and collaboration, principal practices in teacher evaluation tend to become habits and carry over

from year to year with little revision or growth. With the demands of the job, there is an incentive for principals to find the most efficient manner to complete the observations rather than to focus on effective coaching to help teachers engage in a continuous improvement process.

CPS is divided into 17 networks, each with its own Chief of Schools who is responsible for developing principals. Some principals have successfully applied to be more autonomous, and these Independent School Principals are tasked with finding their own professional development. Network Chiefs possess broad autonomy with regard to professional development for principals. As a result, professional development for principals is inconsistent across the district. While all principals engage in the practice of training teachers on the evaluation process, which forces them to learn the procedures, some principals receive training on best practices in adult learning and coaching from their network staff, while others do not. Support from the district ISSs is not compulsory and done primarily on request.

Developing the teachers' competency for engaging in the evaluation process is equally important to that of the principal. Teachers are only required to attend one annual session on the content and procedures of the evaluation program. There is no mandate for ongoing professional development, no required content for engaging in collegial inquiry or peer observations, and no ongoing support for navigating the evaluation program and its content.

Conditions. Principals must complete the statewide training modules and recertify every five years in order to be able to evaluate teachers. This training offers the best opportunity for administrators to improve their practice in teacher observation and evaluation. In the interim, CPS offers optional professional development opportunities in addition to the yearly presentation that administrators facilitate with teachers. Administrators are motivated to help teachers be as effective as possible, as their own evaluation is based, in part, on student results. Administrators

conduct a large number of observations each year, so they have the opportunity to observe a wide range of practices and learn best practices from effective teachers.

There is, however, a significant disparity between the ratings of teachers in schools with a high number of students living in poverty and those teachers in schools with a low number of students living in poverty. CCSR has identified this disparity in the distribution of highly rated teachers across schools of various poverty levels. Teachers at high performing schools tend to earn higher ratings than teachers at low performing schools. Jiang and Sporte (2016) noted "On observation scores, teachers in lower poverty schools have substantially higher scores on average than teachers in higher poverty schools" (p. 2).

The distribution of teacher evaluation scores is inequitable across poverty level quintiles. Teachers with the highest observation scores are overrepresented in schools with lower poverty levels, and teachers with the lowest observation scores are overrepresented in schools with higher poverty levels, as seen in Table 2.

Table 2

Percentage of Teachers with the Highest Observation Scores

Poverty %	% of Teachers
0%-20%	34%
20%-40%	22%
40%-60%	22%
60%-80%	15%
80%-100%	6%

(Jiang and Sporte, 2016, p. 10)

As seen above, only 6% of teachers with the highest observation scores work schools with the highest poverty levels, while 34% of teachers with the highest observation scores work in the schools with the lowest poverty levels.

Table 3

Percentage of Teachers with the Lowest Observation Scores

Poverty %	% of Teachers
0%-20%	9%
20%-40%	21%
40%-60%	18%
60%-80%	23%
80%-100%	30%

(Jiang and Sporte, 2016, p. 10)

Conversely, Table 3 shows that 9% of teachers with the lowest observation scores work in schools with the lowest poverty levels, while 30% of teachers with the lowest observation scores work in the schools with a poverty level in the top quintile. Fifty three percent of schools were in the top two quintiles of school poverty.

Table 4 reveals that teachers in low poverty schools tended to earn higher ratings than teachers in high poverty schools. As the school poverty level diminishes, the average teacher evaluation score increases

Table 4

Average Teacher Evaluation Scores by School Poverty Level

School Poverty Level	Average Score (SD)	Controlling for Teacher Background
1-Lowest	332 (42)	331
2	312 (49)	308
3	312 (48)	305
4	304 (48)	298
5-Highest	289 (44)	288

(Jiang and Sporte, 2016, p. 12)

On the other hand, teachers in low poverty schools earn significantly higher scores than their counterparts in high poverty schools, when controlling for teacher background, with an average score of 331, versus an average score of 288 for their counterparts in schools with a higher level of poverty.

This uneven distribution of teachers creates inequity for the students of Chicago, yet the correlation does not prove causation. As was the case with the principals who rated their teachers' practices higher than the outside evaluator, there may be other factors in play here. Principals in schools with lower poverty levels and higher achievement face less pressure to increase student test scores. Attribution of the higher test scores to higher quality instruction may be misplaced, but the inequity is a cause for concern. Certainly our students with the greatest need should benefit from our highest quality teachers, as those students in the lower poverty schools currently do. Linking teacher evaluation to student standardized test scores is a factor in this inequity, in spite of the fact that the cause for the inequity is unproven.

Not only do students in higher poverty schools potentially suffer from lower quality instruction, but the teachers in these schools also experience inequity in how teachers in higher poverty schools are generally rated. Furthermore, the students in higher poverty schools tend to score lower on standardized tests, which, in turn, lowers teachers' VAM scores and by extension, their evaluation scores. This is a vicious cycle in which teachers are victimized for conditions which are not within their control, creating an incentive for teachers to move to a lower poverty level school where teachers tend to earn higher observation ratings and student results tend to be higher.

Culture. As a result of the conflicts of the last several years between CPS administration and the CTU, many teachers are distrustful of CPS administration, inclusive of central office and school administrators. The teacher and administrator attitudes towards the REACH Students program, however, have remained largely positive. The responses in Table 3 indicate that a majority of teachers believe that this program will lead to improved student learning, are satisfied with the program, and have made instructional modifications based on feedback that

they received. Over 85% of administrators believe that the program will lead to improved

instruction.

In the two years studied, 60%-65% of teachers either agreed or strongly agreed that the

REACH Students observation framework would lead to improved student learning.

Table 5

Teacher Perception of REACH Students Framework Impact on Instruction

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	Year
Overall the evaluation system will lead to improved student learning.	8%	27%	51%	14%	2013-2014
	9%	30%	48%	13%	2014-2015
Overall the evaluation system will lead to better instruction.	8%	24%	51%	14%	2013-2014
	9%	27%	51%	13%	2014-2015

(Jiang and Sporte, 2014, p. 3, Jiang and Sporte, 2016, p. 4)

Teacher support for the program remained consistent over the two years, which indicates

that year over year REACH was implemented in a manner which gained and maintained the trust

of almost two out of three teachers, as seen in Table 5.

Table 6

Administrator Perception of REACH Students Framework's Impact on Instruction and Teacher Effectiveness

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	Year
The observation process will lead to better instruction at my school.	1%	10%	66%	23%	2013-2014
	1%	14%	65%	20%	2014-2015
Is a useful tool for identifying teacher effectiveness in this school	2%	5%	53%	40%	2012-2013
	1%	11%	63%	25%	2013-2014

(Jiang and Sporte, 2014, p. 3, Jiang and Sporte, 2016, p. 5)

Table 6 shows that administrator support for REACH was even stronger than teachers'

support, with 85%-95% of administrators expressing the belief that using the framework would

improve instruction in their school. By the same margin, administrators expressed the belief that

REACH was a useful tool to assess teacher effectiveness.

During the 2014-2015 school year, a majority of teachers indicated that they were

satisfied with the teacher evaluation process, and more than 80% of teachers indicated that they

had used feedback from the process to make changes to their instruction, as seen in Table 7.

Table 7

Teacher Perception of the REACH Students Framework

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	Year
Overall, I am satisfied with the	8%	21%	49%	23%	2012-2013
Teacher evaluation process at this	12%	25%	51%	11%	2013-2014
school.	12%	28%	52%	10%	2014-2015
I have made changes to my teaching	3%	10%	56%	30%	2013-2014
as a result of this observation					
process.*	3%	9%	51%	37%	2014-2015

(Jiang and Sporte, 2014, p. 3, Jiang and Sporte, 2016, p. 6)

Table 8 shows that while more than 90% of teachers expressed the belief that their

administrator was able to accurately assess their instruction, teachers stated that the evaluation

process increased their stress level, with the percentage of teachers who strongly agreed that it

increased their stress level nearly doubling from 20% to 38% from year one to year two of the

program. During the 2014-2015 school year more than half of teachers, 57%, believed that the

evaluation process required more effort than it was worth.

Table 8

Teacher Perception of REACH Students Framework Evaluators and Value

Question	Strongly Disagree	Disagree	Agree	Strongly Agree	Year
My evaluator is able to accurately assess my instruction.	2%	6%	24%	67%	2014-2015
The evaluation process has increased my level of stress and anxiety.	4%	18%	39%	38%	2014-2015
	6%	36%	39%	20%	2013-2014
The evaluation process takes more effort than the results are worth.	6%	38%	39%	18%	2014-2015

(Jiang and Sporte, 2014, p. 3, Jiang and Sporte, 2016, p. 7)

Understanding that observations ultimately result in an evaluative score creates stress for teachers, and many do not see the process as an opportunity for growth, rather they see it as a judgement of their teaching practice. This stress inhibits open dialogue and diminishes a willingness to discuss teacher needs or opportunities for growth, as teachers have reported that they "fear that they will be unfairly judged" (Jiang and Sporte, 2014, p. 12). Teachers are motivated to be seen in the best possible light, so if they are aware of the observations in advance, they will likely prepare differently than they would for an unobserved lesson. These observed lessons are not likely to be a true representation of the teacher's daily practice, rather they are more of what is commonly referred to as a "Dog and Pony Show". I have observed this practice many times over the past seven years.

With experience, teachers develop an approach to teaching that is comfortable and familiar, which presents a barrier to teachers viewing the observation process as an opportunity for growth when it pushes them beyond their comfort zone. I have experienced this resistance to change with teachers whom I have evaluated year after year. We find ourselves engaging in repeated conversations as I push them on specific components of their practice. Domain 3:

Instruction, Component b: Using Questioning and Discussion Techniques, of the CPS

Framework for Teaching, presents a challenge for many teachers. Releasing control and fostering

true student to student discussion of the lesson content is a challenge for teachers who do not

want to give up control of the content or allow students to speak freely during class, fearing lost

instructional time or off-topic student discourse.

To build on the strengths of the current REACH Students teacher evaluation program, it

is necessary to consider the concerns of the practitioners and the tension between the current

state of teacher evaluation and the recommendation of the body of research. Through this

consideration, a set of recommendations emerges to improve the current state of teacher

evaluation in Chicago for the benefit of the teachers and, most importantly, for the benefit of the

students.

Interpretations

In interpreting these data, it is worthwhile to note that the REACH Students teacher

evaluation program employs several best practices and enjoys broad support from teachers and

administrators with over 60% of teachers and over 80% of administrators supporting the

program, however, here are opportunities for improvement, and surely inequities exist within and

as a result of the program. Overall, there is significant evidence indicating that the program

fosters continuous improvement, collegial inquiry, and fair and valid teacher evaluation.

While teacher trust is undermined by the use of VAM scores in the overall teacher

evaluation score, the REACH Students program's existing strengths create a foundation upon

which to build trust. The most glaring inequity is evident in the distribution of highly rated

teachers across schools with varying levels of poverty, with merely 6% of these highly-rated

teachers in the most impoverished schools and 34% of these teachers at the schools with the

lowest levels of poverty. The disproportionate lack of highly rated teachers in schools with high poverty levels must be addressed in order to achieve equal access to high quality teachers in Chicago across boundaries of socioeconomic status.

By building upon existing support for the teacher evaluation program, CPS can create ongoing and job-embedded professional development opportunities which will build trust and encourage teacher participation and ownership in the REACH Students teacher evaluation program. CPS can further build teacher ownership in the program by using best practices in professional development and adult learning, including: collegial inquiry, peer observations, and ongoing opportunities to engage with colleagues and give feedback.

Teacher and administrator perceptions demonstrate widespread support for the program; a majority of teachers trust their administrator and believe that the teacher evaluation program will lead to improved student outcomes. On the other hand, most teachers feel a significant amount of stress and do not believe that the benefits of this program outweigh the costs. By building ownership and offering opportunities for safe practice with colleagues, teachers will have more of an opportunity to experience the benefits of the program and enjoy diminished stress as they become more familiar and comfortable with it.

Finally, the use of VAM in teacher evaluation produces inequitable and inaccurate evaluation scores. Teachers are held accountable for test scores that are influenced by student demographic and environmental factors, despite the VAM calculation which is designed to control for these factors. In addition, the benefits that teachers at high performing schools enjoy in the VAM incentivize the higher rated and more experienced teachers to move to higher performing and lower poverty schools. This incentive does a disservice to the neediest students who would benefit most from the more experienced and higher rated teachers.

Judgements

This inequitable access to highly rated teachers may be a consequence of the use of VAM scores in teacher evaluation. Since student scores comprise 30% of a teacher's overall evaluation score, teachers are incentivized to work in schools where students earn higher scores as this will lead to higher evaluation scores. In general, students in schools with low poverty levels tend to score higher than their counterparts in schools with higher poverty levels. In addition, the improved environment and diminished pressure to raise test scores in higher performing schools creates a less stressful environment. As a result of these incentives, more experienced and highly-rated teachers are motivated to move towards the schools where there are fewer students in poverty.

The inclusion of VAM scores in teacher evaluation undermines teacher trust in the program as it fails to fully control for students' environmental factors (Darling-Hammond et al., 2012, Rothstein, 2010). This practice punishes teachers who work in schools where students face more barriers to success in education, and unfairly benefits teachers whose students do not face as many demographic and environmental challenges.

The current system of support and professional development for teachers with regard to evaluation is inadequate. Teachers are only required to participate in one professional development session each year, which focuses on the evaluation process and includes only cursory coverage of content. There is little support for administrators and teachers to engage in collegial inquiry and a process of continuous improvement with professional learning that focuses on compliance over content.

Recommendations

In order to address the equity concerns, VAM scores should be removed from teacher evaluation. The necessity to comply with the student growth requirement of PERA, in light of the changes in ESSA in 2016 that removed the VAM requirement, grants CPS the ability to allow schools to develop curriculum-based assessments and demonstrate student growth in a more accurate and equitable manner. Exclusion of VAM scores from teacher evaluation could affect the distribution of highly-rated teachers across schools of varying levels of poverty, and it would increase teacher trust in the program. To improve trust, CPS should create systems to include teacher voice in the decision making process about what is included in teacher evaluation and how it is implemented.

To further strengthen the REACH Students teacher evaluation program, CPS should create more opportunities for administrators and teachers to engage in collegial inquiry as well as opportunities for teachers to engage in professional learning communities. These practices would help refine the program, and would improve instruction through non-evaluative, regular, peer observations. A picture of an ideal state, considering these recommended changes, follows in the next section.

CHAPTER FIVE: TO-BE FRAMEWORK

Envisioning Success: To Be

Wagner et al. (2006) guide leaders through the process of revising the As Is to the To Be, or goal state, by asking the question, "What would success look like if the problem you identified were solved?" (p.119). Upon extensive examination of the current state, published research, and best practices in teacher evaluation, a picture of a comprehensive and effective teacher evaluation program emerges.

On the surface, the REACH Students CPS program for teacher evaluation seems to have many of the necessary components of effective teacher evaluation, as defined by Darling-Hammond et al. (2012): "multiple classroom observations and data sources, expert administrators, rubrics, mentors, collaboration with other teachers, and professional development" (p. 8). In order to engage teachers in continuous improvement, each of these components must be in place and implemented correctly. While administrators in CPS engage in pre-observation and post-observation conferences with teachers, these conversations must be conducted according to best practices in order to have the greatest possible impact. Danielson (2010) states that in order for teacher evaluation to improve practice, evaluators must "engage teachers in those activities that promote learning—namely self-assessment, reflection on practice, and professional conversation" (p. 37). Self-assessment is notably missing in the CPS teacher evaluation program. Drago-Severson (2009) contends that engaging in collaborative reflective practice will have a positive effect on school communities and systems (p. 154). Effective coaching of teachers and the use of Drago-Severson's adult learning theory will guide the nature of professional conversations in the improved teacher evaluation program.

Context. Relations between the CPS administration and the Chicago Teacher's Union (CTU) will have improved dramatically. With a unified vision of an effective teacher evaluation program and an understanding that the primary purpose of this program is to engage teachers in a continuous improvement process, CPS and CTU will have become thought partners with the common goal of effective teachers leading well-educated students.

The leadership at CPS is stable, with a former CPS educator, Dr. Janice Jackson, as CEO who has created a five-year vision for the district. Having survived two mayors, an anomaly for the CPS CEO in Chicago, Dr. Jackson appears to be on track to be a long term leader for the district. The district is more focused on the needs and responsibilities of individual schools. School teams set goals that align with the district vision and ensure long term success. Every member of the organization is focused on teaching and learning, and test scores are just one indicator of success, not the final determination. Support has replaced remediation, and teachers are free to be creative and make mistakes while continuing to grow and innovate.

CPS has created comprehensive and relevant frameworks for each teaching position that provide clear expectations and a vision of success. Principals use the components of these frameworks to plan for teacher development, based on individual teacher needs.

CPS will have eliminated the use of VAM scores in teacher evaluation, resulting in many benefits for teachers and students. Teachers are free to expand their curriculum beyond the narrow focus of reading and math without as much pressure from the end of the year standardized tests, teachers use more interactive instruction and engage students in project-based learning. Teachers measure student progress using classroom assessments and portfolios, which demonstrate mastery of standards and accurately demonstrate student growth over time.

Competencies. Instructional Effectiveness Specialists provide regular and ongoing support to principals, and all principals participate in professional learning communities in which they collaborate with other school administrators, visit other schools and observe teachers for the purpose of norming their ratings. As a result of this practice, principals have a common vision of high-quality instruction, and they rate their teachers' performance in a similar manner as their peers and outside observers.

In addition to providing increased validity to teacher observation ratings, participating in professional learning communities provides administrators with the opportunity to regularly discuss, adjust, and improve their teacher evaluation practices. Sharing ideas and troubleshooting problems with other administrators in a systematic manner fosters continuous improvement in the continued implementation of the teacher evaluation program.

CPS provides training to all principals on the expectations for teacher evaluation and the frameworks for instruction. All networks across the district implement a unified curriculum for developing the capacity of principals to effectively coach teachers, employing best practices in adult learning both for the principal learning sessions and to improve the principals' coaching sessions with teachers. Teachers independently engage in self-assessment and reflection, and administrators and teachers engage in effective professional conversations focused on continuous improvement for teachers.

Conditions. Principals participate in professional learning communities with Instructional Effectiveness Specialists and other principals to ensure a unified vision of effective teacher practice. CPS provides ample training and access to resources for each of the individual frameworks, as well as yearly update sessions to elicit feedback and make needed changes to the resources or the process.

Principals build trust with teachers by engaging in productive professional conversations as well as regular, non-evaluative classroom observations with timely feedback in order to support the teachers in their continuous improvement process. Principals and teachers work together to create and maintain a shared vision of effective teaching and plan intentionally for improvement in teacher practice.

This additional support for the principals and administrators who observe and evaluate teachers serves to build trust between teachers and administration. The CTU and CPS administration participate in yearly assessments of the evaluation program. These assessments include feedback from teachers and administrators, an examination of aggregate teacher scores to uncover any potential bias, and norming of teacher ratings using videos of teachers at each performance level. CPS trains teachers on the Framework for Teaching and teachers participate in PLCs in which they observe each other, assess their practice using the rubric, and debrief the process with CPS administration. By working together and setting yearly goals, the teachers and CPS administration take ownership of the process, which increases trust and investment in the program.

Culture. Stability in CPS administration has improved its credibility and increased trust between administrators and teachers. With the changes made to the implementation of the teacher evaluation program, teachers understand and appreciate the potential benefits and opportunities to improve their teaching practice. By engaging with evaluators who are trained and adept at conducting unbiased and productive observations and effective coaching conversations based on adult learning theory, teachers gain trust in the evaluation process. Teachers are able to focus more on improving their teaching practice to promote student success rather than on evaluation scores.

As a result of this collaborative process, teachers do not feel stress when engaging in the observation process, instead they relish the opportunity to engage in reflection and self-improvement. Teachers continue to trust their administrators and incorporate feedback into their instruction.

As a result of this increased trust, teachers have opened their doors to colleagues and administrators. Evaluation remains legally mandated, but the resources dedicated to it in Chicago Public Schools are intended to engage teachers in a continuous improvement process to the benefit of the students.

This vision of the current state of teacher evaluation in Chicago can be achieved by increasing the level of trust that the teachers possess in the teacher evaluation program and by strengthening the support for teachers and administrators to engage in a process of continuous improvement.

Figure 1. *To Be Diagnostic Tool*

"To Be" 4 C's Analysis for Improving Teacher Evaluation Implementation

Context

- The school is a part of a large district, in which there is a productive and trusting relationship between teachers and the district.
- Leadership is stable from year to year, with a clear focus on factors that contribute to success.
- There is ongoing, job-embedded support for teacher evaluators.
- Teacher evaluation no longer includes student standardized test scores.

Culture

- There is a high level of trust between teachers and school and district administration.
- Teachers view observations as opportunities for growth rather than a judgement on their practice.
- Teachers are willing to share and receive feedback on their practice.

The teacher evaluation process supports professional growth, develops trust, and motivates teachers to engage in continuous improvement.

Conditions

- Principals have access to manuals, rubrics, and ongoing training.
- Principals are motivated by having effective teachers.
- Principals evaluate all teachers at least once per year.
- Some observations are non-evaluative, and teachers observe each other.
- Specific rubrics exist for each type of practitioner.
- There is a unified vision among teachers and evaluators of the different performance levels
- Teachers participate in the decision making process about teacher evaluation.

Competencies

- The district trains evaluators continuously on the evaluation process and rubric.
- The district trains evaluators on best practices in adult learning and coaching.
- Experienced Instructional Effectiveness Specialists support evaluators on a regular and ongoing basis.
- Principals participate in professional learning communities with each other to share practices and norm ratings.
- Principals engage in continuous improvement to make adjustments to teacher evaluation based on yearly feedback from evaluators and teachers.
- Teachers participate in professional learning communities to learn the process, better understand the frameworks, and conduct peer observations.
- Teachers offer feedback on the process and make suggestions for improvement.
- CTU and CPS work together to gather feedback and make adjustments from year to year

CHAPTER SIX: STRATEGIES AND ACTIONS

The organizational improvement process described by Wagner et al. (2006), provides a framework to create a vision of future success. An examination of the current state, when compared with this vision of the ideal state, brings to light strategies and actions to bridge the gap between the two states. Teacher evaluation in CPS is an enormous initiative which affects the daily life of tens of thousands of teachers and administrators as well as the lives of hundreds of thousands of students. As such, it is somewhat inured to change by virtue of the magnitude of the program, while at the same time, it is a highly impactful and essential program for the teachers, administrators, and most importantly, the students of Chicago Public Schools.

Taking into account the recommendations of researchers and the vision of the ideal state, CPS could implement a series of modifications to its teacher evaluation program in order to create a more equitable and effective program which truly serves the needs of all of its stakeholders.

Strategies and Actions

Three main areas of improvements to the teacher evaluation process in CPS would greatly improve its equity and impact. First, CPS should create a comprehensive system of professional learning for teachers and administrators in order to ensure that the system is implemented properly and all stakeholders possess the capacity for engaging in it with the maximum positive impact. Secondly, CPS should take steps to intentionally build trust with teachers, both to overcome challenges from prior interactions and to create a sense of shared purpose with the benefit of students as the primary focus. Lastly, CPS should create professional learning communities of teachers in order to promote better understanding of and investment in the evaluation process through non-evaluative peer observations, collaborative learning, and

opportunities for teachers to inform adjustments to the implementation plan for the REACH

Students teacher evaluation program.

In March of 2016, The Chicago Public Education fund published a brief titled, *Voices from the Field: Strategies from School Leaders to Improve REACH*, which made three recommendations for improving teacher evaluation in Chicago, based on feedback from administrators:

- Get out of their way.
  - Provide leaders with more flexibility around observations and protect their time during the instructional day.
- Help them leverage the tools they already have.
  - Offer clear resource recommendations for common instructional challenges and produce timely data on REACH outcomes.
- Support their individual development.
  - Acknowledge and account for unique school context and provide customized training. (p. 2)

Allowing greater flexibility for administrators will not only give them the opportunity to be more responsive to the needs of the individual teachers, it will, in turn, foster more trust in the system from teachers as it becomes more responsive and flexible. The second and third recommendations relate to professional development opportunities for administrators to be able to leverage what already exists, and for the district to be more responsive and meet the individual needs of each administrator. In addition, "school leaders suggested that the district could offer quarterly sessions on effective coaching conversations" (p. 8). Explicitly teaching leaders how to meet the needs of adult learners will improve the impact of the REACH program.

A comprehensive system of professional learning is essential for sustained success in any school system. Drago-Severson (2009) noted that much of the work in schools is done in isolation "without the benefit of a supportive yet critically thoughtful observer" (p. 15). Furthermore, Drago-Severson (2009) recognizes that the district, school leader, and teacher all

play a role in building collaborative learning communities. By creating the systems and structures to support professional learning, including regularly scheduled professional development sessions with all stakeholders and hiring a team of expert evaluators to support both the administrators and teachers in the process, CPS can create the conditions for stakeholders to engage in collegial inquiry and reflective practice. Creating teams of learners "can be a support to both individual and organizational learning and development" (Drago-Severson, 2009, p.154).

In order to better support teachers and administrators in the evaluation process, CPS should schedule professional development sessions at least four times per year with stakeholder groups. In addition to the time that teachers spend in PLCs with other teachers, they would benefit from meetings with teachers from other schools during which the district can share best practices and foster collegial inquiry with regard to the evaluation program. This would also be an opportunity for teachers to give feedback to the district on the content and implementation of the teacher evaluation program.

Building on the current practice of the Instructional Support Specialists, CPS should expand this program in order to provide better support to administrators and teachers. These ISSs would work with administrators in order to norm both observation practices and interpretation of the Framework for Teaching. Ensuring that administrators across the district rate teacher observations similarly improves the validity of all of the observation ratings. Additionally, CPS should recruit teacher specialists to support teachers in peer observations and conduct deep dives into the components of the evaluation rubric.

As noted by Bryk & Schneider (2003), in order to build trust, teachers must have their feedback incorporated into the implementation process. As such, the teacher evaluation program must include opportunities for teachers to have input into how the program is structured and

implemented. Creating systematic opportunities for teachers to have a voice in the implementation of teacher evaluation will increase teacher trust and buy-in in the program. This opportunity can be offered during the quarterly professional development sessions, through surveys, and focus groups as well as at an annual meeting with CPS and CTU.

Increasing the time that administrators spend in classrooms by including non-evaluative observations will further enhance teacher trust in the program. These non-evaluative observations will serve as an opportunity for administrators and teachers to experience the process in safe practice and have a conversation about teaching and learning without the distraction of an associated evaluation. The time spent with the teacher and in the classrooms will serve to strengthen the rapport and relationship between the administrator and teacher. White et al. (2013) found that increasing the amount of time that administrators spend in the classrooms will "alleviate concerns about the accuracy of ratings, build teacher trust, and promote improvement" (p. 13).

Importantly, removing the use of VAM scores from teacher evaluation will increase teacher trust by eliminating the potential inequity caused by environmental and demographic factors which affect student assessment results. Darling-Hammond (2013) noted that the use of standardized test scores in teacher evaluation undermines teacher trust, "because they can't control or disentangle influences on student progress; they inconsistently rate teachers, and they don't account for students assigned to teachers in a particular year" (p. 6). In lieu of using standardized test scores in teacher evaluation, Darling-Hammond (2013) suggests that appropriate use of student work samples and portfolios is a better indicator of teacher effectiveness than student assessment results.

In addition to the professional learning opportunities that CPS creates for all stakeholders, and in order to further increase trust and teacher buy-in in the program, CPS should create professional learning communities of teachers to create opportunities for teachers to better understand the process and expectations of evaluation while engaging in low-stakes, non-evaluative peer observations and feedback. Peer observation embeds professional learning into the teachers' daily work, which Darling-Hammond (2013) noted was most effective when it is ongoing and embedded in the teacher's practice. Students of teachers who engage in peer observation have been shown to benefit from the process. Burgess et al. (2019) found, "the observed teachers do benefit from peer observation, as measured by the test score gains of their pupils" (p. 29). Furthermore, Burgess et al. (2019) state that the "pupils of the observing teachers also benefit, and perhaps to a greater extent than the teachers they observe" (p. 29). By creating the opportunity for teachers to engage in peer observation, CPS can further strengthen the teacher evaluation program and simultaneously improve outcomes for the students.

Strategy	Action
Create a comprehensive system to support observation validity, alignment, and proper implementation of the teacher evaluation program	• Create a system for teachers to systematically self-assess and reflect on their practice. • Hire a team of expert evaluators, who participate in regular norming observations with administrators and support a small number of school-based administrator evaluators, 5-8 evaluators each, in order to ensure best practices. • Evaluation experts participate in norming observations and review feedback with evaluators each month. • Train all evaluators on adult learning theory and effective coaching strategies.
Intentionally build trust between the district and teachers with regard to teacher evaluation	• Eliminate the use of VAM scores in teacher evaluation. Implement curriculum-based assessments to demonstrate student growth. • Enlist teachers to be a part of the team that designs and implements the teacher evaluation program and that supports evaluators. • Gather regular feedback from teachers to identify opportunities to increase trust and improve the evaluation program. • Conduct non-evaluative observations with feedback to facilitate teacher development and help teachers to understand the value of the process.
Create a system of professional learning support for teachers to build capacity, ownership, and investment in the evaluation process	• Teacher leaders and evaluators plan together in order to provide teachers the opportunity for safe practice in the evaluation process, including non-evaluative peer observations. • Teachers participate in professional learning communities to learn the process and to better understand the frameworks • Teachers offer feedback on the process and make suggestions for improvement. • CTU and CPS work together to gather feedback and make annual adjustments to the program.

CHAPTER SEVEN: IMPLICATIONS AND POLICY RECOMMENDATIONS

Introduction

PERA, which the Illinois General Assembly passed in January 2010, altered the landscape of teacher evaluation in Illinois, identifying student growth indicators as key metrics for teacher evaluation. In Chicago Public Schools, students have individual goals, based on their performance on the NWEA assessment from the previous spring, and teachers are evaluated by how much growth their students demonstrate the following spring, as well as the percentage of the students who meet their growth goal. This measured growth for each student is used to calculate a VAM score, which comprises 30% of the teacher's summative evaluation score. Teachers have expressed mixed feelings about the use of student growth on test scores (Jiang, Sporte, & Luppescu, 2015).

Significant potential problems exist in the use of VAM scores in teacher evaluation. According to Toch (2009), less than half of all public school teachers teach either of the subjects that are tested and many of the skills that the tests assess are low level skills. Most importantly, as Darling-Hammond (2012) pointed out, there are myriad personal and societal factors which persist in affecting student performance after being controlled for in the VAM calculation.

Policy Statement

I recommend the exclusion of VAM scores in teacher evaluation in Chicago Public School. As noted above, the current use of these scores breeds distrust with teachers. This distrust is well placed, as the correlation between VAM scores and teacher quality is not direct or without complicating factors. Various societal factors have an impact on student outcomes, and

teachers should not be held accountable for these VAM scores over which they have limited control (Darling-Hammond et al., 2012, Rothstein, 2010).

Analysis of Needs

Educational Analysis. We rely on teachers to prepare students to be critical thinkers and to be capable of complex problem solving. It does students a grave disservice to make standardized assessments a primary focus for teachers. Students lose precious instructional time on test preparation, and teachers focus on transferring low-level skills to ensure their students' success on high-stakes standardized tests. As noted above, Cochran-Smith and Lytle (2006) observed that the presence of high-stakes testing in teacher evaluation "narrows the curriculum" and "oversimplifies the process of teacher learning and practice" (p. 670). Berube (2004) demonstrated that the use of high-stakes tests incentivized teaching low order skills, and that "the more mixed and traditional, 'drill and grill' teachers produced the students with the higher scores" (p. 265). Darling-Hammond (2007) also noted the narrowing of the curriculum and the focus on low level skills.

Teachers should be free to design engaging lessons which help students to learn how to think critically rather than be distracted by the need to prepare students for a low skill level, high-stakes test. Furthermore, if the results of a high-stakes test are a component of the teacher's evaluation, it creates a conflict of interest between instructional best practices and the teacher's desire to earn the highest possible evaluation score. This unnecessary distraction is detrimental to the effectiveness of instruction in every classroom that suffers under this mandate.

Economic Analysis. There is little economic cost to excluding VAM scores as a component of teacher evaluation. Conversely, there is enormous potential for financial benefit.

Standardized testing is big business. Between the years 2014-2018, the state of Illinois paid

Pearson $160 million for four years of Partnership for Assessment of Readiness for College and

Careers (PARCC) testing. Certainly, teacher evaluation is not the sole rationale for high-stakes

testing, as standardized tests are used for student promotion and graduation criteria; however, the

use of these scores in teacher evaluation provides a rationale for continuing or expanding the use

of these costly assessments.

There would be some cost associated with revising the current teacher evaluation

programs to account for the elimination of the use of VAM scores, but there would be potential

savings from additional costly contracts with publishers and testing companies. For the most

part, school districts likely have the structures in place to revise teacher evaluation programs;

however, some districts may choose to hire consultants or other vendors to assist them in the

revision. This potential cost is minuscule compared to the cost of publishing and administering

high-stakes standardized tests.

Social Analysis. We have endured the era of accountability in education for several

decades. The No Child Left Behind Act has expired, yet the 2010 PERA law requires the use of

student test scores in teacher evaluation in the state of Illinois. As school districts begin to

recover from the years of austerity as a result of the financial crisis of 2008, there is an increased

opportunity for change. The body of research supports the elimination of the use of VAM scores

in teacher evaluation (Darling-Hammond et al., 2012, Rothstein, 2010, Briggs & Domingue,

2011), yet there has been no movement in this direction in Chicago Public Schools.

It is time for that to change. For too long, teachers have toiled under the shadow of high-

stakes, standardized testing. While publishing and testing companies enjoy the profits of

continued high-stakes testing, teachers and students are inconvenienced at best, and suffer a deleterious effect to the quality of instruction and quality of student learning at worst (Cochran-Smith & Lytle, 2006, Smith et al., 2001).

In addition to the distraction created by high-stakes standardized testing as a component of teacher evaluation, distrust is sown between teachers and administration. Trust is a necessary component of teacher and administrator relationships, as it impacts the effectiveness of a school district. Judging teachers based on standardized test scores creates inequity and undermines this important trust (Sporte, et al., 2013, Danielson, 2016, White et al., 2012, Byrk and Schneider, 2003).

This policy change would improve the relationship between school districts and teacher unions. The use of VAM scores in teacher evaluation has long been a sticking point between the Chicago Teachers Union and Chicago Public Schools, and it was one of the primary disagreements in the negotiations leading up to the teachers' strike in 2012 (Davey & Yaccino, 2012). Excluding VAM scores from teacher evaluation would be a good faith concession on the part of district leadership, which could change the tenor of labor relations and compel the union to make concessions of their own. With an increased level of trust between the school districts and teachers unions, a more productive and collaborative relationship could ensue, to the benefit of all members of the community.

Political Analysis. From the federal level, President George W. Bush shepherded in the No Child Left Behind Act and President Obama funded the $4.35 billion Race to the Top grant, which resulted in PERA and the mandate to use student test scores as a part of teacher evaluation in Illinois. In 2016, President Obama signed ESSA into law, removing the blanket mandate to

include VAM scores in teacher evaluation. This new autonomy creates opportunity for a change in the policy pertaining to teacher evaluation. This autonomy has affected the way that teachers are evaluated in many states, as the practice of including standardized tests as measures of student growth in teacher evaluation is waning. As of 2019, only 25 states required the inclusion of standardized test scores, down from 37 in 2015, and only 15 states currently use the Value Added Model, while 23 states have stopped including VAM scores in teacher evaluation (National Council on Teacher Quality [NCTQ], 2019).

Lobbyists promote the interests of their employers - publishers and testing companies - and they have played a role in shaping much of the existing education policy. Working in the interest of organizations seeking to make a profit, these lobbyists do not have an incentive to promote policies that are necessarily in the best interest of students. As stated above, Pearson, an educational publishing and testing company received $160 million over four years in Illinois to administer one test. Faced with a potential loss of profit, these politically connected companies could be a formidable barrier to any policy change that would reduce standardized testing administration.

Legal Analysis. There are no federal laws or mandates which currently require the use of VAM scores in teacher evaluation in Illinois. The Performance Evaluation Reform Act (PERA), which was passed in January 2010, requires the use of VAM scores in teacher evaluation. However, since the 2016 ESSA law, local school districts have the autonomy to choose their student growth measure. School districts are able to choose to use curriculum-based, teacher created or teacher chosen assessments to demonstrate student growth. Powerful lobbyists and politicians who promote "holding teachers accountable" through the use of standardized test

scores would provide resistance, but some districts have already made this change and CPS is free to do so.

Moral and Ethical Analysis. Since there are so many societal factors that contribute to these test scores that are outside the teacher's control (Darling-Hammond et al., 2012, Rothstein, 2010, Caldas & Bankston, 1997), it is unfair to evaluate teachers based on VAM scores. When comparing teacher observations scores with student test scores, Marzano (2012) found "very little relationship between teacher observation scores and student results" (p. 11). Of course, teachers have the responsibility to ensure that they follow educational best practices and offer their students the best possible educational experience. However, relying on VAM scores to measure this impact is misguided and unfair.

Furthermore, there is the larger concern that high-stakes testing has a deleterious effect on students, and disproportionately on at-risk students. Myers and Curtiss (2003) noted that Gary Orfield, the co-director of the Civil Rights Project at Harvard University, found that "high-stakes testing penalized low-income and ethnic minority students and is linked to high dropout rates in these groups" (p. 70). By extension, teachers who teach in low income and minority schools earn lower evaluation scores when compared to their counterparts at higher performing schools, where, in Chicago, Jiang & Sporte (2016) found that "students in high-poverty schools are more likely to be taught by teachers with lower observation and value-added scores" (p. 25). This inequity incentivizes teachers to work in higher performing schools, which results in the more qualified and experienced teacher working at higher performing schools, leaving the less qualified and less experienced teachers to work at the lower performing schools, whose students need the highest quality teachers in order to combat the various societal factors which are suppressing their ability to be successful in school.

In addition to implementation changes to increase teacher trust, voice, and engagement, the exclusion of VAM scores in teacher evaluation in CPS would help to make the REACH Students teacher evaluation system more accurate and equitable. While barriers exist to this policy change, recent changes in law as well as national trends support these recommended changes to improve teacher evaluation in Chicago.

CHAPTER EIGHT: CONCLUSION

Introduction

The Illinois PERA law of 2010 forced Illinois School districts to take a more comprehensive approach to teacher evaluation in Illinois. Chicago Public Schools created the REACH Students teacher evaluation program and modified the Danielson framework to create their own Framework for Teaching. In general, teachers and administrators supported this new evaluation program, and there was some degree of trust between the two groups. While student growth scores were mandated under PERA, the inclusion of VAM scores in teacher evaluation undermined teacher trust in this new evaluation program.

Engaging teachers and administrators in effective collegial inquiry and professional conversations about teacher practice is essential to foster a process of continuous improvement and is a necessary component of an effective teacher evaluation program. While the Framework for Teaching provides a shared vision for high quality instruction, CPS could better support the necessary professional development for both teachers and administrators to make the program as effective as possible. Teachers participate in one professional development session per year, which focuses on the REACH Students evaluation and process, and administrators do not participate in any mandated professional development or enjoy any systematic, ongoing support.

Creating a system of support for professional learning for both the teachers and administrators would not only improve the quality and validity of the observations, feedback, and any adjustments to instruction, but it would increase the necessary teacher trust in the program. Using the existing structure of the REACH Students program and adding these supports would greatly improve teacher evaluation in Chicago.

Additionally, the use of VAM scores in teacher evaluation creates inequity and undermines teacher trust in the program. The inclusion of these scores creates an incentive for teachers to work in schools with higher test scores, since, in general, teachers in schools with higher test scores receive higher evaluation scores. As a result of this discrepancy, more experienced and higher rated teachers are overrepresented in schools with lower levels of poverty. Conversely, schools with higher levels of poverty suffer from an overrepresentation of lower rated teachers. The students in these schools with a high level of poverty need the best possible teachers to combat the myriad societal factors that undermine student academic outcomes. This vicious cycle consistently works to the detriment of the neediest students.

Discussion

The examination of the implementation of the teacher evaluation program in Chicago Public Schools through the voices of the practitioners has drawn into focus the tension between best practices and how teacher evaluation is implemented in Chicago. In addition to the equity concerns of including VAM scores, the lack of teacher trust in the program illustrates an underlying barrier to productive and wholehearted teacher engagement in the program.

By building on the current program, creating embedded opportunities for professional growth for both teachers and administrators, and by giving teachers a voice in how teacher evaluation is implemented, CPS could greatly improve the effectiveness of the REACH Students program. The end result would be invested and engaged teams of teachers and administrators who work together to not only improve instruction, but to improve the teacher evaluation system that supports it.

Leadership Lessons

By examining a program in which I participate, I have become more intentional in my personal approach to teacher evaluation. I create regular opportunities to engage my assistant principal in norming observations, and we share feedback strategies with each other. This practice ensures that on the school level, our observation scores - and more importantly our vision for high quality instruction - are aligned and shared. Ensuring validity in our school's teacher observation scores builds trust and encourages teacher engagement in the continuous improvement process.

In order to promote teacher trust in the evaluation program, I have regular discussions with teachers about the benefits of the Framework for Teaching and the opportunities for professional growth that the formal and informal observations present. I remain as transparent as possible about all ratings and focus all discussion, both in pre-observation conferences and post-observation conferences, on those components of the framework which offer opportunities for growth, based on the observation. Focusing on opportunities rather than on judgement supports increased trust and professional learning.

Beyond my school, however, there is much more work to be done. Addressing the issue of student test scores is controversial, and not deeply understood. As I consider my own practice and the processes of the program, I see enormous opportunity for growth, which would not only support stronger teacher development, but would support improved trust in the system at large.

CPS greatly improved teacher evaluation in 2012 with the REACH Students teacher evaluation program. Teachers and administrators now engage in professional conversations about teacher practice, based on a shared vision of high quality teaching, the CPS Framework for Teaching. While this new program is a significant improvement over the former checklist

system, it offers insufficient support for teachers and administrators to engage in professional conversations which foster reflection, self-assessment, and professional growth. Additionally, the use of VAM scores in teacher evaluation not only undermines teacher trust, but it undermines the validity of the program, and has the potential to push the highest rated teachers out of the neediest schools. Lastly, teachers have no voice in how the evaluation program is implemented, which further undermines their trust in the program. By investing in human capital, ameliorating the support system for teachers and administrators, giving teachers a voice, and eliminating the use of VAM scores in teacher evaluation, CPS could dramatically improve teacher evaluation in Chicago and thus help all teachers to engage in continuous improvement for the benefit of all students.

REFERENCES

American Educational Research Association. (2015, November 11). AERA Issues Statement on the Use of Value-Added Models in Evaluation of Educators and Educator Preparation Programs. https://www.aera.net/Newsroom/News-Releases-and-Statements/AERA-Issues-Statement-on-the-Use-of-Value-Added-Models-in-Evaluation-of-Educators-and-Educator-Preparation-Programs.

Assembly, I. G. (2010). Performance Evaluation Reform Act: Laws of Illinois, 2010. (Illinois Public Act 096-0861). Springfield, IL: Author.

Berube, C. T. (2004). Are standards preventing good teaching? *Clearing House*, *77*(6), 264–267.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal*, *9*(2), 27.

Braun, H. I. (2005). Using student progress to evaluate teachers: A primer on value-added models. Princeton, NJ: Educational Testing Service.

Briggs, D. & Domingue, B. (2011). Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles Unified School District teachers by the Los Angeles Times. Boulder, CO: National Education Policy Center.

Bryk, A. S., & Schneider, B. (2003). Trust in schools: A core resource for school reform. *Educational leadership*, *60*(6), 40-45.

Burgess, S., Rawal, S., & Taylor, E. S. (2019). *Teacher peer observation and student test scores: Evidence from a field experiment in English secondary schools*. Working Paper. Cambridge, MA.

Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on

    individual academic achievement. *The Journal of Educational Research*, *90*(5), 269-277.

Chicago Public Schools. (2020, January 6*). CPS Stats and Facts.* https://cps.edu/About_CPS/At-

    a-glance/Pages/Stats_and_facts.aspx

Close, K., Amrein-Beardsley, A., & Collins, C. (2019). Mapping America's teacher evaluation

    plans under ESSA. *Phi Delta Kappan*, *101*(2), 22-26.

Cochran-Smith, M. & Lytle, S. (2006) Troubling images of teaching in no child left behind.

    Harvard Educational Review, 76(4), 668-697.

Corcoran, S. P. (2010). Can teachers be evaluated by their students' test scores? should they be?

    the use of value-added measures of teacher effectiveness in policy and practice.

    Education Policy for Action Series. *Annenberg Institute for School Reform at Brown*

    University (NJ1).

Corcoran, S., & Goldhaber, D. (2013). Value added and its uses: Where you stand depends on

    where you sit. *Education Finance and Policy*, *8*(3), 418-434.

Danielson, C. (2010). Evaluations that help teachers learn. *Educational Leadership, 68*(4), 35–

    39.Assessing teachers: a conversation with Charlotte Danielson. (2012). *Principal*, *91*(4),

    26-27.

Danielson, C. (2011). *Enhancing professional practice: A framework for teaching*. ASCD.

Danielson, C. (2012). It's your evaluation -- collaborating to improve teacher practice. *Education*

    Digest, *77*(8), 22–27.

Danielson, C. (2016). Charlotte Danielson on rethinking teacher evaluation. *Education Week*, *35*(28), 20-24.

Darling-Hammond, L. (2007). Race, inequality and educational accountability: the irony of "No Child Left Behind." *Race, Ethnicity & Education*, *10*(3), 245–260.

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

Darling-Hammond, L. (2012). Value-added teacher evaluation: the harm behind the hype. *Education Week*, *31*(24), 32–24.

Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. Teachers College Press.

Davey, M., & Yaccino, S. (2012). Teachers end Chicago strike on second try. The New York Times.

DeNisco, A. (2014). Removing weak links. *District Administration*, *50*(2), 50–54.

Desimone, L. M., & Pak, K. (2017). Instructional coaching as high-quality professional development. *Theory Into Practice*, *56*(1), 3–12.

Donaldson, M. L., Woulfin, S., LeChasseur, K., & Cobb, C. D. (2016). The structure and substance of teachers' opportunities to learn about teacher evaluation reform: promise or pitfall for equity? *Equity & Excellence in Education*, *49*(2), 183–201. https://doi-org.nl.idm.oclc.org/10.1080/10665684.2016.1144831

Drago-Severson, E. (2006). Learning-oriented leadership. *Independent school*, *65*(4), 58-64.

Drago-Severson, E. (2009). *Leading adult learning: Supporting adult development in our schools*. Corwin Press.

Fullan, M. (2011). *Choosing the wrong drivers for whole system reform* (pp. 3-4). Melbourne Centre for Strategic Education

Fulmer, Gavin W. 2012. Value-added measures in education: what every educator needs to know. *Science Education* 96 (2): 364–66.

Gardner, D. National Commission on Excellence in Education (1983). A nation at risk: The imperative for educational reform. An open letter to the American people. A report to the nation and the Secretary of Education.

Hanushek, E. A. (1979). Conceptual and empirical issues in the estimation of educational production functions. *Journal of human Resources*, 351-388.

Jiang, J. Y., Sartain, L., Sporte, S. E., Steinberg, M. P., & Society for Research on Educational Effectiveness, (2014). The impact of teacher evaluation reform on student learning: success and challenges in replicating experimental findings with non-experimental data.

Jiang, J., Sporte, S., & Luppescu, S. (2015) Teacher perspectives on evaluation reform. *Educational Researcher*, 44(2), 105 – 116

Jiang, J. Y., & Sporte, S. E. (2016). Teacher Evaluation in Chicago: Differences in Observation and Value-Added Scores by Teacher, Student, and School Characteristics. Research Report. *University of Chicago Consortium on School Research*.

Jiang, J., Sporte, S., & Luppescu, S. (2015) Teacher perspectives on evaluation reform.

    Educational Researcher, 44(2), 105 – 116.

Little, J. W. (2012). Professional community and professional development in the learning

    centered school. *Teacher Learning that Matters* (pp. 42-64). Routledge.

Liu, Y., Visone, J., Mongillo, M. B., & Lisi, P. (2019). What matters to teachers if evaluation is

    meant to help them improve? *Studies in Educational Evaluation*, *61*, 41–54. https://doi-

    org.nl.idm.oclc.org/10.1016/j.stueduc.2019.01.006

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, *70*(3),

    14–19.

Marzano, R. J. (2013). How to show student learning. *Educational Leadership*, *71*(2), 82–83.

Marzano, R. J., & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model

    for teacher growth and student achievement*. ASCD.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., & Hamilton, L. (2003). Evaluating value-added

    models for teacher accountability. Santa Monica, CA: The RAND Corporation.

Moskal ACM, Stein SJ, Golding C. Can you increase teacher engagement with evaluation simply

    by improving the evaluation system? *Assessment & Evaluation in Higher Education*.

    2016;41(2):286-300. doi:10.1080/02602938.2015.1007838.

Myers, M. A., and D. Curtiss. 2003. Failing the equity test. Principal Leadership. 3, no. 2.

    (October): 70–73.

Paige, M., Amrein-Beardsley, A., & Close, K. (2019). Tennessee's national impact on teacher evaluation law & policy: an assessment of value-added model litigation. *Tennessee Journal of Law and Policy*.

Patton, M. Q. (2008). *Utilization-focused evaluation*. Sage publications.

Quinn, A. E. (2014). Looking at the bigger picture with Dr. Robert Marzano: teacher evaluation and development for improved student learning. *Delta Kappa Gamma Bulletin*, *81*(1).

Reid, D. B. (2018). How schools train principals to use new teacher evaluation systems: A comparative case study of charter school and traditional public school principals. *Journal of School Choice*, *12*(2), 237–253.

https://doiorg.nl.idm.oclc.org/10.1080/15582159.2018.1447737

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

Saeki, E., Segool, N., Pendergast, L., & Embse, N. (2018). The influence of test-based accountability policies on early elementary teachers: School climate, environmental stress, and teacher stress. *Psychology in the Schools*, *55*(4), 391–403. https://doiorg.nl.idm.oclc.org/10.1002/pits.22112

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. *Research Report*. Consortium on Chicago School Research. 1313 East 60th Street, Chicago, IL 60637.

Smith, J. B., Lee, V. E., & Newmann, F. M. (2001). Instruction and achievement in Chicago elementary schools. Improving Chicago's Schools.

Sporte, S. E., Stevens, W. D., Healey, K., Jiang, J., & Hart, H. (2013). *Teacher evaluation in practice: Implementing Chicago's REACH Students*. University of Chicago Consortium on Chicago School Research. 1313 East 60th Street, Chicago, IL 60637.

Voices from the field: strategies from school leaders to improve REACH. The Chicago Public Education Fund. (2016), 1–16.

Wagner, T., & Kegan, R. (2006). *Change leadership: A practical guide to transforming our schools*. San Francisco: Jossey-Bass.

Toch, T. (2009). The Perils of Merit Pay. *Phi Delta Kappan*, *91*(2), 99

APPENDIX A: Chicago Public Schools Framework for Teaching

The CPS Framework for Teaching

CHICAGO PUBLIC SCHOOLS | CPS

Adapted from the Danielson *Framework for Teaching*

Domain 1: Planning and Preparation

Component	Unsatisfactory	Basic	Proficient	Distinguished
1a: Demonstrating Knowledge of Content and Pedagogy Knowledge of: • Content Standards Within and Across Grade Levels • Disciplinary Literacy • Prerequisite Relationships • Content-Related Pedagogy	Teacher demonstrates little to no knowledge of relevant content standards within and/or across grade levels. Teacher demonstrates no knowledge of the disciplinary way of reading, writing and/or thinking within the subject area. Teacher demonstrates little understanding of prerequisite knowledge important to student learning of the content/skills. Teacher's plans reflect little or no understanding of the range of pedagogical approaches suitable to student learning of the content/skills being taught.	Teacher demonstrates knowledge of the relevant content standards within the grade level but displays lack of awareness of how these concepts relate to one another and/or build across grade levels. Teacher demonstrates some knowledge of the disciplinary way of reading, writing, and/or thinking within the subject area. The teacher demonstrates some understanding of prerequisite learning, although knowledge of relationships among topics may be inaccurate or incomplete. Teacher's plans reflect a limited range of pedagogical approaches suitable to student learning of the content/skills being taught.	Teacher demonstrates knowledge of the relevant content standards, within and across grade levels. Teacher demonstrates knowledge of the disciplinary way of reading, writing, and/or thinking within the subject area. Teacher demonstrates accurate understanding of prerequisite learning and relationships among topics and concepts. Teacher's plans reflect a range of effective pedagogical approaches suitable to student learning of the content/skills being taught.	Teacher demonstrates knowledge of the relevant content standards within the grade level and across grade levels, as well as how these standards relate to other disciplines. Teacher's plans demonstrate extensive knowledge of the disciplinary way of reading, writing, and/or thinking within the subject area. Teacher demonstrates deep understanding of prerequisite learning and relationships among topics and concepts. Teacher's plans include a range of effective pedagogical approaches suitable to student learning of the content/skills being taught and anticipate student misconceptions.
1b: Demonstrating Knowledge of Students Knowledge of: • Child and Adolescent Development • The Learning Process • Students' Skills, Knowledge, and Language Proficiency • Students' Interests and Cultural Heritage • Students' Special Needs and Appropriate Accommodations/ Modifications	The teacher demonstrates little to no understanding of how students learn and does not attain information about levels of development. Teacher does not gather knowledge about students' backgrounds, cultures, prior knowledge, skills, language proficiencies, learning styles, interests, and special needs and does not indicate that such knowledge informs teacher's practice.	The teacher displays generally accurate knowledge of how students learn and attains information about levels of development for the class as a whole. Teacher gathers some knowledge about some students' backgrounds, cultures, prior knowledge, skills, language proficiencies, learning styles, interests, and special needs.	The teacher demonstrates an understanding of the active nature of student learning and attains information about levels of development for groups of students. Teacher purposefully gathers information from several sources about most students' backgrounds, cultures, prior knowledge, skills, language proficiencies, learning styles, interests, and special needs.	The teacher demonstrates an understanding of the active nature of student learning and attains information about levels of development for individual students. Teacher purposefully and continually gathers information from several sources about all students' individual backgrounds, cultures, prior knowledge, skills, language proficiencies, learning styles, multiple intelligences, interests, and special needs.

Component	Unsatisfactory	Basic	Proficient	Distinguished
1c: Selecting Learning Objectives • Clarity of Objectives • Sequence and Alignment of Objectives • Balance of Objectives	Learning objectives are not standards-based, are unclear, or are stated as activities rather than as student learning outcomes, prohibiting a feasible method of assessment. Teacher does not sequence and align learning objectives to build toward deep understanding and mastery of the standards. Objectives reflect only one type of learning and/or only one discipline.	Learning objectives are partially standards-based, clear, written in the form of student learning outcomes, aligned to methods of assessment, and/or are only written for the class as a whole. Teacher demonstrates an attempt to sequence and align some standards-based learning objectives, but does not build toward deep understanding or mastery of the standards. Objectives reflect more than one type of learning, but teacher has made no attempt at coordination of the disciplines.	Learning objectives are standards-based, clear, written in the form of student learning outcomes, aligned to methods of assessment, and varied to account for the needs of groups of students. Teacher sequences and aligns standards-based learning objectives to build toward deep understanding and mastery of the standards. Objectives reflect several different types of learning and invite opportunities for coordination within and across the disciplines.	Learning objectives are standards-based, clear, written in the form of student learning outcomes, aligned to methods of assessment, and varied in whatever way is needed to account for individual students' needs. Teacher sequences and aligns standards-based objectives to build toward deep understanding, mastery of the standards, and meaningful authentic application. Objectives reflect several different types of learning and provide multiple opportunities for coordination and integration within and across the disciplines.
1d: Designing Coherent Instruction Design Incorporates: • Knowledge of Students and Their Needs • Learning Tasks • Materials and Resources • Instructional Grouping • Lesson and Unit Structure	Teacher does not coordinate knowledge of content, students, and resources to design units and lessons. Learning tasks are not aligned to objectives. Tasks are not cognitively challenging and do not require students to provide evidence of their reasoning. There is no evidence of scaffolding and differentiation for students to access the content/skills. The progression of tasks is not coherent and has unrealistic time allocations. Units and lessons do not include grade-appropriate levels of texts and/or other materials and do not represent a cognitive challenge. The lesson or unit does not have a recognizable structure and makes no use of instructional groupings to support the learning objectives.	Teacher coordinates some knowledge of content, students, and resources to design units and lessons. Learning tasks are partially aligned to objectives. Tasks are cognitively challenging, designed for the class as a whole, and occasionally require students to provide evidence of their reasoning. There is some evidence of scaffolding and differentiation for some students to access the content/skills. The progression of tasks in units and lessons is not always coherent, and some time allocations are unrealistic. Units and lessons include grade-appropriate levels of texts and other materials that represent a moderate cognitive challenge. The lesson or unit has a recognizable structure with some evidence of instructional groupings that partially support the learning objectives.	Teacher coordinates knowledge of content, students, and resources to design units and lessons. Learning tasks are aligned to objectives. Tasks are cognitively challenging, designed for groups of students, and require students to provide evidence of their reasoning. There is evidence of scaffolding and differentiation for most students to access the content/skills. The units and lessons are paced appropriately. Units and lessons include grade-appropriate levels of texts and other materials, representing a cognitive challenge. The lesson or unit has a clear structure with intentional and structured use of instructional groupings that support the learning objectives.	Teacher coordinates in-depth knowledge of content, students, and resources (including technology) to design units and lessons. Learning tasks are aligned to objectives. Tasks are cognitively challenging for individual students and require students to provide evidence of their reasoning. There is evidence of scaffolding and differentiation for all students to access the content/skills. The units and lessons are paced appropriately. Units and lessons include grade-appropriate levels of texts and other materials so every student can access the content/skills. The lesson or unit has a clear structure that incorporates student choice, allows for different pathways of instruction aligned with diverse student needs, and uses instructional groupings intentionally.

Component	Unsatisfactory	Basic	Proficient	Distinguished
1e: Designing Student Assessment • *Congruence with Standards-Based Learning Objectives* • *Levels of Performance and Standards* • *Design of Formative Assessments* • *Use for Planning*	The plan for student assessment is not aligned with the standards-based learning objectives identified for the unit and/or lesson. Assessments contain no criteria or descriptors aligned to student expectations. Teacher does not select or design formative assessments that measure student learning and/or growth. Teacher does not use prior assessment results to design units and lessons.	The plan for student assessment is partially aligned with the standards-based learning objectives identified for the unit and/or lesson. Assessments do not clearly identify and/or describe student expectations. Teacher selects or designs formative assessments that measure only part of student learning or growth. Teacher uses prior assessment results to design units and lessons that target the class as a whole.	The plan for student assessment is aligned with the standards-based learning objectives identified for the unit and/or lesson. Assessment methodologies are designed or adapted for groups of students as needed. Assessments clearly identify and describe student expectations and provide descriptors. Teacher selects and designs formative assessments that accurately measure student learning and/or growth. Teacher uses prior assessment results to design units and lessons that target groups of students.	The plan for student assessment is aligned with the standards-based learning objectives identified for the unit and lesson. Assessment methodologies have been designed or adapted for individual students as needed. Assessment criteria are thorough, describe high expectations for students, and provide clear descriptors. Teacher's formative assessments are complex, well designed or selected, and tailored for individual students, when necessary, in order to measure varying degrees of each student's learning and growth effectively. Teacher uses assessment results to design units and lessons that target the diverse needs of every student.

Domain 2: The Classroom Environment

Component	Unsatisfactory	Basic	Proficient	Distinguished
2a: Creating an Environment of Respect and Rapport • *Teacher Interactions with Students* • *Student Interactions with Other Students*	Patterns of classroom interactions, both between the teacher and students and among students, are mostly negative and disrespectful. Interactions are insensitive and/or inappropriate to the ages and development of the students, and the context of the class. The net result of interactions has a negative impact on students emotionally and/or academically.	Patterns of classroom interactions, both between the teacher and students and among students, are generally respectful but may reflect occasional inconsistencies or incidences of disrespect. Some interactions are sensitive and/or appropriate to the ages and development of the students, and the context of the class. The net result of the interactions has a neutral impact on students emotionally and/or academically.	Patterns of classroom interactions, both between the teacher and students and among students, are friendly and demonstrate caring and respect. Interactions among students are generally polite and respectful. Interactions are sensitive and appropriate to the ages and development of the students, and to the context of the class. The net result of the interactions has a positive impact on students emotionally and academically.	Patterns of classroom interactions, both between the teacher and students and among students, are highly respectful, reflecting genuine warmth and caring. Students contribute to high levels of civility among all members of the class. Interactions are sensitive to students as individuals, appropriate to the ages and development of individual students, and to the context of the class. The net result of interactions is that of academic and personal connections among students and adults.
2b: Establishing a Culture for Learning • *Importance of Learning* • *Expectations for Learning* • *Student Persistence* • *Student Ownership of Learning*	The teacher creates a classroom culture that reflects a lack of teacher and/or student commitment to learning. The teacher fosters a culture in which there is little or no investment of student energy into the task at hand. The teacher does not convey that practice or perseverance is expected or that it results in student success. The teacher's expectations for student learning are medium to low, with high expectations reserved for only a few students. Students do not show interest in task completion or quality.	The teacher creates a classroom culture that reflects some teacher and/or student commitment to learning. The teacher fosters a culture in which there is some investment of student energy into the learning task at hand. The teacher conveys that student success is the result of natural or innate ability rather than practice and perseverance. Teacher conveys high learning expectations for some students. Students indicate that they are interested in completion, rather than quality, of a task.	The teacher creates a classroom culture that reflects teacher and student commitment to learning. The teacher fosters a culture in which classroom interactions indicate learning and hard work. The teacher conveys that with practice and perseverance students can reach desired goals. Teacher conveys high learning expectations for all students. Students take some responsibility for their learning by indicating that they want to understand the content or master the skill rather than simply complete a task.	The teacher creates a classroom culture that reflects a shared belief in the importance of learning and hard work. The teacher conveys high learning expectations for all students and develops structures that enable practice and perseverance for each individual student. Students assume responsibility for high-quality work by persevering, initiating improvements, addressing critiques, making revisions, adding detail and/or helping peers.

Component	Unsatisfactory	Basic	Proficient	Distinguished
2c: Managing Classroom Procedures • *Management of Instructional Groups* • *Management of Transitions* • *Management of Materials and Supplies* • *Performance of Classroom Routines* • *Direction of Volunteers and Paraprofessionals*	Ineffective classroom routines and procedures lead to loss of much instructional time. The teacher's management of instructional groupings, transitions, and/or the handling of materials and supplies are ineffective, leading to disruption of learning. There is little evidence that students know or follow established routines. The teacher does not give volunteers and/or paraprofessionals clearly defined duties.	Partially effective classroom routines and procedures lead to loss of some instructional time. The teacher's management of instructional groupings, transitions, and/or the handling of materials and supplies are inconsistent, leading to some disruption of learning. With regular guidance and prompting, students follow established routines. The teacher assigns duties to volunteers and/or paraprofessionals during portions of class time.	Effective classroom routines and procedures lead to minimal loss of instructional time. The teacher's management of instructional groupings, transitions, and/or the handling of materials and supplies are consistent and effective with little disruption of learning. With minimal guidance and prompting, students follow established classroom routines. The teacher engages volunteers and/or paraprofessionals with clearly defined duties that support student learning.	Effective classroom routines and procedures maximize instructional time. The teacher orchestrates the environment so that students contribute to the management of instructional groupings, transitions, and/or the handling of materials and supplies without disruption of learning. Students follow classroom routines without the teacher's prompting. Teacher productively engages volunteers and/or paraprofessionals in tasks that make a substantive contribution to student learning and are well integrated into the classroom community.
2d: Managing Student Behavior • *Expectations and Norms* • *Monitoring of Student Behavior* • *Fostering Positive Student Behavior* • *Response to Student Behavior*	Teacher has not established standards of conduct. Teacher engages in little to no monitoring of student behavior. Teacher does not use positive framing to model and reinforce appropriate behavior or redirect inappropriate student behavior. Teacher does not respond to students' inappropriate behavior, or the response is negative, repressive, and/or disrespectful.	Teacher has established standards of conduct, but there is inconsistent implementation so some student behaviors challenge the standards of conduct. Teacher inconsistently uses positive framing to model and reinforce appropriate behavior and redirect inappropriate student behavior. Teacher tries, with uneven results, to monitor student behavior. Teacher's response to students' inappropriate behavior is inconsistent and is sometimes disrespectful.	Teacher has established standards of conduct with consistent implementation so most students follow the standards of conduct most of the time. Teacher monitors student behavior against established standards of conduct. Teacher uses positive framing to model and reinforce appropriate behavior and redirect inappropriate student behavior. Teacher's response to students' inappropriate behavior is consistent, proportionate, respectful to students, and effective.	Teacher and students establish and implement standards of conduct. Students follow the standards of conduct and self-monitor their behaviors. Teacher's monitoring of student behavior is subtle and preventive. Teacher uses positive framing to model and reinforce positive behavior for individual students. Teacher's response to students' inappropriate behavior is sensitive to individual student needs and respects students' dignity.

Domain 3: Instruction

Component	Unsatisfactory	Basic	Proficient	Distinguished
3a: Communicating with Students • *Standards-Based Learning Objectives* • *Directions for Activities* • *Content Delivery and Clarity* • *Use of Oral and Written Language*	Teacher neither clearly communicates standards-based learning objective(s) to students nor addresses their relevance to learning. Teacher's directions and procedures are confusing to students. Teacher's explanation of content is unclear or inaccurate; explanations do not connect with students' knowledge and experience. Teacher's spoken and written language is unclear and incorrect. Vocabulary is vague, incorrect, or inappropriate for the students' ages and levels of development, leaving students confused.	Teacher does not communicate the standards-based learning objective(s) to students or does not address their relevance to learning. Teacher clarifies directions and procedures after initial student confusion. Teacher's explanation of content contains minor errors, and/or some portions are clear while other portions are difficult to follow; explanations occasionally connect with students' knowledge and experience. Teacher's spoken and written language is unclear or incorrect. Vocabulary is limited or inappropriate for the students' ages or levels of development.	Teacher clearly communicates standards-based learning objective(s) to students and addresses their relevance to learning. Teacher clearly communicates directions and procedures. Teacher's explanation of content is clear and accurate, and connects with students' knowledge and experience. Teacher's spoken and written language is clear and correct. Vocabulary is appropriate for the students' ages and levels of development.	Teacher clearly communicates standards-based learning objective(s). Teacher guides students to articulate the relevance of the objective(s) to learning. Teacher clearly explains directions and procedures, and anticipates possible student misunderstanding. Teacher's explanation of content is thorough, accurate, and clear, enabling students to develop a conceptual understanding of content while making connections to their interests, knowledge, and experience; students contribute to extending the content by explaining concepts to their classmates. Teacher's spoken and written language is expressive, and builds on students' language development and understanding of content. Vocabulary is appropriate for the students' ages and levels of development; students contribute to the correct use of academic vocabulary.
3b: Using Questioning and Discussion Techniques • *Low- and High-Level Questioning* • *Discussion Techniques and Explanation of Thinking* • *Student Participation*	Teacher does not ask questions, or all questions are of low cognitive challenge, requiring only short, specific, right or wrong answers. Questions are not developmentally appropriate. Teacher does not require students to construct viable arguments. Questions are asked in rapid succession with no "wait time" for student processing and response. The discussion is irrelevant to the content under study or predominantly in the form of recitations, with the teacher mediating all questions and answers. Teacher accepts all contributions without asking students to explain or provide evidence for their thinking. Few students are listening and responding to questions and answers from either the teacher or peers.	Teacher's questions lead students through a single path of inquiry where answers are seemingly pre-determined, with few high-level or open-ended questions. Questions are not always developmentally appropriate. Questions are asked with limited "wait time." Teacher attempts to create a discussion among students to engage with the content under study, with uneven results. Teacher sometimes requires students to provide evidence of their thinking or construct viable arguments based on evidence. Some students are listening and responding to questions and comments from their teacher or peers, and/or a few students dominate the discussion.	Teacher's questions are low- and high-level, open-ended, and developmentally appropriate, requiring student thinking, and promoting understanding. Teacher creates a genuine discussion among students, providing adequate "wait time" for students to engage with the content under study and stepping aside when doing so is appropriate. Teacher requires students to provide evidence of their thinking and construct viable arguments based on evidence. Most students are listening and responding to questions and answers from their teacher and peers. Teacher ensures that most voices are heard in the discourse.	Teacher uses a variety of low- and high-level, open-ended, and developmentally appropriate questions to challenge students cognitively, advance high level thinking and discourse, and promote metacognition. Teacher's discussion techniques enable students to engage each other in authentic discussions about the content under study. Students formulate questions and challenge one another using viable arguments based on evidence. All students are listening and responding to questions and answers from their teacher and peers. Students themselves ensure that all voices are heard in the discourse.

Component	Unsatisfactory	Basic	Proficient	Distinguished
3c: Engaging Students in Learning • *Alignment of Learning Objectives* • *Task and Text Complexity* • *Scaffolding and Differentiating Instruction* • *Pacing and Sequencing* • *Grouping*	Tasks do not align with standards-based learning objectives. Tasks and/or text require only rote responses, do not result in active engagement, and do not challenge student thinking. Teacher does not scaffold or differentiate instruction so that all students access complex, grade-level, and/or developmentally appropriate text or tasks. The teacher's pacing of the lesson is too slow or rushed, and tasks are not sequenced to build students' depth of understanding. The teacher's grouping of students is unintentional and inhibits student mastery of the content/skills.	Tasks partially align with standards-based learning objectives. Tasks and/or text minimally challenge student thinking, and result in active engagement of only some students while allowing others to be passive or merely compliant. Teacher occasionally scaffolds and/or differentiates instruction so that only some students access complex, grade-level, and/or developmentally appropriate text and/or tasks. The teacher's pacing of the lesson is partially appropriate, and/or tasks are partially sequenced to build students' depth of understanding. The teacher's grouping of students is intentional but does not lead to student mastery of the content/skills.	Tasks align with standards-based objectives. Tasks and text are complex and challenge student thinking, resulting in active engagement of most students. Teacher scaffolds and differentiates instruction so that most students access complex, grade-level and/or developmentally appropriate text and tasks. The teacher's pacing of the lesson is appropriate, and tasks are sequenced to build students' depth of understanding. The teacher's grouping of students is intentional and leads to student mastery of the content/skills.	Tasks align with standards-based learning objectives and are tailored so virtually all students are intellectually engaged in challenging content. Tasks and text are complex and promote student engagement through inquiry and choice. Students contribute to the exploration of content. Teacher scaffolds and differentiates instruction so that all students access complex, grade-level, and/or developmentally appropriate text and/or tasks. The teacher's pacing of the lesson is appropriate, and tasks are sequenced not only to build students' depth of understanding, but also to require student reflection and synthesis of the learning. Teacher's grouping of students is intentional and students serve as resources for each other to achieve mastery of the content/skills.
3d: Using Assessment in Instruction • *Monitoring of Student Learning with Checks for Understanding* • *Assessment Performance Levels* • *Student Self-Assessment* • *Feedback to Students*	Teacher does not use formative assessment, neither to check for completion of work nor to monitor progress and check for student understanding. Students cannot explain the criteria by which their work will be assessed and do not engage in self- or peer-assessment. Teacher's feedback is absent or of poor quality.	Teacher sometimes uses formative assessment to monitor progress toward student understanding of the learning objectives and/or teacher checks for completion of work rather than student understanding of the learning objectives. Students can explain some of the criteria by which their work will be assessed; few engage in self- or peer-assessment. Teacher's feedback is general and/or doesn't advance specific learning.	Teacher uses formative assessment during instruction to monitor progress and to check for student understanding of the learning objectives. Students can explain the criteria by which their work will be assessed; some of them engage in self- or peer-assessment. Teacher provides accurate and specific feedback to individuals and groups of students to advance learning.	Teacher fully integrates formative assessment into instruction, and uses it to monitor progress, and to check for understanding for individual students. Students can explain, and there is some evidence that they have contributed to, the criteria by which their work will be assessed. Students self- and peer-assess to monitor their progress. Teacher and students provide individualized feedback that is accurate, specific, and advances learning.
3e: Demonstrating Flexibility and Responsiveness • *Response to Student Needs* • *Teacher's Persistence* • *Lesson Adjustment*	The teacher ignores students' questions, needs, learning styles and interests; when students have difficulty learning, the teacher blames them or their home or the external environment for their lack of success. The teacher makes no attempt to adjust instruction during the lesson to meet student needs, even when students don't understand the content or have not mastered the skill.	The teacher attempts to accommodate students' questions, needs, learning styles and interests during instruction and accepts responsibility for the success of all students. When formative assessments show a need for intervention or enrichment, teacher attempts to adjust instruction during the lesson, but impromptu adjustments are ineffective.	Teacher accommodates students' questions, needs, learning styles and interests during instruction. The teacher accepts responsibility for student learning and persists in seeking approaches for all students. When formative assessments show a need for intervention or enrichment, teacher makes effective impromptu adjustments to instruction.	Teacher seizes opportunities to enhance learning, building on a spontaneous world or local event and/or student interests. Teacher persists in adjusting instruction so individual student misunderstandings or advanced needs are successfully accommodated. When formative assessments show a need for intervention or enrichment, teacher makes effective impromptu adjustments that individualize instruction for students.

Domain 4: Professional Responsibilities

Component	Unsatisfactory	Basic	Proficient	Distinguished
4a: Reflecting on Teaching and Learning • Effectiveness • Use in Future Teaching	Teacher does not describe whether or not a lesson or unit was effective or achieved its objective, or teacher misjudges the success of a lesson or unit and its impact on student learning. Teacher is not able to analyze the aspects of his/her practice that led to the outcome of the lesson and the impact on student learning. Teacher makes no suggestions about how practice could have been altered to improve the lesson or future similar lessons.	Teacher accurately describes whether or not a lesson or unit was effective but does not describe the extent to which it achieved its objective or its impact on student learning. Teacher is able to analyze some aspects of his/her practice that led to the outcome of the lesson and the impact on student learning. Teacher makes general suggestions about how a lesson could have been altered to improve the lesson or future similar lessons.	Teacher makes an accurate assessment of a lesson's or unit's effectiveness and the extent to which it achieved its objective and impact on student learning and can provide evidence to support the judgment. Teacher analyzes aspects of his/her practice that led to the outcome of the lesson and the impact on student learning. Teacher makes specific suggestions about how a lesson could have been altered to improve the lesson or future similar lessons.	Teacher makes an accurate assessment of a lesson's or unit's effectiveness and the extent to which it achieved its objective and its impact on student learning, citing many specific examples and evidence. Teacher is able to analyze many aspects of his/her practice that led to the outcome of the lesson and the impact on student learning. Teacher offers specific alternative practices, complete with the probable success of each aspect of practice could have improved the lesson or future similar lessons.
4b: Maintaining Accurate Records • Student Completion of Assignments • Student Progress in Learning • Non-Instructional Records	Teacher has a disorganized system or no system for maintaining information on student completion of assignments, student progress in learning, and non-instructional records, resulting in errors and confusion.	Teacher has a rudimentary system for maintaining information on student completion of assignments, student progress in learning, and non-instructional records, requiring frequent monitoring to avoid errors.	Teacher has an organized system for maintaining information on student completion of assignments, student progress in learning, and non-instructional records, requiring little monitoring to avoid errors.	Teacher has a detailed system for maintaining information on student completion of assignments, student progress in learning, and non-instructional records, requiring no monitoring for errors. Students contribute information and participate in maintaining the records.

Component	Unsatisfactory	Basic	Proficient	Distinguished
4c: Communicating with Families • Information and Updates about Grade Level Expectations and Student Progress • Engagement of Families as Partners in the Instructional Program • Response to Families • Cultural Appropriateness	Teacher does not communicate with families to inform them of class activities, to convey an individual student's progress, nor to solicit the family's support in relationship to grade level expectations. Teacher does not engage families in the instructional program. Teacher does not respond to families' concerns, neither professionally nor in a timely manner. Teacher's communication with families is not conveyed in a culturally appropriate manner. Teacher's communication is one-way, not interactive. Teacher only communicates with families for behavioral concerns or about academic failure.	Teacher rarely communicates with families to inform them of class activities, to convey information about an individual student's progress, and/or to solicit the family's support in relationship to grade level expectations. Teacher engages families in the instructional program only for attendance at activities or events. Teacher sometimes responds to families' concerns in a professional and/or timely manner. Teacher's communication with families is not always appropriate to the cultural norms of students' families. Teacher's communication is interactive only when a family member initiates communication.	Teacher regularly communicates with families in a two-way interactive manner to discuss class activities, individual student's progress, and to solicit the family's support in relationship to grade level expectations. Teacher meaningfully engages families as partners in the instructional program (e.g. through classroom volunteering, working at home with their child, and involvement in class projects in and out of school). Teacher responds to families' concerns professionally and in a timely manner. Teacher's communication with families is appropriate to the cultural norms and needs of the students' families.	Teacher frequently communicates with families to convey information about class and individual activities, individual student's progress and to solicit and utilize the family's support in student learning. Teacher meaningfully and successfully engages families as partners in the instructional program (e.g. through class and home volunteering, working at home with their child, involvement in class and school projects in and out of school, and parent workshops and training). Teacher responds to families' concerns professionally and in a timely manner, providing resources and solutions to address the concerns. Teacher's communication with families is sensitive to cultural norms and needs, with students contributing to the communication as appropriate.
4d: Growing and Developing Professionally • Enhancement of Content Knowledge and Pedagogical Skill • Collaboration and Professional Inquiry to Advance Student Learning • Participation in School Leadership Team and/or Teacher Teams • Incorporation of Feedback	Teacher rarely, if at all, engages in professional growth activities to enhance content knowledge or pedagogical skill to improve practice. Teacher rarely meets with colleagues or resists meeting and collaborating with colleagues. Teacher rarely, if ever, makes an effort to participate in team-based professional inquiry to advance student learning. Teacher does not volunteer to participate in a leadership and/or teaching team. Teacher resists feedback from colleagues or administrators and makes no effort to incorporate it to improve practice and advance student learning.	Teacher participates in required professional growth activities to enhance content knowledge or pedagogical skill to a limited extent and/or when it is convenient, using new knowledge inconsistently to improve practice. Teacher reluctantly meets to collaborate with colleagues, and reluctantly provides or accepts support to/from them. Teacher participates in team-based professional inquiry to advance student learning and participates in a leadership and/or teaching team only when invited. Teacher accepts feedback from colleagues and administrators with some reluctance, using feedback inconsistently to improve practice and advance student learning.	Teacher seeks opportunities for professional growth to enhance content knowledge and pedagogical skill and uses new knowledge to improve practice. Teacher regularly collaborates with and provides/receives support to/from colleagues. Teacher participates actively in team based professional inquiry that advances student learning and makes substantial contribution to the school leadership team and/or grade-level/content/department teaching team. Teacher accepts and consistently uses feedback from colleagues and administrators to improve practice and advance student learning.	Teacher initiates opportunities for professional growth and makes a systematic effort to enhance content knowledge and pedagogical skill of self and colleagues. S/he uses new knowledge to improve practice of self and colleagues. Teacher invites meetings and initiates collaborations with colleagues. Teacher provides and accepts collegial support and feedback to/from colleagues. Teacher participates in and facilitates professional inquiry with school team to advance student learning and serves on a leadership and/or teaching team. Teacher welcomes and uses feedback from a variety of stakeholders (e.g. colleagues, administrators, students, parents, external education partners) to improve practice and advance student learning.

Component	Unsatisfactory	Basic	Proficient	Distinguished
4e: Demonstrating Professionalism • Integrity and Ethical Conduct • Advocacy • Decision-Making • Compliance with School and District Regulations • Attendance	Teacher does not hold student or required school information confidential, and is dishonest in professional and student/family interactions. The teacher is not alert to students' needs, contributes to practices that result in some students being ill-served, and does not ensure that students are prepared to succeed in school, college, career, and life. Teacher makes decisions and recommendations that are based on self-serving interests. Teacher does not comply with school and district regulations. Teacher does not have a responsible or professional attendance record.	Teacher holds student and required school information confidential, and is honest in professional and student/family interactions most of the time. The teacher's attempts to serve students are inconsistent, and unknowingly contribute to some students being ill-served. Teacher sometimes ensures students are prepared to succeed in school, college, career, and life. The teacher's decisions and recommendations are based on limited, though genuinely professional, considerations. Teacher complies minimally with school and district regulations, doing just enough to get by. Teacher has a minimally responsible or professional attendance record.	Teacher always holds student and required school information confidential, and displays high standards of honesty, integrity, and confidentiality in interactions with colleagues, students, and the public. The teacher is active in serving students and works to ensure that all students receive a fair opportunity to succeed in school, college, career, and life. The teacher maintains an open mind in decision-making and helps ensure that such decisions are based on professional considerations. Teacher complies fully with school and district regulations. Teacher has a responsible and professional attendance record.	Teacher has the highest standards of integrity, always holds student and required school information confidential, and is honest in professional and student/family interactions. The teacher is proactive in serving students, seeking out resources when needed. The teacher makes a concerted effort to challenge negative attitudes or practices so that all students, particularly those traditionally underserved, are honored in the school and prepared for success in school, college, career, and life. Teacher takes a leadership role in decision-making for the school and helps ensure that such decisions are based on the highest professional considerations. Teacher complies fully and takes a leadership role with school and district regulations. Teacher has a responsible and professional attendance record.

APPENDIX B: Chicago Public Schools Teacher Evaluation Checklist

CLASSROOM TEACHER VISITATION

This form is required. It should be used in conjunction with the "Post-Observation Framework Feedback Form" (Form 5B).

Teacher's Name: _____ Room_____ Date _____

_____ School_____ Subject/Grade_____

(Place a (✓) or brief comment in the appropriate column.)

	Strength	Weakness	Does Not Apply
I. Instruction			
a) Provides written lesson plans and preparation in accordance with the objectives of the instructional program.	_____	_____	_____
b) Establishes positive learning expectation standards for all students.	_____	_____	_____
c) Periodically evaluates pupils' progress and keeps up-to-date records of pupils' achievements.	_____	_____	_____
d) Applies contemporary principles of learning theory and teaching methodology.	_____	_____	_____
e) Draws from the range of instruction materials available in the school.	_____	_____	_____
f) Exhibits willingness to participate in the development and implementation of new ideas and teaching techniques.	_____	_____	_____
g) Provides bulletin board and interest areas reflective of current student work.	_____	_____	_____
h) Exhibits and applies knowledge of the curriculum content related to subject area and instructional level.	_____	_____	_____
i) Shows evidence of student performance and progress.	_____	_____	_____
II. School Environment			
a) Establishes and maintains reasonable rules of conduct within the classroom consistent with the provisions of the Student Code of Conduct.	_____	_____	_____
b) Maintains attendance books, lesson plan, seating chart(s), and grade book accurately.	_____	_____	_____
c) Uses recommendations and suggestions from conference and special education staffings.	_____	_____	_____
d) Encourages student growth in self discipline and positive self-concept.	_____	_____	_____
e) Makes students aware of the teacher's objectives and expectations.	_____	_____	_____
f) Practices fairness in teacher-pupil relationships.	_____	_____	_____
g) Exhibits an understanding and respect for students as individuals.	_____	_____	_____
III. Professional and Personal Standards			
a) Presents an appearance that does not adversely affect the students' ability to learn.	_____	_____	_____
b) Demonstrates proper diction and grammatical usage when addressing students.	_____	_____	_____
c) Uses sound and professional judgment.	_____	_____	_____

IV. Local School Unit Criteria

a) CPS Framework for Teaching and related process

b) _____

c) _____

Comments: _____

_____

APPENDIX C: As Is Diagnostic Tool

"As Is" 4 C's Analysis for Improving Teacher Evaluation Implementation

Context
- The school is a part of a large district, in which there is an adversarial relationship between teachers and the district.
- Teachers object to the use of student test scores in their evaluation.
- There have been many leadership changes, resulting in a lack of focus.
- There has been inconsistent and shifting support for teacher evaluation.

Conditions

+
- Principals have access to manuals, rubrics, and professional development once per year
- Principals must be certified, which includes norming with external evaluators
- Principals are motivated by having effective teachers
- Principals evaluate all teachers at least once per year

−
- The highest rated teachers work in schools with the lowest poverty levels
- Teachers in the schools with the highest poverty levels receive lower ratings
- All observations are evaluative
- One size fits all rubric does not work for all practitioners
- There is a wide range of vision of the different performance levels between evaluators

Culture

+
- Teachers believe that their evaluators are capable, and they recognize the value in the REACH Framework, but observations cause stress and teachers feel that the potential benefits are not outweighed by the costs
- All observations are seen as evaluative
- Teachers are protective of their practice

−
- Lack of trust between teachers and the district
- Lack of trust between teachers and school administration

The teacher evaluation process has the support of teachers and administrators, but its effectiveness is undermined by a lack of trust, consistency, and commitment.

Competencies

+
- Principals are trained on the evaluation process and Framework for Teaching
- Some principals are trained on adult learning and coaching
- Experienced Instructional Effectiveness Specialists support principal evaluation in a limited manner

−
- Principals and teachers receive scant ongoing training and support
- Learning opportunities for evaluators are not consistent across the district
- Principals tend to rate their teachers higher than external evaluators
- Little changes year to year, resulting in continuation of habits and practices

Appendix D: To Be Diagnostic Tool

"To Be" 4 C's Analysis for Improving Teacher Evaluation Implementation

Context

- The school is a part of a large district, in which there is a productive and trusting relationship between teachers and the district.
- Leadership is stable from year to year, with a clear focus on factors that contribute to success.
- There is ongoing, job-embedded support for teachers and evaluators.
- Teacher evaluation no longer includes Value Added Model scores.

Culture

- There is a high level of trust between teachers and school and district administration.
- Teachers view observations as opportunities for growth rather than a judgement on their practice.
- Teachers are willing to share and receive feedback on their practice.

The teacher evaluation process supports professional growth, develops trust, and motivates teachers to engage in continuous improvement.

Conditions

- Principals have access to manuals, rubrics, and ongoing training.
- Principals are motivated by having effective teachers.
- Principals evaluate all teachers at least once per year.
- Some observations are non-evaluative, and teachers engage in peer observation.
- Specific rubrics exist for each type of practitioner.
- There is a unified vision among teachers and evaluators of the different performance levels
- Teachers participate in the decision making process about teacher evaluation content and implementation.

Competencies

- The district trains evaluators continuously on the evaluation process and rubric as well as best practices in adult learning and coaching.
- Teachers engage in self-assessment and reflection in a process of continuous improvement.
- Experienced Instructional Effectiveness Specialists support evaluators on an ongoing basis.
- Principals participate in professional learning communities with other administrators and evaluators to share practices and norm ratings.
- Principals engage in continuous improvement to make adjustments to teacher evaluation based on yearly feedback from evaluators and teachers.
- Teachers participate in professional learning communities to learn the process, better understand the frameworks, and conduct peer observations.
- Teachers offer feedback on the process and make suggestions for improvement.
- CTU and CPS work together to gather feedback and make adjustments from year to year.