

Use of Teacher Evaluation for Improving Instruction: A Mixed Methods Investigation

In High-Poverty New Jersey Charter High Schools

By

Christian Mathews

A dissertation submitted to the

Graduate School of Education

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Education

Graduate Program in Teacher Leadership

written under the direction of

Dr. Angela O'Donnell

Dr. Linda Reddy

Dr. Drew Gitomer

New Brunswick, NJ

October 2017

USE OF TEACHER EVALUATION FOR IMPROVING INSTRUCTION

Abstract

In the state of New Jersey and nationally, teacher evaluation is a current focus of education reform efforts (Hallinger et al., 2014; Schulman, 2014). A well-executed teacher evaluation system produces data regarding educator effectiveness. It is the intention of the 2012 *AchieveNJ* teacher evaluation reform law that teacher evaluation data be used to make human capital management decisions, including those associated with the instructional improvement of teachers (NJ DOE, 2014). These teacher evaluation systems (TES) are intended to increase student achievement by improving the quality of teacher practice (Schulman, 2014). Although this fact is established in the literature, it is seen that educators report individual, small group, and school-level instructional improvement programs that are disconnected from their practice and TES data usage (Danielson, 2011; Garret et. al, 2001; Hallinger et. al., 2014; Schulman, 2014; Wilson & Berne, 1999).

Using a mixed methods design this investigation seeks to examine how teacher evaluation data in a school reform grant (School System Improvement Project (SSIP); PIs Reddy, Kettler & Kurz) funded by the US Department of Education - Teacher Incentive Fund Program awarded to Rutgers University) is utilized for instructional improvement. School partners in the project include 14 high needs charter schools across the state of New Jersey. The sample included 155 teacher survey respondents, 16 semi-structured teacher interviews, and 11 semi-structured administrator interviews. The Rutgers School Systems Improvement Project (SSIP) is a focused on enhancing Human Capital Management Systems and Teacher Evaluation Systems to improve principal and teacher effectiveness to maximize student growth and achievement.

There were two aims of this mixed methods study. The first aim was to examine school administrators' use (nonuse) of teacher evaluation data to inform instructional improvement in

USE OF TEACHER EVALUATION FOR IMPROVING INSTRUCTION

educators. The second aim of the study was to examine teachers' experiences with school administrators' use of teacher evaluation data to support instructional improvement. These aims were addressed via qualitative semi-structured interviews of school administrators, semi-structured focus group interviews of teachers, and quantitative analysis of survey data exploring teacher attitudes, beliefs, and experiences regarding teacher evaluation..

In addressing the aims of the study, four assertions emerged as supported by both the qualitative and quantitative data results:

- Administrators and teachers report data are used to inform collaborative and specific feedback most frequently.
- Teachers and administrators report TES data are most used to inform individual instructional improvement, specifically informing Professional Development Plans and future development goals.
- Teachers and administrators report Teacher Evaluation Systems data are sometimes used to inform group and whole school level improvement, but barriers interfere with its use.
- The barriers of time, knowledge, and training of both administrators and teachers are interfering with data use for instructional improvement.

The results of this investigation inform the Rutgers School Systems Improvement Project in current and future support of partner schools in the implementation of teacher evaluation and instructional improvement programs. Recommendations to address these assertions within the problem of practice include improving administrator training, improving administrator onboarding programs, enhanced support surrounding data based decision making, and future study to confirm findings and establish clear generalizability to the broader field.

USE OF TEACHER EVALUATION FOR IMPROVING INSTRUCTION

Dedication

This work is dedicated to my loving wife Christine and loving dog Cody for their constant support and endless tolerance; to all those educators past and present who have lead me down this road and showed me the way at Queen of Peace Grammar School, Queen of Peace High School, Felician University, Seton Hall University, and Rutgers University; to my colleagues at the Rutgers University School Systems Improvement Project without whom none of this would be possible, especially my director for opening the door to this world. Finally, with special thanks to the members of my committee and dissertation chair who have committed themselves above and beyond the call of duty to mentor me through this dissertation process: Dr. O'Donnell, Dr. Gitomer, and Dr. Reddy.

Table of Contents

1. Title page	
2. Abstract.....	i-ii
3. Dedication.....	iii
4. Table of contents.....	iv
5. Tables.....	v
6. Chapter I: Problem Statement.....	1-9
7. Chapter II: Literature Review.....	10-35
8. Chapter III: Methodology.....	36-45
9. Chapter IV: Results.....	46-69
10. Chapter V: Conclusion	70-80
11. References	81-89
12. Appendix A: Interview protocol.....	90-91
13. Appendix B: Focus group protocol.....	92-93
14. Appendix C: Examined TEES-T Survey tions.....	94-95
15. Appendix D: Remaining quantitative survey tables.....	96-108

USE OF TEACHER EVALUATION FOR IMPROVING INSTRUCTION

Tables

Table 1. <i>Sample of Administrators</i>	40
Table 2. <i>Sample of Teachers</i>	41
Table 3. <i>Descriptive Statistics for Charter High School 1-4 Quantitative Sample TEES-T</i>	65
Table 4. <i>Average response of sample for each subscale</i>	65
Table 5. <i>Item: The evaluation feedback provided information for professional development opportunities</i>	66
Table 6. <i>Item: The evaluation system helped to improve the quality of instruction</i>	66
Table 7. <i>Item: The evaluation system improved my professional growth</i>	67
Table 8. <i>Item: The evaluation system informed changes in my classroom practice</i>	67
Table 9. <i>Item: The evaluation system helped to improve student learning</i>	68
Appendix D Tables 10-42.....	96-108

CHAPTER 1

Problem Statement

The focus of educational reform has shifted over the past two decades from school-level accountability measures that were central to No Child Left Behind (2001) to enhancing individual teacher practice as a way to increase student learning and outcomes (Hallinger, Heck, & Murphy, 2014). Teacher-centered reforms have increased recently as a result of federal programs such as Race to the Top (2009). As a result, states began requiring school districts to implement new teacher evaluation systems that seek to improve the instructional quality of teachers. However, institutional factors in a school or district may prohibit teacher evaluation from fulfilling this purpose. This is because teacher evaluation can only be an effective method of guiding instructional improvement in circumstances where it is implemented with high fidelity (Ho & Kane, 2013). Teachers will ultimately be the final point of execution of reform in schools and teachers are the key to reform efforts over the past decade by leaders, administrators, politicians, and other reformers. The push in teacher evaluation reform reflects this shift in emphasis.

As school reform initiatives change their focus, numerous teacher evaluation efforts connected to broader school improvement programs were initiated throughout the United States (Hallinger et al., 2014). Major reforms of teacher evaluation systems have been undertaken within the past decade in forty three states (National Council on Teacher Quality, 2012; Rotherham & Mitchel, 2014). There are still more states considering major teacher evaluation reform initiatives within the next few years (Hallinger et al., 2014). Teacher evaluation is and will continue to be a major component of education reform efforts.

New Jersey joined the national reform movement in 2012 and passed the AchieveNJ law, which initiated a new Teacher Evaluation System (TES). Further, the law standardized a teacher effectiveness formula in which classroom observations count towards 70-80% (depending on teacher placement) of a teacher's annual evaluation score. Teachers receive three to four observations that are combined to produce a score using one or more state approved evaluation instruments such as the *Stronge Model* (2013), *Marshall Teacher Evaluation Rubric* (2003), *Danielson Framework for Teaching* (2011) and numerous others. These models were designed to focus on highly effective, research-supported teaching strategies (NJDOE, 2014). AchieveNJ further urges districts to use teacher effectiveness determinations to inform human capital decisions, such as tenure and retentions, as well as to create instructional improvement plans and professional development programs (Schulman, 2014). Evaluating educators is not alone enough to influence student outcomes. Teachers need instructional improvement to adjust academic and behavior management strategies based on evaluation data (Darling-Hammond et. al, 2012; Donaldson & Papay, 2012 Ho & Kane, 2013; Rademaker, 2008; Taylor & Tyler, 2012). Ultimately, these systems should improve student achievement by improving the quality of teacher practice. The feedback teachers receive from evaluations, coupled with targeted professional development, helps them improve in controlled circumstances (Danielson, 2011; Hallinger et al., 2014; Reddy, Fabiano, Dudek, & Hsu, 2014). Although many researchers believe that teacher formative assessments are best practice, evidence suggests they are rarely implemented at schools (Hallinger et al., 2014).

Despite the intention of AchieveNJ and other teacher evaluation reform initiatives across the United States, there is reason to believe that teachers are not receiving support for instructional improvement that is linked to Teacher Evaluation System (TES) data. Teachers

often report a disconnect between the instructional improvement support they receive within their buildings and the evaluations and feedback they receive (Danielson, 2011; Garret et. al, 2001; Wilson & Berne, 1999). Furthermore, research on the use of teacher observation systems to date would suggest that while evaluations are taking place, they may not be used in ways that lead to improvement (Donaldson & Papay, 2012; Hamilton et al., 2008;). Organizational factors and implementation challenges within the school have an impact on how evaluation data are utilized (Hamilton et. al, 2008). A qualitative semi-structured interview study of ten principals, ten academic vice-principals, and 72 teachers found that teachers and principals report that teacher evaluation can be a beneficial process when conducted collaboratively and focused on a process of formative feedback (Hamilton et. al., 2008). In cases where this did not occur, instructional improvement was e far more limited (Hamilton et. al., 2008). In a quantitative survey review of over two thousand teacher responses and several hundred administrator responses across Georgia, Pennsylvania, and California addressing teacher accountability systems, researchers found that organizational factors and implementation challenges within the school have an impact on how evaluation data are utilized (Hamilton et. al, 2008). Notably, high quality professional development is key to teacher instructional improvement (Garret et. al, 2001; Little, 2006; Wood, 2007), but uneven utilization of evaluation system data in professional learning opportunities impacts the growth of teachers (Hamilton et. al., 2008).

There are significant implementation issues that impact the creation and utilization of teacher evaluation data. Implementation issues include too few observations conducted, inadequately trained observers, a lack of pre and/or post conferences, and a failure to utilize evaluation data for meaningful instructional improvement (Donaldson et. al, 2012; Hamilton et. al, 2008; Reddy et. al, 2014). Eliminating implementation issues leads to better results in

evaluation and observations as well as potentially alleviating some of the organizational drawbacks associated with a high stakes teacher evaluation system. Ho and Kane (2013) found that implementing observations with fidelity leads to stronger student outcomes and improved teacher instructional improvement. Challenges in implementing teacher evaluation systems (TES) lead to questions regarding data quality in a literature base that is currently shallow, requiring additional research (Hallinger et al., 2014). Many teachers report that teacher evaluation systems are disconnected from instructional improvement programs, and that the high stakes nature of teacher evaluation and its associated issues may substantially impact the quality of benefits received from TES (Kimball & Milanowski, 2009). Literature on professional development methods similarly report a disconnect between teacher evaluation and any type of instructional improvement programs on the individual or school wide level (Darling-Hammond, 2009; Garret et. al, 2001; Little, 2006; Wood, 2007). These challenges illustrate the need for additional study and for support and training addressing teacher evaluation implementation and data use. Teacher evaluation data has the potential to contribute to real change, but these challenges can strongly inhibit its potential benefits in schools.

It is important to note that teacher evaluation does not exist in a policy vacuum, without controversy, or as an unquestioned benefit. Many teachers report that teacher evaluation systems are disconnected from instructional improvement programs, and that the high stakes nature of teacher evaluation and its associated issues may substantially impact the quality of benefits received from TES systems (Kimball & Milanowski, 2009). Literature on professional development methods similarly report a disconnect from teacher evaluation and instructional improvement programs on the individual or school wide level (Darling-Hammond, 2009; Garret et. al, 2001; Little, 2006; Wood, 2007). These challenges illustrate the need for additional study

and for support and training addressing mentioned issues. It is of key importance to recognize the potential drawbacks of teacher evaluation when implementing or examining an evaluation focused intervention.

The primary aim of this investigation was to examine how teacher evaluation data in a large-scale school reform grant (School System Improvement Project (SSIP); PIs Reddy, Kettler & Kurz; funded by the US Department of Education - Teacher Incentive Fund Program (Grant S374A120060 awarded to Rutgers University) was utilized. School partners in the project included 14 high poverty charter schools across the state of New Jersey. This study specifically examined four NJ charter high schools partnering with the project. The Rutgers School Systems Improvement Project (SSIP) is a multi-faceted intervention support system for its partner NJ charter schools in the implementation of their teacher evaluation systems. For example, SSIP helps partners plan appropriately for instructional improvement within their schools, implementing an educator evaluation system, and then using the resulting data to make decisions in their organizations. The Rutgers University SSI Project is designed as an evaluation system intervention and partnership among the US Department of Education, Rutgers University, and local education agencies (in this particular circumstance high need charter schools). The Rutgers SSI Project provides training in implementing the teacher evaluation system, assistance in providing instructional improvement programs, training in data-based decision making, and overall support in applying teacher observation methods.

My role in the project was that of a Senior Program Administrator. I provided individualized support for teachers and administrators in all aspects of the teacher evaluation system and its implementation, including training in observation systems, training in data-based decision making, trouble-shooting, technical assistance, and consultation. Supporting teachers

and administrators through all aspects of teacher evaluation, from training to implementation to application of data, is essential to the success of the SSI project intervention. The theory of action behind the SSI project is that if the teacher evaluation system is implemented with fidelity to best practices and all parties receive proper support, and the data produced by the SSI project will be applied in decision making in instructional improvement on an individual and school-wide basis.

This study aimed to examine principals use of teacher evaluation data to inform instructional improvement programs. I define “Instructional Improvement” as programming provided by administrators with the intention of improving student learning outcomes (Coe, Aloisi, Higgins, & Major, 2015; Farrell, 2015). This includes individualized feedback, professional development programs for individual teachers, groups, and the whole school, professional improvement plans, and any other actions or programming that administrators design to guide teachers towards improving their instruction with the goal of improved student learning (Coe et al., 2015; Farrell, 2015). Student learning outcomes can ultimately be measured in a variety of ways, including but not limited to state-level summative evaluations, benchmark assessments such as the *Measures of Academic Progress* (MAP), and other classroom-level summative and formative assessments provided by the teacher (Coe et al., 2015; Farrell, 2015;; Reddy et al, 2013). The literature that exists details challenges that principals experience in using teacher evaluation data as a potential barrier to improving student achievement (Hallinger et al., 2014; Kimball & Malinowski, 2009). For example, administrators have expressed challenges and reservations in being able to effectively collect and utilize teacher evaluation data within their schools (Kimball & Milinowski, 2009). However, few studies describe systems that

implement formative systems fully. More research is needed to examine and bridge these gaps and barriers.

The problem of practice addressed in this study is whether SSI project partner charter schools are, in fact, utilizing teacher evaluation system data to inform instructional improvement. As the intervention enters its third year of implementation, it is crucial to the Rutgers School Systems Improvement Project that the significant data collected by the evaluation system is utilized to these ends. If not, it will seriously hamper the effectiveness of the teacher evaluation system and associated data-based decision making (Hallinger et al., 2014; Ho & Kane, 2013). On a broader level, this study will add to a currently shallow research base that explores the connection between the use of teacher evaluation data and instructional improvement. This study examines whether the supports provided by the SSI project in teacher evaluation are leading to data-based decision making for instructional improvement. This includes whether principals are using data as intended to guide teacher instructional improvement, considering the current climate of teacher evaluation reform.

There were two aims of the study. The first aim was to examine school administrators' use (nonuse) of teacher evaluation data to inform instructional improvement in educators. The second aim of the study was to examine teachers' experiences with school administrators use of teacher evaluation data to support instructional improvement.

Pilot Study Description

In a preliminary pilot study of four administrators and four teachers, results of qualitative interviews portend that while administrators believe in utilizing teacher evaluation data for instructional improvement, they may not be in fact effectively utilizing TES data to inform individual, small group, or whole school level instructional improvement programs. While

administrators use TES data for collaborative feedback on the individual level, few administrator responses revealed any significant usage of TES data in a rich and in-depth manner.

Potential barriers of time, talent, institutional support and resources emerge as factors that are creating this possible disconnect. More data are needed from administrators and especially teachers to create a fuller picture of the teacher evaluation system within partner schools.

Utilizing archived survey data regarding attitudes and beliefs regarding teacher evaluation system use and implementation within their schools will also create a fuller picture.

The pilot study was conducted with the same methodology of the qualitative interview portion. It informed this study in several key ways. It was utilized to refine and improve research questions and aims, specifically in the areas of potentials barriers. The pilot study identified potential barriers and assisted in the construction of a hypothesis to support research aims and questions. This hypothesis is that administrators utilize TES data for instructional improvement on an individual basis, and are able to provide specific/collaborative feedback within the TES system, but are less able to do so within the context of instructional improvement for small groups of teachers and whole school level programming. The pilot study was also utilized to test and refine the qualitative protocols for teachers and administrators. The pilot study allowed for several questions to be rephrased and five others to be removed from the protocol for clarity, as they did not produce any actionable data for analysis. The pilot study also provided a way to test data analysis methodology in Dedose (Dedose, 2016).

Significantly, the pilot study provided an initial codebook for the qualitative portion of this study. Preliminary codes were developed based upon review and analysis of qualitative interview data. It was refined for alignment with research aims. The final product of the pilot study produced a complete codebook that was utilized to analyze the qualitative data for this

study. The pilot study provided a codebook, research protocol, and hypothesis produced from testing and analysis of a live research sample that laid the foundation for this study moving forward.

Research Questions

AIM 1

- 1) Do school administrators use teacher evaluation data to inform their instructional improvement decisions?
- 2) What do school administrators report as opportunities and barriers when using teacher evaluation data for improving instruction and classroom management?
- 3) How do school administrators use teacher evaluation data to inform their decisions to improve educators' instruction and classroom management?

AIM 2

- 1) What do teachers report about administrators' use of teacher evaluation data to support their instructional improvement?
- 2) What do teachers report as opportunities and barriers when using teacher evaluation data for improving instruction and classroom management within their context?
- 3) What do teachers report about school administrator's use of teacher evaluation data to inform their decisions to improve educator's instruction and classroom management?

CHAPTER II

Literature Review

This literature review compiles and analyzes relevant research about Teacher Evaluation Systems (TES). In examining the literature, four major areas emerge that are relevant to this study. The first is teacher evaluation outcomes, which discusses how research findings are used to justify the value of teacher evaluation systems through evidence of improved student achievement. Second, literature on the relationship between instructional improvement and professional development that is linked to TES is essential to analyze for this study. The third section focuses on literature examining TES and instructional leadership. The study is framed by the concept of Data-Based Decision Making (DBDM), which comprises the fourth area in this literature review. Conceptually, principals and administrators receive data from the TES and are tasked with using these data to make decisions. It is a key understanding of the literature and a foundation of the Rutgers SSI Project that administrators need to use the TES to make decisions. TES can be used to make decisions about human capital or, and specifically for this study, instructional improvement. Using TES data to make decisions in schools is a central component of the rationale for this study.

Methodology

An effective literature review should include a comprehensive and systematic search for research (Lauer, 2004). The literature review began with a search for relevant literature through the Rutgers University Library database and *Google Scholar*. Search terms used were “Teacher Evaluation”, “Educator Evaluation”, “Data Based Decision-Making” “Instructional Improvement” and “Educator Evaluation Systems.” The search identified additional resources

utilizing the bibliographies and works cited from the sources described, and a recent literature review provided additional sources (Hallinger et. al., 2014). Recommendations for relevant literature were requested from SSI Project staff members, and the “resource sections” of SSI project designed trainings were searched for relevant articles. The search criteria used identified many of the most significant teacher evaluation studies currently available. Scholarship impact ratings were examined through *Google Scholar* and the frequency of citations of major studies was noted. Cross-checking sources found with Hallinger et al., (2014) matched many of the sources found in the literature search. The literature selected for inclusion within this review met Cresswell’s (2013) criteria for inclusion including scholarly impact, relevance in the field, appropriateness of publication source, and freedom from author bias.

Teacher Evaluation and Student Achievement Outcomes

TES scores can potentially correlate with student achievement under certain conditions. Several studies find a positive correlation between teachers’ scores on their teacher evaluations and student achievement score gains, though it is critically important to note that these studies are not in any way centrally examining teacher evaluation systems (Gallagher, 2004; Wright, Horn, & Sanders, 1997). In a qualitative study of one high poverty urban elementary charter school that serves 1,400 students, an evaluation system in which administrators were properly trained to conduct observations, conduct pre- and post- conferences, and met reliability testing criteria, demonstrated increased student achievement outcomes measured on state level testing and improved teacher observation scores on a subject specific teacher evaluation rubric. This study did not set out to demonstrate this item, and as such has methodological and inferential gaps that must be considered (Gallagher, 2004). The study also found that when teachers and observers were properly trained in the use of the observation rubric, and observers utilized the

data produced by the observation to produce high quality feedback reports and then used this information to inform instructional improvement programs, instructional outcomes for teachers and student achievement outcomes could potentially increase for individual classroom teachers (Gallagher, 2004). These outcomes include scores on state level standardized assessments in literacy and mathematics, as well as school-level benchmark assessments in these same content areas. The correlation is statistically significant in the area of literacy, but not in the area of mathematics. This study is limited as it is a case study of a single elementary school, but the findings are cited by many of the more recent studies in the field as identifying a potential connection between well-utilized teacher evaluation systems and student achievement scores. In a major large-scale study of the Tennessee state evaluation system, another study found that teachers' scores on the *Danielson Framework for Teaching* were moderately correlated to student achievement on state-wide standardized testing measures in mathematics and literacy (Ho & Kane, 2013). The Tennessee state evaluation system examined had introduced training for administrators, a required number of observations, pre and post conferencing, and the utilization of teacher data in instructional improvement decisions. This is significant because it suggests that for teacher evaluation systems to reflect any kind of correlation with student achievement outcomes, these factors of best practices in implementation, such as training, number of observations, and a pre/post conference, need to be in place (Wright et al., 1997). These studies point towards a potential correlation between teacher quality, as observed by teacher evaluation systems, and student achievement outcomes. It is important to note that the studies in question were not designed specifically to measure the impacts of TES system implementation, as such this inferential leap should be considered within this context.

Teacher evaluation systems have limitations as currently applied in the literature. It is important to qualify this literature by making clear that teacher quality is not the only significant factor impacting student achievement outcomes; externalities impacting students and communities outside of the classroom, such as economic conditions, as well as students' baseline scores entering a classroom are also described as significant factors (Hallinger et al., 2014). These studies demonstrate a correlation between scores from teacher evaluation system and student learning outcomes in literacy and mathematics for students, when the system is executed under conditions of fidelity with administrators who conduct an adequate number of observations, conduct these honestly, have been properly trained, solicit information in advance with post conferencing and provide feedback. One limitation of these studies is that they are either of a one-school sample, or reflect the potential externalities at play that impact student outcomes measured (Gallagher, 2004; Ho & Kane, 2013; Wright et al., 1997). This reflects a weakness of the literature that is often relied upon by policy makers to support teacher evaluation reforms. High stakes applications of evaluation systems still are problematic, and can lead to negative teacher outcomes when applied without integrity (Hallinger et al., 2014; Kimball & Milanowski, 2009). The research calls for teacher evaluation to be applied in a formative process, however, systems as applied do not effectively address this. The literature base reflects that in circumstances where teacher evaluation systems are not meeting quality criteria or best practices implementation in creating the data in the first place, they can be counterproductive, especially in terms of teacher attitudes and morale (Hallinger et al., 2014; Kimball & Milanowski, 2009). There is a clear disconnect between the potential benefits of teacher evaluation, and its high stakes implementation without the full formative components in place (Hallinger et al., 2014).

If there is a gap in implementation (such as observers do not use the entire instrument), the correlation between scores and growth in student achievement drops (Casbianca et. al, 2013; Danielson, 2011; Ho & Kane, 2013). This is of concern because, according to the NJ State Department of Education, some schools in the state use altered or unofficial versions of teacher observation rubrics or models (NJDOE, 2014). Studies also point to the importance of utilizing feedback during post conference and using this information to inform teachers' instruction as an essential component of their evaluation systems (Casbianca et. al, 2013; Gallagher, 2004; Ho & Kane, 2013; Wright et al., 1997). The research base suggests key issues to focus on include examining how observers are implementing their teacher evaluation systems and how they utilize the data produced, considering factors of training, number of observations, and use of data to inform instructional improvement programs.

The most comprehensive study of TES so far, The Measures of Effective Teaching (MET; 2013), was conducted by the Bill and Melinda Gates Foundations across multiple urban school districts. The MET Study is significant because of its large scale. It analyzed data on 3,000 teacher evaluations across numerous districts. This study found a correlation between teacher evaluation system scores and student learning outcomes as measured on state wide standardized testing in Mathematics and Literacy when the data is produced according to conditions of fidelity (Ho & Kane, 2013). Specifically, the MET study selected at least five teacher observations per year as a mark of demarcation in order to strengthen the correlation between evaluation scores and student outcomes. It also calls for trained observers and for the observation instruments to be used with integrity. Integrity is operationalized in the MET study in several ways. Observations must have a sufficient number of observers, be conducted by trained observers who, based on a reliability criteria test, are recorded within a 72 hour period,

and be conducted with intent to use the instrument authentically (no fake scores, bias, or favoritism). Considering the findings in this manner is significant because data from the MET study were used to influence current public policy across the United States and significant to this study because it demonstrates the implementation needs of a teacher evaluation system.

Although the MET study calls for a rigorous implementation of teacher observation, most states, including NJ, have required less extensive applications (Schulman, 2014). In fact, none of the state wide evaluation systems described in the research on this topic require more than four evaluations per year. In New Jersey, three is the norm for tenured teachers. Even the weaker correlation that occurs from doing three observations relies on the evaluation systems and associated measures being implemented with integrity and fidelity.

If the system is implemented with integrity, results of the TES system can successfully demonstrate links to student achievement outcomes. This is reflected in the research associated with the *Danielson Framework for Teaching*. In a study of 105 mid-career teachers in the Cincinnati school district (using the *Danielson Framework for Teaching*), teacher evaluations conducted according to best practices correlated to increased student achievement scores (Taylor & Tyler, 2012), pointing to a possible connection that could be confirmed with more research. Teacher evaluation observations had to be conducted in sufficient number, by trained observers, and in accordance with an established observation protocols. These were evidenced in state-wide student achievement testing in Ohio. This finding indicates that the *Danielson Framework for Teaching* and similar instruments show promise for improving instruction as evidenced by improvements in student achievement if they meet this basic level of adherence to these established factors.

The themes of “teacher evaluation as potentially important to increase student achievement,” “the importance of implementing teacher evaluation with fidelity to see these potential gains,” and “the potential correlation between the implementation of comprehensive teacher evaluation systems and student achievement outcomes” are revealed within the literature. A story emerges that when teacher evaluation systems are implemented according to best practices, they reveal that highly rated teachers evidencing effective and highly effective teaching practice demonstrate higher student outcomes and that teachers improve instructionally in an effective teacher evaluation system (Ho & Kane, 2013; Taylor & Tyler, 2012). The potential link between teacher evaluation and instructional improvement as reflected in Gallagher’s case study is explored in the next section. The literature reveals a potential correlation between implementing a teacher evaluation system with integrity and improved student learning outcomes on standardized state level assessments in Math and Literacy. However, there have only been a few major studies that examine this potential correlation that are frequently cited in the literature (Ho & Kane, 2013; Taylor & Tyler, 2012; Wright et al., 1997). These studies focus on justifying the existence of teacher evaluation data by links to student outcomes and focus on what produced a high quality teacher evaluation system. They do not exist in a deep enough base to say without any qualifications that a fully-implemented teacher evaluation system will lead to increased student learning, more study is needed (Hallinger et al., 2014). More large scale study is needed to explore this correlation, indeed the SSI project intervention overall is attempting to add to this literature base (Reddy, Dudek, & Hsu, 2014). Nonetheless, enough evidence currently exists of a potential link in the literature to continue study in this area and the literature shows enough potential to call for continued progress (Taylor & Tyler, 2012).

Teacher Evaluation and Instructional Improvement

Simply having a Teacher Evaluation System does not alone improve student achievement outcomes. This information must inform change for individuals and groups within school organizations (Ho & Kane, 2013; Tyler & Tyler, 2012). The previous section explored whether teacher evaluation systems correlate to student achievement gains, this section explores whether teacher evaluation systems are related to increased teacher instructional improvement. The following section must be considered in light of the fact that outside of the two studies cited in this paragraph; the studies in question were not specifically addressing similar research aims to this literature review. As such, an inferential leap is taken to consider the results on these studies through this lens.

High quality observations, defined as observations that include a pre and post conference, use a research supported rubric, and are conducted by a trained and tested- to- reliability criteria observer, provide the opportunity for higher quality instructional improvement for teachers overall (Ho & Kane, 2013). As discussed in the problem statement, many teachers receive instructional improvement programs and feedback they rate as disconnected from practice and not relevant to improving student learning (Garret et. al, 2001; Mahar & Strobert, 2010; Rademaker, 2008; Wilson & Berne, 1999).

Teacher instructional improvement programming is most effective when targeted towards the individual teacher's needs and made relevant (Garret et. al, 2001). In a case study of one school's instructional improvement program, teachers experienced positive results when connected to in-depth teacher evaluation, as reported in qualitative semi-structured interviews and focus groups (Rademaker, 2008). The more in-depth feedback from observations teachers receive and are able to discuss in a meaningful way, the more they find the information to be

meaningful and useful to their instructional improvement. In a mixed methods study of one large urban school district conducted by Mahar and Storbert (2010), teachers affirmed that evaluation scores that were more in-depth were more beneficial for their professional growth than evaluations lacking depth and clarity. This finding was observed through surveys, teacher self-reporting in interviews, and classroom observations. Teachers also reported a desire to engage in instructional improvement programming that addressed areas of need identified by their evaluations. Teachers self-reported that instructional improvement is most beneficial when connected to their evaluations. These findings appear to show an important conclusion in the literature relevant to this study, but this must be qualified that these three studies (Garrett et. al., 2001; Mahar & Storbert, 2010; Rademaker, 2008) utilize relatively small samples and rely primarily on observation of professional development, classroom teaching and teacher self-report in surveys and interviews. More study is needed to explore the link between teacher evaluation and instructional improvement.

In a dissertation study focused on teacher evaluation and connection to teacher professional development goals and Student Learning Objectives (SLO), a survey of over 1400 teachers and 41 administrators in a Virginia school district found that teachers generally do not see a connection between teacher evaluation systems and professional goals or SLOs as they are applied within their districts. Teachers reported that they would prefer TES data to be more relevant in their teacher evaluation. Teachers also reported less favorable opinions of the teacher evaluation system than their administrators, in applicability, data use, process, feedback, and overall impressions.

The best professional development is targeted at what teachers need, and the evaluation process. Observation feedback can be a valuable component of targeting professional

development to teacher needs. Other studies have findings in a non-evaluative context. In a meta-analysis of state-wide survey results on teacher professional development from Alabama, Georgia, Missouri, and Arizona, teachers self-reported that high quality professional development improves student learning (Darling-Hammond, 2009). Teachers also self-report that development programs that do not meet high quality criteria of relevance, connection to work, and connection to relevant classroom data do not demonstrate the same kind of connection to instructional growth, shown in a case study of one school district as well as the same meta-analysis of survey data (Darling Hammond, 2009; Rademaker, 2008). The research suggests that both individual teacher improvement plans and school wide professional development programs should be informed by teacher evaluation data. However, more evidence and study is also needed in this area. The research base appears to be focused on teacher self-report only, or is too narrow to be generalizable to the entire population (Casbianca et. al, 2013; Hallinger et al., 2014).

The *Danielson Framework for Teaching* and other teacher observation instruments can be utilized within a teacher evaluation system to help guide school wide professional development and instructional improvement. The instrument guides observers to give feedback for teachers relevant to each of the instruments' domains. A study conducted of an instrument used in the Rutgers School System Improvement Project demonstrated improvement in teacher performance on the Danielson Framework and the *Classroom Strategies Scale* (CSS) when teachers received specific, ongoing instructional feedback (Reddy et al., 2014). The CSS is a classroom observation rubric that serves as a primary evaluation instrument in the SSI project intervention. As an example for the *Framework for Teaching*, Danielson calls for multiple observations and in-depth discussion of results to lead to increased teacher practice outcomes (Danielson, 2011). These recommendations are aligned with what is known as best practices for

teacher feedback, specific and targeted instructional support leads to better teacher outcomes (Garrett et al., 2001; Rademaker, 2008). Teacher evaluation systems when utilized to inform legitimate professional development improve the gains that teachers bring to students. One study of 105 mid-career teachers in the Cincinnati public school district found an improvement in student achievement during the year that teachers were evaluated using the *Danielson Framework for Teaching*, and received feedback and participated in discussions of the evaluation data. Teachers were evaluated using the framework the same year students took a state-level assessment and then received feedback and instructional improvement tied to the evaluation data. The data examined were state-level standardized tests in Mathematics. These data were used to inform individual and school wide instructional improvement programs as a central component of the evaluation intervention. The data for the teachers were then examined six years before and six years after the evaluation. The positive effect on achievement was sustained in subsequent years as measured on the same state-level Mathematics standardized assessments (Taylor & Tyler, 2012). Of the literature base examined, this study demonstrates the clearest connection between both student achievement outcomes and teacher instructional improvement based on teacher evaluation. The limitation of the study is its focus on one subject area, only mid-career teachers, and outcomes being measured only in state-level standardized assessment. It is a promising study, but reflects a need for a more developed research base to solidify the justification for this connection.

Despite this, teachers have been reporting for a lengthy period of time that most of the feedback to improve they receive is disconnected from their self-perceived development needs (Cochran-Smith & Lytle, 1999; Garret et. al, 2001; Wilson & Berne, 1999; Wood, 2007). These studies primarily discuss instructional improvement in general, and when they discuss evaluation

data, it is in the context of data use in instructional improvement overall or in relation to a specific professional development intervention, such as a professional learning community (Wood, 2007). There is significant literature available describing what effective instructional improvement programs and professional development look like (Curry, 2008; Little, 2006; Wood, 2007). The research base is relatively limited in addressing the use of evaluation data to inform professional development programs specifically, or instructional improvement actions in the school overall.

Teacher data used in communities in practice can inform the goals of professional development programs and set goals for improvement (Curry, 2008; Darling-Hammond, 2009; Little, 2006). When connected to individual and school based development plans, teachers can evidence improvement in practice (Darling-Hammond, 2009). Professional development activities need to be integrated into the school-wide feedback system (Little, 2006). It also needs to be relevant to practice, otherwise even the best designed and intentioned improvement programs can fall apart (Curry, 2008). Meaningful professional development can be made relevant to teachers; teacher evaluation system data assists in building that connection. The Cincinnati study produces the most compelling evidence for direct correlation between evaluation that guides instructional improvement and gains in student outcomes (Taylor & Tyler, 2012). Other studies indirectly suggest a connection between teacher evaluation and instructional improvement either through teacher self-reports (Darling-Hammond, 2009; Garrett, 2001; Rademaker, 2008), examining professional development interventions (Cochran-Smith & Lytle, 1999; Curry, 2008; Garret et. al, 2001; Little, 2006; Wilson & Berne, 1999; Wood, 2007), and through examining teacher evaluation systems directly (Casbianca et. al, 2013; Ho & Kane, 2013). Good practice in professional development tied to TES includes instructional leaders who

guide it in data-based decision making. The largest identified gap is the need for more literature specifically examining the link between instructional improvement and teacher evaluation akin to the promising Cincinnati study.

Teacher Evaluation and Instructional Leadership

The state of the literature base in instructional leadership overall can best be described as disjointed. It is distributed among educational leadership, organizational leadership, change, management, and many other fields (Neumerski, 2013). The literature base involving teacher evaluation and instructional leadership is similarly dispersed among different areas. There is virtually no literature base specifically addressing the use of teacher evaluation in instructional leadership, and the literature that does exist is drawn from these different fields. As detailed in a meta-analysis of instructional leadership literature, many state and national leadership standards call upon instructional leaders to use evaluation data in their decision making process (Neumerski, 2013). Instructional leaders are identified by many relevant exterior organizations as being responsible for utilizing teacher evaluation data.

Instructional leaders are also tasked with guiding instructional improvement. In several meta-analyses of literature available in the instructional leadership field, numerous qualitative studies of educational institutions state the importance of the instructional leader or principal in the instructional improvement of teachers within the school (Firestone & Riehl, 2005; Neumerski, 2013). The literature base described by both of these analyses relate a much stronger qualitative base present in the literature focusing in on the instructional improvement role of the instructional leader. In these studies, teacher evaluation is a component of this instructional improvement role, but is not the focus of these studies overall. The instructional leader's role in

the teacher evaluation process as a potential positive force to lead to improved teacher and student outcomes is not clearly defined within the literature base described.

In a literature review article addressing teacher evaluation and instructional leadership specifically, the motivations behind teacher evaluation systems and their relevance to school reform and instructional leadership are explored (Firestone, 2014). Examining the major studies currently available in teacher evaluation, the analysis of the literature revealed that teacher evaluation systems have the greatest impact when leveraging trust and a formative development atmosphere. Firestone writes when discussing recent teacher evaluation reforms and major studies such as the MET, “Evaluation programs will only improve teaching where the programs are introduced with a great deal of trust between agents and authorizers” (Firestone, 2014, p. 105). Emphasized further in the review is the need for teacher evaluation to focus on teachers’ ability to utilize data to make decisions, and for administrators to emphasize the formative aspects of teacher evaluation. Of high relevance for the current study, the analysis concludes that administrators should use the formative aspects of teacher evaluation to guide teacher instructional improvement to maximize the positive aspects of teacher motivation while minimizing the negative (Firestone, 2014). The motivations of teachers in an evaluation need more study to be properly optimized to exploit the potential benefits of teacher evaluation while avoiding potential pitfalls associated with a high stakes system. It is a major danger for instructional leaders to apply the high stakes side of teacher evaluation while neglecting these formative motivations and aspects.

The instructional leader’s central role in the development of teachers through formative processes appears within the school change literature as well. For change to be as effective as possible instructional leaders should engage in a formative evaluation process (Fullan 2007;

Reeves, 2009). Change on an institutional scale also calls for development over numerous academic years, and utilizing data as a focus point of change (Fullan, 2007; Reeves, 2009). This call is supported in the literature, but primarily attached to qualitative studies. More large-scale study is needed in these areas to specifically tie together change, teacher evaluation, and instructional leadership in both quantitative and qualitative forms. The current study examines this link within a teacher evaluation centered change initiative.

The literature connecting teacher evaluation and instructional leadership is varied, and evidences gaps in the literature. Especially, more confirming studies are needed to solidify this link. However, the literature that exists illustrates a potential link between applying teacher evaluation in a formative manner and positive school level student achievement outcomes (Firestone 2014; Neumerski, 2013) and teacher engagement in instructional improvement programs opposed to only applying teacher evaluation systems from a financial incentive or a human capital management perspective (Firestone, 2014). The current study is intended to further explore this link.

Effective Instructional Leadership

It is important to consider not only the links in the literature between teacher evaluation and instructional leadership, but it is also important to consider what the literature portends as effective instructional leadership in and of itself. Considering the links between instructional leadership and effective instructional improvement programs is important to examine. It is important not only to consider whether instructional improvement programs are aligned with teacher evaluation data, but also what is the current state of best practice within the literature itself for effective instructional leadership.

Effective instructional leadership occurs through the prism of a larger human capital management system. Planning effective instructional improvement programs through the larger context of an effective evaluation, staff retention, and broader human capital system is envisioned as a way of leading to broader education reform (Odden, 2011). According to Odden, “That means that recruitment, selection, distribution, induction, professional development, performance management and evaluation, compensation, and career progression are all restructured to boost teacher and principal effectiveness in ways that dramatically improve instructional practice and student learning (Odden, , 2011, p. 11). Alignment of teacher evaluation with instructional improvement is of key importance to the effective execution of any educational enterprise. One major drawback of Odden’s assertion is there is currently limited literature supporting this assertion in large scale, empirical ways. The literature that makes this connection is explored in other sections of this literature review at length. Total quality management, as envisioned by Deming (1986), can be applied to effective instructional leadership. Effective instructional leaders follow the cycle of planning, doing, acting, checking, and reevaluating plans consistently throughout (Sallis, 1996). Deming’s vision of quality management in business is one of several paths in which human capital management can be applied to effective instructional leadership.

Practitioner change and sustaining practitioner change are important parts of instructional leadership. Both Fullan and Reeves assert that for an educational organization to be effective, it must be constantly changing and adapting to best serve learner needs. An important element to sustain practitioner change is the ability to constantly assess and adapt during the implementation of any change initiative (Fullan, 2007; Reeves, 2009). Instructional improvement and staff development are central not only to any individual change push, but in the overall growth and

sustainability of the institution at large. For change to be sustained it must be adaptable and consider organizational input, communication and information gathering are essential components of this process (Reeves, 2009). Historically, this organization is open to change, but at a very slow pace. It is crucial to build the kind of long-lasting cultural change necessary to positively impact a great number of organizations over time. At this point, it is necessary to bring in all elements of the school organization. If they are not included, the change will not be sustained (Fullan, 2007). Having the ability to develop staff to promote instructional improvement as well as adapt to organizational challenges is a central component of instructional leadership.

Best practices in adult learning must also be considered in any well-constructed instructional leadership program. Consulting teachers directly through a focus group is a powerful way to allow teachers the ability to contribute to the development of any change initiative or PD plan. Teachers benefit from opportunities to talk about teaching and feel as though their views are expressed, and taken seriously, within their buildings (Wilson & Berne, 1999). It is important to give adults a say in their learning, the proposed focus group as a needs assessment is an avenue to include them within the process. Also, according to Knowles et. al. (1998), it is important for adult learners to have a say in the construction of a learning experience. Related to this larger point about adult learners, teachers have less resistance to professional development experiences when they are included in shaping the process (Brookfield, 1989). Effective instructional leadership includes teacher input through the development process.

It is of key importance to consider the unique professional development (PD) needs of teachers and adult learners when designing a program or engaging in any aspect of instructional

leadership. There are potential pitfalls to successful PD programs that must be considered. The designer must encourage learning from experience, and consider the needs of teachers in professional development, relevancy, and learner transformation when designing a comprehensive PD plan for teachers (Curry, 2008; Garret et. al, 2001; Knowles, 1998; Woods, 20007). There are challenges in design that must be overcome to fully implement this type of professional development.

It is important to avoid a major potential pitfall that can impact professional development programs, development and instructional improvement programs being delivered in a “top-down” fashion limits teacher buy-in and engagement in the improvement process (Little, 2006; Woods, 2007). Teachers reported in a study of six professional learning communities that high-level professional learning directly leads to instructional improvement when delivered in a manner that they see as directly relevant to teacher practice (Curry, 2008). It is important to engage in professional development activities that are relevant and targeted towards the work teachers do in the classroom, in this case, improving instruction. In a self-reported survey of 1027 math and science teachers, findings indicate that professional development programs should be relevant to practice and content area oriented (Garret, Porter, Desimone, Birman, & Yoon, 2001). Teachers, and adult learners in general, need to see the practicality and relevance of the learning experience at hand. (Curry, 2008; Garret et. al, 2001; Knowles, 1998; Woods, 2007). One way of including teacher input is through experiential learning. Learning by doing is an effective strategy for gaining the interest and engagement of adult learners (Silberman & Auerbach, 2006). This is in line with a case study design examining the impact of an experiential professional learning experience in a World Languages department. The case study demonstrated a potential positive impact of experiential learning based professional learning (Burke, 2013).

Adult learners benefit from experience driven, relevant learning and should be learning through activities that reflect their practice. The activities of this development plan must include teachers as active participants.

Instructional leaders must consider that teachers want professional development that makes them better, transformed teachers, as evidenced through gains in student learning. Teachers need to implement and experience the positive results of the change to fully experience transformational learning (Silberman & Auerbach, 2006). The exact moment of transformation for the teacher may be intentional or unintentional. It may be sudden or incremental over a period of time (Mezirow, 2000). There are numerous ways in which transformation potentially occurs, but it appears that exposure and assimilation can be potential triggers. Observation and experience can lead to important shifts in thinking (Burke, 2013; Silberman & Auerbach, 2006). Engaging in a new environment or a new process can be also be a trigger for transformational learning (Burke, 2013). It is essential to create opportunities for teachers to help develop change for themselves and observe practices in others, engage collaboratively to improve practice, such as in data driven professional learning communities and peer observation groups.

Professional learning communities that are properly executed have the potential to be a catalyst for transformation of teacher practice (Little, 2006, Woods, 2007). As seen in a qualitative case study involving monthly meetings with teachers and observations, teachers reported that work groups can become a part of their practice and culture, as well as linking teacher evaluation, student achievement, and classroom practice together in their minds (Franke, Kazemi, Shih, Biagetti, & Battey, 2005). Schools in which teachers are supported in their use of classroom data and in their professional development with adequate time, expectations, and resources in professional learning communities saw gains in student outcomes in a variety of

different cases and measures (Franke et al., 2005; Little, 2006; Thibodeau, 2008; Wood, 2007).

Professional learning communities, given the right conditions, have the potential to engage teachers in the development process to ultimately improve student learning outcomes.

There are key elements for a successful professional learning community that must be considered and carefully put into place when utilizing this development strategy. Data must be presented in an understandable manner, using a structure such as a discussion protocol guiding discourse that serves as a catalyst for teachers to pursue their own professional growth with clear direction (Boudett, City, & Murnane, 2005; Curry, 2008; Little, 2006; Wood, 2007). The analysis of student data from English classes focused on areas of need is one potential strategy for solving this problem. Professional learning communities become communities of practice when centered on a careful analysis of student work (Himley & Carini, 2000; Little, 2006). Peer observations are another way of affecting teacher gains when carried out under certain conditions. Strategies such as utilizing self-reflection and observations among peers in a professional learning community are supported by teachers and can lead to teacher growth (Darling-Hammond, Wei, Andree, Richardson, Orphanos, & National Staff, D. C., 2009; Franke et. al., 2005; Wood, 2007). It is important to note, however, that any gains made as a result of peer observations are only maintained when teachers are supported with time and resources (Wood, 2007). Peer observations as a part of teacher communities can lead to improvements in teacher practice in all areas, including content instruction and behavior management (Franke et. al., 2005; Wood, 2007).

One way of providing a framework for peer observations is using protocols focusing on specific areas (Casabianca et al., 2013; Franke et. al., 2005; Thibodeau, 2008). In a case study of a district implementing a comprehensive evaluation system utilizing a content specific

observation protocol, a potential correlation was found between system implementation and improved student outcomes on standardized assessments (Casabianca et. al., 2013). By observing new practices and applying them within their own context, it provides an experience for potential transformational learning (Burke, 2013).

According to the available literature, best practices instructional leadership includes a teacher centered focus on professional development and instructional improvement. Rooting this learning within the best practices of professional development, adult learning, human capital management, and change processes has the potential to improve instructional learning with the ultimate goal of improving student learning. The literature is much deeper in establishing the positive effects of each individual practice cited, rather than including systematic studies investigating whether these processes working in concert or were connected to a data-driven Teacher Evaluation System would give increased impact. This is certainly an area for future study.

Data-Based Decision Making

Data-Based Decision Making is the utilization of data in school level decisions. The literature base is growing, but little exists that specifically links it to teacher evaluation data. The majority of the work that exists is focused on teacher utilization of classroom data (Morrison, 2009). This serves as a major framework for this study because as proposed, this study examines whether administrators are using teacher evaluation data to make instructional improvement decisions. The literature in this area tied to teacher evaluation is pulled from broader school level DBDM studies.

Data-driven instructional improvement can potentially improve both student outcomes and teacher perceptions of relevance, but again, the literature base is narrow on the connection

between TES and direct student achievement. Administrators and teachers who make decisions based upon data (including student achievement and TES) available within their schools demonstrate higher student and building level outcomes (Doran, 2015; Morrison, 2009). Professional development and instructional improvement have improved effectiveness outcomes for teacher ability and student achievement when data-based and focused on teams seen in a qualitative study of a teacher work group that operated within a medium sized urban school district (Morrison, 2009). The study included observations of teacher work groups and teacher semi-structured interviews, but is limited in its small sample size of twenty teachers. (Morrison, 2009). The results point to the importance of effective data based decision making. It could help teacher instructional improvement as seen on observation rubrics such as the *Danielson Framework for Teaching* and potentially lead to increases in student achievement rates seen on state testing as seen in a meta-analysis of Tennessee state evaluation reports and testing data. These are important outcomes for those concerned about teacher quality and student achievement (Doran, 2015). These studies suggest a potential correlation between data-based decision making, teacher quality, and student achievement but are far too limited to draw definitive conclusions. A significant increase in the research base is needed. Data are essential to making instructional improvement decisions. Differentiated development for teachers improved their efficacy and capability to improve student outcomes utilizing evaluation data (Doran, 2015). A case study of three district level data based decision making training interventions was conducted in schools where administrators utilized data to influence professional-learning opportunities, showed that there were increased student outcomes on state level mathematics and literacy assessments (Gold, 2005). Additionally teachers reported professional learning to be more

relevant to their practice in self-reported surveys (Gold, 2005). Significantly more study is needed to explore this potential link to draw more concrete conclusions.

The New Jersey and Rhode Island state evaluation systems were envisioned and designed with data based decision making in mind (Gist, 2013; Shulman, 2014). Both the New Jersey and Rhode Island statutes mention explicitly that professional development and instructional improvement is an intended use of teacher observations and teacher evaluation system scores (Gist, 2013; Shulman, 2014;). Research reports were written for both states utilizing qualitative input from a sample of teachers, administrators, and other stakeholders consulted within the process of the development of the evaluation systems. According to Shulman (2014) and Gist (2013) there is a need to encourage a cultural shift in the use of evaluation systems to promote teacher development. Current reform initiatives focused on Educator Evaluation Systems (EES) all establish teacher quality as a significant factor in student growth outcomes. The TES has a stated goal of leading teacher instructional improvement by using the information and scores that they produce to inform instructional improvement programs and targeted professional development As stated in Shulman's report on the NJ Teacher Evaluation System, a goal of this initiative is to see sustained teacher instructional improvement throughout the observations, aligning as implemented with Rutgers SSI Project partner schools.

Data-based decision making alone is not going to directly lead to instructional leadership. Data must be supplemented with qualitative supports and information to lead to school-wide instructional improvement (Elmore, 2003). Teachers must be trained in the effective use of data and have the necessary professional development in which to implement data analysis effectively. Otherwise, any data-based decision making occurring at the instructional level will not be effective. School leadership has a key role to play in supporting the use of data-based

decision making by providing effective, job imbedded professional development centered on the ability to utilize data to make classroom decisions (Elmore, 2003).

School leaders and observers are tasked with the implementation of the TES including the use of TES data. This includes deciding which observation measures to use, how to train staff, how long to conduct observations, how many informal professional development observations to conduct, and how to engage in authentic pre/post observation conferences with staff. Charter School leaders whose schools are participating in the Rutgers SSI Project are expected to utilize TES data to make decisions.

Making Sense of Policy in School Reform

Practitioners must be able to effectively navigate policy in order to make decisions utilizing best practices. Doehering and Volkmar (2016) make the case in their analysis that knowledge gaps in an area related to education research, in their examination Autism Spectrum Disorders, impact the ability of practitioners and policy makers to make decisions according to best practice. Knowledge gaps are identified as a key barrier. Similarly, Wagner and Sconyers (1996) in their Center on Families report identify knowledge gaps in policy and connections between practitioners, family members, school leaders, and policy makers as a potential derailing force in school reform efforts. Examining school reform history, practitioners have been integral to instances of success within various school reform movements (Reese, 2005). It is potentially significant to consider practitioners as central to school reform movements, such as the current endeavor in teacher evaluation systems and policy.

Conclusion

School leaders and observers are tasked with providing instructional improvement to teachers on an individualized and school-wide basis (Firestone & Riehl, 2005). The literature supports an instructional program that is relevant, authentic, and connected to teacher practice as well as connected to the TES process, and may lead to improved teacher and student outcomes under certain conditions (Cochran-Smith & Lytle, 1999; Curry, 2008; Doran, 2015; Firestone, 2014; Garret et. al, 2001; Gold, 2006; Little, 2006; Wilson & Berne, 1999; Wood, 2007). However, the literature also finds many teachers experience a disconnect between the TES and those factors they see as relevant to practice and the development and instructional improvement programs they actually engage in within their schools (Cochran-Smith & Lytle, 1999; Garret et. al, 2001; Wilson & Berne, 1999; Wood, 2007). Bridging the gap between evaluation data use and instructional improvement within schools is essential, thus providing a more high quality experience. Understanding more fully the motivations taken by school leaders and observers in implementing the data for instructional improvement on an individual and school-wide basis will allow for potential interventions and to align the programs closer to best practices of partner schools (Firestone, 2014).

Applying the evaluation system with integrity in the project and utilizing data-based decision making may produce gains in student outcomes (Ho et. al, 2013; Sparks, 2013). The evaluation data produced should be utilized to target teacher instructional improvement, also leading towards improved student outcomes (Cochran-Smith & Lytle, 1999; Curry, 2008; Doran, 2015; Garret et. al, 2001; Gold, 2005; Little, 2006; Morrison, 2009; Sparks, 2013; Wilson & Berne, 1999; Wood, 2007). It is also important to consider the potential drawbacks of applying

teacher evaluation systems in a high-stakes, non-formative, and incomplete manner ((Hallinger et al., 2014; Kimball & Milanowski, 2009).

There are significant gaps in the literature to be addressed by the current study and strengths that may be significant. A major strength in the literature is that several important studies and meta analyses point to a potential link between teacher evaluation systems and school improvement under controlled circumstances and when taking a formative approach (Firestone, 2014; Ho & Kane, 2013; Taylor & Tyler, 2012). A continued weakness in the literature is a lack of major studies to confirm this potential link (Hallinger et al., 2014), as well as a lack of definitive understanding of the instructional leaders role in the teacher evaluation process (Neumerski, 2013). Many of the teacher evaluation studies that do exist focus on student achievement, and do not specifically focus on the teacher's role or the administrators' role in best developing through the evaluation process. Understanding how SSI project partner administrators are utilizing teacher evaluation data will allow project staff to intervene to produce better results for teachers, administrators, and ultimately students throughout the project.

CHAPTER III

Methodology

In this mixed methods study, quantitative survey data of 155 responses out of 534 teachers, voluntarily participating from the fall semester of the 2016-2017 academic year, were analyzed, coinciding with a qualitative interview study of 11 administrators and 16 teachers connected to their school's teacher evaluation system that conducted at the beginning of the 2016-2017 school year. Quantitative data were collected in all SSI project partner schools with qualitative interviews collected from four similar Rutgers SSI project partnered high schools. The qualitative administrator interviews address the first research aim and its associated research questions. Sixteen interviews of teachers from these same four high schools each were conducted to address the second research aim and its associated research questions, to provide a triangulation perspective from teachers on administrator responses and to provide in depth data to address the second research aim and its associated research questions. Qualitative semi-structured interviews were selected as a methodology due to its ability to examine issues in greater depth and provide a richness of data to answer research questions that look at in-depth use and process of teacher evaluation systems (Merriam, 2009). Interviews were approximately thirty minutes long, and were conducted with administrators who had connection to the teacher evaluation system and instructional improvement within the school organization. Teachers had to be evaluated fully under the teacher evaluation system.

Questions focused on the implementation of the school's evaluation system as well as the application of teacher evaluation data within their building context. The protocol was flexible enough to allow interviewees to tell their story while allowing checks by the interviewer to prevent the conversation from moving too far from the intended task (Merriam, 2009). The semi-

structured interview protocol provided evidence of best practices, and was followed up after the interviews by member checks of the interview transcripts to ensure comfort of school partners and accuracy (Merriam, 2009). As an additional data source, reviews of school PD calendars and TES data reports were analyzed to provide another source to triangulate interview and focus group data.

The quantitative survey portion utilized results of the TEES-Teacher Form. This survey is currently used as a part of the Rutgers School System Improve Project to include teachers' perspectives in a comprehensive teacher evaluation system. The items, located in Appendix C, capture teachers' attitudes and beliefs about teacher evaluation and its implementation within their context (Reddy, Dudek, Kettler, Kurz, & Peters, 2016). Survey items are specifically analyzed in consideration to relevance to the research questions and aims, as discussed in Chapter IV. The survey results were examined to provide further information about the implementation of teacher evaluation within teachers' buildings as well as the follow up and instructional improvement taking place. It was examined to provide additional context for probing questions for teacher focus groups and potential codes for data analysis. The TEES-T is intended to provide school districts with a measure to assess teacher attitudes and experiences of teacher evaluation systems (Reddy et al., 2016). It was also developed under the operating hypothesis that teacher evaluation is an essential component of professional development (Finnegan, 2013). Questions include probes on the nature of engagement with the teacher evaluation system, whether it was collaborative, and how the teacher engages with the teacher evaluation system, complementing many of the qualitative interview questions to provide a more complete answer to the survey questions. Teachers were asked to appraise their teacher evaluation process, judge the value and usage of teacher evaluation system feedback (including

towards instructional improvement), perceived impact, and their experiences of implementation and effectiveness. Having a broader selection of teachers responding to the survey allowed for the broader sample to be discussed. Both the full breadth of quantitative responses across the project is analyzed for project wide implications, and in light of the qualitative interviews, the surveys from four high schools were analyzed separately to supplement the qualitative data findings in the four high school sample..

Site/Context

In response to the AchieveNJ law recently passed in New Jersey, the Rutgers University SSI Project implements a system in fourteen NJ charter schools overall that is aligned to public school teacher evaluation system utilizing approved teacher evaluation measures focused on data-based decision making. The Rutgers SSI Project provides training and support in the implementation of all aspects of the Teacher Evaluation System, including training, data support, technical assistance, troubleshooting, and reliability testing. The New Jersey Teacher Evaluation system as applied by the School Systems Improvement Project is designed to be in alignment with the NJ state evaluation model (School Systems Improvement Project, 2014). My individual role was described earlier within the problem statement.

The current study samples all high schools partnering with the Rutgers University SSI Project that use the *Classroom Strategies Scale* (four schools) for the qualitative portion. The total sample of teachers utilized for the survey component was derived from archived data. For the qualitative interviews and focus groups, a selection of 11 administrators taken from four representative schools of convenience was selected. Central to this choice was to ensure that all schools were utilizing the same teacher evaluation system and that data were not impacted due to individual observation measure characteristics. Choosing high schools was essential to ensure

that schools have similar characteristics within the sample. All schools are Charter Schools operating in urban areas. All schools have similar demographics, with Free and Reduced Lunch Percentiles above 50%. Schools are located in urban areas within New Jersey. Schools vary in their administrative structure and in number of teachers, grade-level, organizational structure, evaluation system components, and standardized assessments. The selection criterion was that all four schools were implementing the same teacher evaluation system. The selection of administrators was to ensure a representative sample. Focus groups of teachers were comprised of a selection of teachers from the same four schools for the same reasons as the administrators.

Sample

The quantitative sample came from all SSI Project schools and represents the totality of teachers within the project. The qualitative sample of administrators consisted of 11 administrators selected from four Rutgers SSI Project High Schools that were chosen to be representative of school size, school grade level, and administrator structure in the overall project. A sample that is more reflective of the staff will produce data more representative of the whole project (Jackson, 2007). The administrators selected were the totality of administrators from target schools involved within the teacher evaluation system. The selection criteria dictate that administrators must be directly involved in instructional evaluation; at a minimum conducting teacher evaluations, conducting observations, and designing instructional improvement or professional development programs. The 11 administrators were selected from all types of administrative roles (CEO, Principal, Vice-Principal, Supervisor, Department Head, Director, etc.) as long as they are directly involved in the evaluation of teachers. Administrators selected possessed, at minimum, a NJ Supervisory Certificate and a Master's Degree. Some participants had a NJ Standard Principal Certification and/or a Doctoral Degree in addition to the

minimum qualifications. All administrators also received training in the *Classroom Strategies Scale*, passing an observation test based on five instructional videos with a minimum reliability score and a second observation instrument determined by the partner school. The administrators received a twenty-hour training course on implementing the observation measures used by the majority of schools in the project. These included a twenty-hour training course in implementing *Classroom Strategies Scale* and a twenty-five hour training course in implementing the *Danielson Framework for Teaching*, including reliability assessments. The administrators selected were a representative sample of the proportionality of administrator roles connected to teacher evaluation present within the sample.

Table 1.

Sample Table of Administrators.

<i>School</i>	<i>Grade Levels</i>	<i>Total Number of Instructional Administrators</i>	<i>Available Instructional Administrators</i>
Charter School 1	9 to 12	3	3
Charter School 2	9 to 12	2	2
Charter School 3	9 to 12	2	2
Charter School 4	6 to 12	4	4
Overall Sample		11	11

The administrators selected for the sample included three instructional supervisors, two academic vice-principals, four principals, and two department heads. Most of the administrators are in their first administrative assignments, having between two and three years' experience in their positions. Two administrators had over three years' experience, and had prior administrative experience before joining their current organizations. Most have been "farmed

from within” their organizations and served as teachers before joining the administrative teams. About half of the sample had experience in organizations other than charter schools; including the private sector and traditional public school districts.

Table 2.

Sample Table of Teachers

<i>School</i>	<i>Grade Levels</i>	<i>Total Number of Teachers</i>	<i>Available Teachers</i>
Charter School 1	9 to 12	4	24
Charter School 2	9 to 12	4	25
Charter School 3	9 to 12	4	16
Charter School 4	6 to 12	4	52
Overall Sample		16	117

The teacher interviews of the sample were recruited from these same four institutions to provide the teacher perspective on the experiences of faculty within the same site contexts. They included four volunteer teachers from each school, representing a sample of convenience. The 16 teachers’ interviews represent the totality of teachers who volunteered to participate. Operating within the sample of convenience who are recruited due to scheduling constraints (outside of regular school hours) teachers represented a variety of content areas, years’ of experience, and degree/certification status. The school selection criteria for teachers are the same as the administrator sample. Teachers from the quantitative sample received the TEES-T survey by email and represent voluntary respondents. Surveys were distributed during the Fall 2016 semester in November. All teachers across all schools across the SSI project received the quantitative survey link. Responses were voluntary. Out of 538 teachers across the SSI project,

155 teachers responded to the survey. Responses were collected by the end of the Fall semester. Of these 155 teachers, 23 responded to the survey from the high schools targeted in this study. Given this reality, both the larger 155 sample and specific 23 sample are discussed within the results and conclusions.

Instrumentation and Data Collection

The interview protocol was developed and tested before seeing service with the research sample. It was piloted with four administrators from the Rutgers SSI Project using the same TES in a study during the 2015-2016 academic year. Piloting the protocol before use in the sample ensured that the interviews conducted indeed gathered data that will answer the research questions. Piloting the protocol before utilizing the instrument to collect data helped ensure the questions I asked were clear and effective at getting people to talk about topics that will answer the research questions (Jackson, 2007). This strategy was essential to maximizing the interview time available and producing the highest amount of usable data within the time constraints.

Interviews were conducted in a manner in alignment with best research practices that respects the often thorny schedules of school administrators. Qualitative interviews were conducted in a one-on-one setting at times of convenience for the members of the sample, usually before or after the school day. A semi-structured interview protocol was selected because it encourages participants to give in-depth responses in a deep, rich manner while hitting on key points aligned to the research questions (Patton, 2002). The protocol was used with administrators in similar demographic circumstances to ensure data collected is relevant to the research questions with smooth implementation. These data were used to refine questions before embarking on the full study. Having necessary descriptive data helps establish validity and reliability (Merriam, 2009). These responses were checked against other surveys and interviews conducted by the Rutgers SSI Project, such as the Teacher Evaluation Attitudes Survey

conducted on project teachers and administrators. General consistency in responses was apparent, at least in terms of general themes of responses. These checks helped ensure validity and reliability of the interview protocol (Jackson, 2007).

The interview protocol itself was based on the research question and was informed by the current state of the literature base. The first research question of “How are partner school administrators utilizing the results of teacher evaluation measures to support instructional improvement?” reflects that the literature reveals a potential correlation between teacher evaluation systems that are properly implemented and improved student learning and teacher outcomes such as evaluation scores. It also reflects the literature base on instructional improvement, stating that teachers self-report higher quality instructional improvement when it is tied to teacher evaluation. I explore these links within the literature review. The literature review also reveals a gap in this area, and the areas of instructional leadership and data-based decision making in this area. To better understand this area, sub questions of “What kinds of instructional improvement decisions do the data inform?”, “If partner school administrators are not utilizing teacher evaluation data in their schools for instructional improvement, what barriers are interfering?”, and “How do partner school administrators engage with and produce the data of teacher evaluation systems?” were explored to better understand how administrators make decisions using this data, create this data, or are prevented from using the data. In summary, the protocol explored how administrators produce evaluation data, how they utilize the data produced to inform instructional improvement actions, and what kinds of actions they take using these data within their building. This approach was informed by the suggestions of the current literature of potential benefits of these actions as well as the current need for additional study to confirm these potential benefits, especially in a climate nationally where state’s and intervention

agencies are dedicating time to these reforms. The protocol is present in Appendix A with potential probing questions. Questions are aligned with the conclusions of the literature, gaps in the literature, and the research questions.

To triangulate administrator data with those from teachers, especially considering many of the potential controversies and drawbacks of teacher evaluation systems as currently implemented (Hallinger et al., 2014), the question “What do teachers report about administrators' use of teacher evaluation data to support their instructional improvement?” was considered an important topic to explore in the research protocol. Having complementary questions across both interview protocols was essential to ensure transferability between protocols when considering research questions (Patton, 2002).

Interviews for administrators and teachers were audio recorded and transcribed. Data were then analyzed in light of the research questions and objectives. Interviews were recorded using a personal digital audio recorder with “Windows Media Player” as a backup recording device on a laptop computer. Interviews were transcribed by the researcher, checking for quality following each transcription.

The survey component was collected and archived by the Rutgers SSI Project. The survey is entitled the Teacher Evaluation Experience Scale. The survey was sent to teachers via email during the fall semester of the 2015-2016 school year through a Qualtrix survey. All teachers who chose to respond to the survey protocol were included within the data sample. It examines areas of system, feedback, process, and motivation to change. Teachers responded to the survey utilizing a 5-point Likert scale. The TEES-T collects beliefs and attitudes related to four constructs verified by factor analysis: (a) System (relevance of the evaluation system), (b) Feedback (value of the evaluation feedback), (c) Process (appraisal of the evaluation process),

and (d) Motivation to Change (perceived impact and importance of evaluation on classroom practice). The four constructs were based on theory and research examining the limitations of current evaluation systems as well as the previous aforementioned studies examining teachers' reactions to new evaluation system (Reddy et al., 2016). The construction of the TEES-T is rooted in a Likert scale. All items were rated on a 5-point Likert scale: 1 (strongly disagree), 2 (disagree), 3 (neutral), 4 (agree), and 5 (strongly agree). All teachers who responded did so voluntarily. The TEES-T has been tested and designed in best practices of psychometric validity and reliability following successful pilot testing and implementation over multiple school years. This includes stakeholder input, testing, review of other scales, peer-review processes, and review of other peer-reviewed measures (Reddy et al., 2016). The TEES-T was rated as internally consistent, but is more useful with four and five factor models than a one factor or global scale. It also did not reflect bias when considering for demographic factors of teachers taking the survey (Reddy et. al., 2016). Considering these psychometric properties, the TEES-T is an appropriate measure for data to be utilized in this study.

CHAPTER IV

Results

Qualitative Analysis Methodology

Transcribed interviews were uploaded into the qualitative analysis program “Dedoose.” A codebook was developed from the transcribed interviews. During the 2015-16 school year, a pilot study of 5 administrators and 4 teachers was conducted with the same semi-structured interview protocol applied in the current study. During the review of pilot study data, a tentative codebook was developed including codes and subcodes relevant to the research aims and questions. Codes were revised during multiple reviews of pilot data.

During the current study’s analysis; the pilot data codebook was the initial start point for review. After the initial review, several codes were removed due to lack of relevance and usefulness to the current study, with several sub codes added for clarity. Notes were added as necessary to clarify data. From this information, a final codebook was developed and used on sample interview data to test ability to produce useful assertions regarding the data.

A final codebook was then applied to the interview data producing a final set of transcribed, coded interviews for analysis. From these interviews and coded segments, four assertions addressing the research aims and questions emerged as most significant. These assertions, in this mixed methods study, were then considered in light of quantitative results to produce conclusions and recommendations.

Qualitative Results

Based upon the interview data collected, several assertions can be made that address the study research questions. The qualitative interviews with teachers and administrators produce a

few major assertions to answer the research questions and aims of this study. They are as follows:

- Administrators and teachers report data is used inform collaborative and specific feedback most frequently.
- Teachers and administrators report the data is most used individual instructional improvement, specifically informing PDP plans and future development goals.
- Teachers and administrators report TES data is sometimes used to inform group and whole school level improvement, but barriers interfere with its use.
- The barriers of time, knowledge, and training of both administrators and teachers are interfering with data use for instructional improvement.

Coding Structure

An initial codebook was created during the pilot study. This codebook includes the major codes of Specific Feedback, Collaborative Feedback, Directive/Negative Feedback, Evaluation Process, Observations, Pre Conference, Post Conference, Formal Professional Development, Individual Instructional Improvement, PLCs, Small Group Instructional Improvement, School-Level Instructional Improvement, Recommendations for Change, Positive Impressions, Negative Impressions, Barriers (Subcodes of: Training, Time, Knowledge, Resources, Buy-In), and Important Quotes. These codes were developed by extracting common themes from the pilot study, which utilized the same interview protocol. Teachers and Administrators each received their own version of a code (with a T or A respectively). These data was analyzed separately and as one set of coded data in different reviews to determine cross over.

Codes were assigned during an initial review of the data. Following this initial review data were organized and re-evaluated for coding integrity in alignment with best practices. Based upon the coded data, assertions emerged based upon administrator and teacher responses.

What Kinds of Instructional Improvement Decisions do the Data Inform?

***Assertion.** Administrators and teachers report data is used inform collaborative and specific feedback most frequently.*

All of the principals and administrators surveyed described instances of using collaborative and specific feedback to teachers. Collaborative feedback is feedback that is produced with the input of the teacher incorporated and discussed during a feedback meeting. In cases of this study, this shows up largely surrounding the formal observation post conference. Specific feedback describes feedback that is detailed, tied to areas where the teacher can improve, and includes examples from the lesson. Evidence was collected during the observation and presented, usually, during the observation feedback meeting (Danielson, 2011; Darling-Hammond, 2008; Donaldson & Papay, 2012; Garret et. al, 2001; Reddy et. al, 2014, Wilson & Berne, 1999). Every principal or administrator interviewed provided statements that described at least one instance of collaborative feedback and one instance of specific feedback meeting this definition and coded by the researcher.

In the teacher interviews, all teachers reported at least one instance of specific feedback in their responses. In the majority of cases (13 out of 16) teachers reported that this specific feedback was also associated with collaborative feedback. Three teachers, from different schools (one from School 3 and two from School 4) report instances of what was coded as directive feedback. Directive feedback is feedback that is described as “top-down” given without feedback or opportunity from reflection from the teacher. Two out of four teachers interviewed in School 4

reported directive feedback in one or more instances. In this school, all four administrators interviewed reported instances of collaborative and specific feedback and described this as a priority.

Administrators and teachers are in agreement throughout the interviews that in the majority of cases, specific and collaborative feedback are occurring. This speaks to the overall execution of the teacher evaluation system in accordance with best practices (Darling-Hammond, 2008). A principal at School 1 related it as "First of all, we try to ... once we start the post-conferences, let me also add this: we try to listen to the teacher first. Try to get what does that teacher feel after that lesson? Getting his insight, his or her thoughts about the lesson and generally we say, "How do you grade your lesson out of ten?", for example. We try to get more from that teacher to have a more friendly conversation rather than just "do this, do that, this is not good", rather than pretty much direct conversation." Administrators were very interested in promoting an evaluation process that was collaborative in nature. One teacher at School 1 stated "What I like about it is I'm able to have a discussion with them. It's not I just listen. I'm able to explain why I did certain things. And then they do this whole detailed evaluation, that's very specific and collaborative." The teacher referred to opportunities in which she shared points with her observer and specific recommendations and feedback provided to her from the observer." Her experience is indicative of the majority of those interviewed. This is one point in which teachers and administrators appear in close agreement. Another teacher at the same school reported "The feedback we get on the observations are tremendously detailed and in depth. We have a lot of PD tied to teacher assessment, teacher development. It's pretty thorough in its rigor, I would say." A number of teachers believed that the individual observations they receive are rigorous and tied to professional development.

All the administrators believed in the value of a post conference and believed in the value of using data to inform teacher instructional improvement. All administrators provided statements in their interviews that made clear that they believed in teacher evaluation feedback as a collaborative process and in the need to be specific in areas to improve. In particular a collaborative evaluation process is specifically valued. For example, an administrator in School 1 put forth the statement, “The post-observation, by the way, is definitely a very intensive study that's one that we're really together, collectively, collaboratively when we are dissecting that lesson and looking at it as educators on the same playing on strengths.” He made numerous statements that reinforce this same idea throughout his interview. All administrators that were interviewed spoke highly of the importance of collaboration, but to varying degrees implemented it.

All administrators interviewed used data for specific feedback during a post conference. Interviewees made statements that belied the utilization of TES data to provide specific feedback to the teachers for example, as an observer from School 1 explains, “That feedback is very specific, very targeted to the actual context of what we're dealing with in the observation.” This statement followed numerous and very specific examples of data usage to provide teachers specific feedback regarding their performance. He continues with another belief statement, “Think of it as, when you're writing a paper and you have to substantiate what you're saying. We kind of draw out a teacher that may need to work on their clarity and directives. We have some bit of evidence behind that. It's something that as I said substantiated our claim and the teacher realizes it, too. They realize the importance, too, behind each decision during that lesson.” The belief in collaborative feedback and specific feedback from the observer is clear and aligns with what is known about best practices in the evaluation process (Darling-Hammond, 2008). Again

and again, this administrator detailed his use of observation data to give feedback to teachers on instruction, and his belief that this feedback is essential to improve practice. There are numerous examples of specific feedback provided by this observer, one particularly clear example follows:

“One teacher actually would provide directions during group activities and they wouldn't lead into it with really clear directives though, so the directions given prior to the activity were pretty vague and lacking real clear direction. Then, once the group activity ensued, then directions were given over the top of the entire class. In that case, two items were addressed, one of which the calm, firm tone of voice. They actually had to grow louder and higher pitched in volume to speak over the class and also clarity directives; clear directives versus vague and the item dealing with directives being one to two steps at a time and also conducted in such a way that it's prior to the activity.”

This quote by the observer illustrated the specific nature of the feedback provided to teachers and represents the most frequent code described within all four administrator interviews. It is similar to this quote from an observer at School 4 describing her collaborative feedback process and its results:

“They even had a chance to go into their classroom, and tell me how much that changed their perspective and had an immediate impact on their classroom where they would say, "Thank you so much, because I was more aware of it and my class is going so much ... It's really going so much better now," or if I did have time to pop in I was very ... I felt it was very motivating to me, because I thought I had an increased ability to influence instruction in classrooms, because I would go back into that teacher and they were doing that.” She believes in the importance of collaborative feedback and believes that the feedback process portends gains within the classroom.”

All administrators spoke at length about how they *used data to provide specific feedback and collaborative feedback*, but did not spend anywhere nearly as much time within their interviews describing how they used the data to make other instructional improvement decisions outside of the direct feedback process. An observer at School 4 brought the evaluation process to PLCs that are intended to dissect data through discussion of the rubric, but not actual data. “During the PLCs, we looked at the Danielson rubric very, very closely. Each domain one, two, three, and four at all different points and I had asked the teacher to self-evaluate and say, “I’m a two in this area. However, I’d like to get to a three, or a four, and five.” Self-reflection and training on the rubric are important parts of teacher evaluation system implementation, but do not reflect usage of data. This is an example of the barrier of knowledge interfering in data usage. The observer’s responses are a clear reflection of belief in data usage for decision making does not necessarily translate into actualization, reflecting impact of barriers.

***Assertion.** Teachers and administrators report the data are most used to inform individual instructional improvement, specifically informing PDP plans and future development goals.*

Evidence for this assertion is seen in all interviews. However, it is only loosely connected to organizational instructional development and school-wide development planning. Two interviews show some small group data decision-making use. Outside of the formal observation process, there is less clear connection between TES data and instructional improvement. One teacher at School 2 reports, “Teacher evaluation data is pretty much just the observations and the feedback.” Several teachers interviewed throughout the four schools made similar statements, even if the administrators at their schools full heartedly believed in the use of TES data to drive instructional improvement programing. This lends itself to the question: If partner school

administrators are not utilizing teacher evaluation data in their schools for instructional improvement, what barriers are interfering?

In one case, top-down directives regarding PD over-ruled any ability to use data to inform school wide professional development. The administrators in the sample have a mixed record of using data for instructional improvement decision-making, falling beneath what the literature and administrators both state are best practice. While most (14/16) administrators interviewed utilized observation data for individual instructional improvement to much more varying degrees they had successes and failures applying these data in the improvement or decision-making process. All four administrators made mention of collaborative and specific feedback extensively in their interviews.

The most prevalent use of data to make instructional improvement decisions was associated with individual instructional improvement. All administrators made use of data for individual instructional improvement for individual professional development plans. An administrator from School 2 stated “From there (post conference) I use data set goals, and so I like to do the goal setting piece with teachers. So I have them bring the goal setting form back so they've had time to actually think about some of the goals and then based on the observation and based on their informal, then we can set goals together, which become part of a professional improvement plan.” This type of instructional improvement was most common across the interviews. For example, an administrator at School 1 details that “After the post-conference we follow up with that teacher throughout their daily instruction, through walk-throughs, where we'll look to see those categories addressed. That's also within their Professional Development Plan. We'll often integrate that in the first round of observations, into their PDP and if they do address that by the end of the year that's relatively cool.” It is clear that this administrator utilized

the individual data to inform individual development plans as well as to inform the goals of walkthroughs and the focus of future individual development and observations. At School 3, also similarly describes the usage of TES data to inform PDPs, pointing to a direct connection between TES data and PDP goals and actions. She described soliciting feedback from teachers as to what individual instructional improvement they would like to receive, linked it to the TES data, and set the PDP collaboratively with the teacher. The administrators described use of a collaborative process to set individual PD plans with an eye to collaboration and an eye to data usage.

***Assertion.** Teachers and administrators report TES data are sometimes used to inform group and whole school level improvement, but barriers interfere with its use.*

The lack of effective grouping and beliefs of the teachers that the district level professional development they receives demonstrates that there is disconnect between what administrators say is best practice in their professional development (TES data use, relevance, effective grouping) and how teachers are experiencing their professional development systems. Many teachers also report that they found their individual instructional improvement programs to be far more effective then the school level or group level professional development they were receiving from their administrators. As quoted by a teacher at School 4 in discussing grouping in professional developments, “But then, when we were separated by groups for departments we all had to develop the annual plans. No, not at all. I was not prepared or trained at all. I was just told to do it, and I was confused. I felt as though the facilitator did not know what they were doing either.” Teachers report in numerous instances throughout all four schools in similar statements that their administrators do not seem to be fully trained in the ability to group or facilitate authentic professional development.

***Assertion.** The barriers of time, knowledge, and training are interfering with TES data use for instructional improvement.*

All engaged in some form of collaboration in the teacher evaluation process, but were mixed when considering specific high quality collaborative actions. In future study, it will be important to see from the teachers' perspective, whether collaboration is happening as put forward. An administrator from School 2 puts forth a similar sentiment. "Collaboration is important, "How do you think the observation went? What do you feel you could have done differently? What areas do you think you are strong in, and what areas do you think you need a little bit more work on?" Then, I provide my feedback and explain why....It really is just a open conversation and dialogue. Pre, in particular, like I said, particularly for our new staff members, it's much more in depth" Generally, administrators believe in collaboration as a worthy goal, but had difficulty in executing follow up in collaborative feedback or instructional improvement beyond the specific, collaborative, and meaningful post conference. Teachers and administrators both reported in numerous instances across all schools that there was no or limited follow up between observations, due to the barriers discussed later in this section.

However, translating this belief into action proved problematic for some administrators even on an individual level. As reported by an administrator at School 4 discussing use of the totality of a teacher's observation record, "I haven't looked at last year's information as closely as I should have or could have." Time is reported by at least one respondent at all schools as a major barrier to fully implementing data-based decision making. Time is an extremely common barrier. It is mentioned by all administrators as a factor holding them back in their implementation and usage of teacher data, with administrators at Schools 1 and 4 specifically lamenting things they wanted to do but could not do using data because of the time constraints

placed upon them. Even when using data, there was a feeling that more could have been done, as illustrated by an administrator from School 1,” The CSSO observations are used in direct correlation with the PDP. Last year we did that. This year we actually had the timing was a little bit off, so we didn't end up using it this year, but it has been used in the past.” There is knowledge of best practice, but an inability to consistently apply data use even in the areas where it is consistently used, in the Professional Development Plans.

Another potential barrier for implementing TES data for instructional improvement emerged from lack of administrator and teacher training surrounding data (Garret et. al, 2001; Darling-Hammond, 2009), as illustrated by this quote from an administrator in School 4’s interview,” I would say two weeks ago, I had given them the time during a PLC to analyze the common assessment data and then I had posed specific questions where they were to identify student deficiencies or weaknesses and create a plan for upcoming lessons. That didn't work very well.” She posed the question and presented data, but was not able to effectively model using data for the teachers’ example. There was little structure in the use of data within the PLC, when the literature reveals the need for having clear goals, trained facilitators, and a structured, collaborative process for examining data in a PLC (Little, 2006; Thibodeau, 2008; Wood, 2007). In his interview, an administrator in School 1 directly stated that his inability to engage fully with the data was a barrier. He did not know proper terminologies and ways of presenting data, despite believing firmly in these responses that using teacher evaluation data for instructional improvement was best practice.

In several instances, teachers reported that administrators seemed to lack the training to fully execute TES data use to inform instructional improvement decisions. One teacher at School 4 reported, “don't feel that (TES Data use for instructional improvement), no. They may, in fact,

maybe it's not just me. Maybe there was data from other teachers that it's being prioritized and is in fact being put into our professional development, but I haven't seen it in the professional development or opportunities outside of the immediate post-conference". The teacher here is reporting that outside of the formal post-conferences, there are fewer opportunities for professional development or instructional improvement utilizing their TES data. Another teacher at School 3 states: "Yeah, I'd say that's somewhat accurate. I don't want to make it sound like there's a total disconnect there, but I just don't see a formal, okay, what you were assessed, here's what we're going to cover. Or we're going to cover this and then you're going to be assessed on it. By the end of a year it all gets ... You know." The disconnect between TES data and the instructional improvement received is clear within the teachers' experience. The teacher did not see a clear link between their evaluation experience and the instructional improvement received.

One way in which this issue is being evidenced is in grouping. As a teacher in school one puts it

"Well once a month we have the District professional development. And I don't like them at all because we have elementary teachers and high school teachers, and we all gotta group together and we supposed to work together, but I don't understand how. We really don't- I understand we are one big charter school chain, but elementary teaching and high school teaching is so different. So we're put in those groups, and it just doesn't make sense to me."

Two teachers in School 3 greatly lamented the lack of training and experience of their administrators in the evaluation. They do not believe that the administrators conducting the evaluation were trained enough in the system to draw valid conclusions or links to professional development. A teacher from School 3 puts forward in this quote rather bluntly: "If they're

saying to me straight up, "Gee, you know, we are new at this, too. It's new for us too." If that is the case, then these scores are not valid to me. They should not be valid to the state because of the fact that you're just admitting that you don't know what you're doing and you're saying this is new for you." Another teacher at the same school expanded on this point, "Yeah, I'd say that's somewhat accurate. I don't want to make it sound like there's a total disconnect there, but I just don't see a formal, okay, what you were assessed, here's what we're going to cover. Or we're going to cover this and then you're going to be assessed on it. By the end of a year it all gets ... You know." This concern over administrator training continues into other schools. At School 2 one teacher reports "It's hard to say if certain people that are doing these evaluations are really utilizing them in the correct way and trained to apply the results in meaningful ways." This teacher's interview demonstrated a lack of confidence not in how the administrator created the data with integrity, but in their inability to apply this data beyond the formal evaluation process.

One potential barrier that arose in one case was the unwillingness of leadership to allow for data use. One administrator felt that "... I felt very isolated because I was the only vice principal. There was no principal. There was a leader, but physically wasn't there a lot and I felt very much like I could not make the kinds of decisions with the data that I needed to make for the school." Although this only occurred in one of the four interviews, leadership buy-in and organizational reluctance to change emerged here as a potential avenue to explore in future research.

Expanding on this point, a number of teachers interviewed reported that they did not feel as though their input into their own instructional improvement on a school or group level was effectively included. For example, a teacher at School 1 reported, "No, I don't really ... I haven't seen too much teacher input into what happens with the professional development. I haven't seen

too much where teachers are asked what areas should be covered, and then those areas are covered. I haven't seen anything like that.” The literature is clear that proper training in executing professional development and instructional improvement programs for school leaders should include teacher voice (Curry, 2008; Little, 2006; Reddy et. al., 2014; Woods, 2007). Not including teacher voice potentially implies a lack of training, not only in using data directly, but also in being able to effectively and directly carry out professional development and instructional improvement programs beyond the formal observation and teacher evaluation process.

When TES data are used to inform school wide professional improvement, it is not as specific in its usage, but rather colors the theme of the workshop. In an administrator interview from School 4 she stated that, “One example was with the high-order thinking questions that the teachers weren't able to scaffold them with the academic corrective feedback, so we did a workshop on having high expectations for all learners.” In this case, the administrator identified an area of common need in the department, and used it to justify a separate program. An issue with the data use in this case relates to whether there is a disconnect and indirect correlation between the two. The data reflected weakness in the area of “*Promotes Students Thinking*”, but the solution was the provision of training on expectations. There are a number of issues that could be explored by delving deeper in the data, such as questioning skill and feedback. This reveals a potential need for expanded training and knowledge base in project administrators. The barriers of training and knowledge appear particularly impactful on this administrator. She utilized specific data, collaborative feedback, and evidenced some beginning data usage, but largely she appeared to lack the overall ability to specifically utilize or model data based decision making beyond establishing broad trends or operating directly in the post-observation feedback session.

One area where two of the four administrators utilized data to inform group level Instructional Improvement was in the pairing of teachers with mentors or into PLCs or professional development initiatives. An administrator in School 1 relates that,

“We'd try to find somebody with certain strengths, but areas for improvement that would align perfectly with the other. We'd use that data to really drive our pairings which was critical as we looked at the academic response opportunities that folly, or the rigor and then match them with a teacher... we'd tried to find somebody with certain strengths, but areas for improvement that would align perfectly with the other. We'd use that data to really drive our pairings which was critical as we looked at the academic response opportunities that folly or the rigor and then match them with a teacher ... Just another example of it.”

The use of data to inform professional development programs for the school is significantly problematic. Every administrator interviewed did not have a clear link between TES data and professional development. An administrator from School 2 finds free PD resources on government websites because no PD budget exists. Not only does this reflect the barrier of resources, but it also precluded her ability to use TES data to determine PD, regardless of whether she believed in it or not. She provided school-level feedback utilizing data from the TES at staff meetings, but did not go in depth to use data to determine PD programming at her school. Describing how PD is developed within his school, an administrator from School 1 describes instructional improvement as “very targeted toward any AP level teachers, especially we're looking at what AP training sessions are instituted throughout from a concept perspective. Then, when it comes to the regular, CORE 10th grade teacher, for instance, I'm just throwing one out there, but that's a little bit more looking at pedagogy and looking at the actual delivery of

instruction.” Essentially, data were not informing the PD selection. It was determined based upon the tracking of the students the teacher teaches. This was in contrast to the small group, PLC assignment, and individual data use seen in his decision-making.

Teachers reported widely different experiences with their TES data and links to professional development within and across the four schools at the heart of the qualitative interviews. For example, one teacher at School 3 reflected positively on her experiences with the teacher evaluation: “ Face-to-face resources are a lot, a lot of it is face-to-face dialogue. Some of it is written resources. In our professional developments we're given printed material. We've seen videos. I think most of it is communication face-to-face. I'd say that's the biggest resource that we have is our administrators and our mentors.” This quote portends a belief in the effectiveness of teacher evaluation at the teacher level. However, from this same school, another teacher reports “ The barrier of training is particularly problematic for the Rutgers SSI project. If teachers do not, at a fundamental level, believe that the administrators do not believe in the validity of the teacher evaluation systems conclusions, it will also not be valid in the eyes of the teacher. This is particularly concerning considering the administrators at School 3 both believe that they are utilizing the system to maximum effectiveness.

These quotes are supported by others throughout the school. One teacher at School 4, described the link between TES data and professional development, “I hadn't really noticed anything yet, truthfully.” In at least one teacher interview per school, this statement is repeated at least one time. At School 4 this attitude is most prevalent, being seen in all of the four interviewed teachers. At School 1 this attitude was the least prevalent, being reported one time. At Schools 2 and 3, this statement was reported two times each. The results here show a mixed

link between instructional improvement programs on the group and school level as experienced by teachers and TES data within these schools.

This trend continues throughout the schools examined. In School 2, a teacher reported in terms of the link between TES data and instructional improvement

“Personally, no. I think it's just a bunch of guys coming together and saying, Okay, let's talk about this, let's talk about that. Not looking and saying, like a lot of thing is like we keep seeing within our evaluation like, "Differentiating or Adaptive Instruction, nothing.”

The clear link between data and instructional improvement seen within specific feedback, collaborative feedback, and one-on-one interactions (like creating a PDP plan) appeared less clear when the conversations turned to discussing instructional improvement within a formal professional development setting. In School 3, all teachers reported a tight link between student data and instructional improvement programs, if not directly connected to TES data. One teacher at School 3 reported when discussing instructional improvement data “It's basically been focused on the students’ data, not on our evaluation data.” While the focus may be on data, TES data appeared to be applied unevenly.

Several teachers went so far as to specifically recommend that administrators receive more explicit training in how to link their evaluation data to instructional improvement when asked for recommendations at the end of their interviews. For example, a teacher at School 4 reports “If they (administrators and the TES data) were more synchronized I think more teachers would benefit from that actually.” This recommendation was a common one, appearing another six times throughout the interviews. Another teacher at the same school similarly reported a desire for more professional development that is relevant and linked to their teacher evaluation data stating “I definitely liked the evaluation system process and how it worked, and I liked how

detailed it was for specific things. I just didn't like that we only used it three times and then we were done. I feel like we could have done more with it in terms of our development.” This sentiment appears again and again throughout the qualitative data, and is a good representation of the types of recommendations and responses teachers gave throughout the schools. Teachers liked their evaluation feedback, but generally desired in most cases that more was done with it to improve their instruction.

The interviews reflect a trend in setting school development programs where some data use occurs, but does not occur in a specific or in-depth manner beyond the immediate post conference in the amounts that are reflected as necessary within the literature practice (Curry, 2008; Little, 2006; Wood, 2007). The gap exists in specific group data usage. Data-use is being used to set broad trends or directions, but is not informing in the rich, specific manner that is established by the literature as best practice (Curry, 2008; Little, 2006; Wood, 2007).

Quantitative Results

In the quantitative survey on Teacher Evaluation Systems (TEES-T) 534 teachers involved in schools partnered with the Rutgers University School System Improvement Project were invited through email to complete a survey on various items about their teacher evaluation feedback, process, and system. Teachers were asked whether they strongly disagreed, disagreed, were neutral, agreed, or strongly agreed with a series of statements (See Appendix C). A total of 155 teachers responded to the survey.

Overall, teachers were largely satisfied with the process, system, and feedback from teacher evaluation. Approximately one third of teachers were neutral, with about 10-20% not believing the system was relevant or useful. This pattern was seen in item after item. However, these numbers of teachers who do not see a connection between TES and the small group,

individual or whole school level instructional development they receive and the relevance of the TES to their individual learning is slightly higher than other responses.

The individuals responding to the items generally felt positively about the statements, agreeing or strongly agreeing with them at a margin of roughly one hundred to fifty who feel neutral with a handful of teachers disagreeing. Of particular relevance to the research question is the response to the prompt seen in the table regarding whether “The teacher evaluation system provided professional development opportunities that motivated me to change my classroom practice.” A large number of teachers (25) disagreed with the sentiment.

For example, six out of one hundred fifty-five who responded did not believe that evaluation feedback was useful as seen in Table 8 Appendix D. Thirty-six were neutral. The remaining teachers either agreed or strongly agreed with the assertion that evaluation feedback was useful. This pattern is seen in the rest of the tables in Appendix D in similar proportions. Impacting the results, out of the sample, 379 did not respond to the email prompt. The responses demonstrate that of the teachers who responded to the survey, roughly two thirds agree that the process, feedback, and system meet standards of being specific, collaborative, useful, and well executed. However, a smaller number believe, especially as seen in Table 39 in Appendix D, that the teacher evaluation system provides professional development opportunities that motivated them to change practice.

Table 3. Descriptive Statistics for Charter High School 1-4 (n=23) Quantitative Sample TEES-T.

<i>Scale</i>	<i>N</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Std. Deviation</i>
Evaluation System	23	20	47	34.57	6.727
Evaluation Feedback	23	30	69	54.09	9..278
Evaluation Process	23	9	40	28.61	6.583
Evaluation Motivation	23	12	25	20.48	3.489

Table 4. Average response of sample for each subscale.

<i>Scale</i>	<i>N</i>	<i>Average</i>
Evaluation System	23	3.7
Evaluation Feedback	23	4
Evaluation Process	23	4
Evaluation Motivation	23	3

To ensure clear alignment with the qualitative sample from four SSI project high schools, it is important to examine a subset of the larger sample from just these schools. Analyzing the 23 responses from teachers among the four-project high schools for comparison to the larger sample, a similar pattern emerges. In the subscales connected to the areas most in line with the teacher evaluation system in terms of feedback, system, and process, teachers generally agree with and like their evaluation system. However, when discussing items associated with motivation for change, including items centered around links to professional development, motivation to change, relevance to teacher evaluation to instructional improvement, etc, the link is not as clear (Appendix C). This is in line with the findings when examining the full sample of

schools with 155 teachers from all grade levels. Teachers agree that the process, system, and feedback of the TES is high quality; but drop from agree to neutral on average when considering whether the TES enacts motivation change and classroom change.

Table 5.

The evaluation feedback provided information for professional development opportunities.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	2	1.2
Disagree	16	10.3
Neutral	45	29
Agree	77	49.7
Strongly Agree	15	9.7
Total	155	100

Table 6.

The evaluation system helped to improve the quality of instruction.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	2	1.2
Disagree	11	6.6
Neutral	44	26.5
Agree	94	56.6
Strongly Agree	15	9.0
Total	155	100

*Table 7.**The evaluation system improved my professional growth.*

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	1	.6
Disagree	10	6.5
Neutral	43	27.7
Agree	86	50.6
Strongly Agree	15	8.8
Total	155	100

*Table 8.**The evaluation system informed changes in my classroom practice.*

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	12	7.7
Neutral	43	27.7
Agree	88	56.8
Strongly Agree	12	7.7
Total	155	100

Table 9.

The evaluation system helped to improve student learning.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	2	1.2
Disagree	15	9.0
Neutral	61	36.7
Agree	72	43.4
Strongly Agree	16	9.6
Total	155	100

As seen in Tables 5, 6, 7, 8, and 9, (The evaluation system helped to improve student learning; The evaluation system improved my professional growth; The evaluation system informed changes in my classroom practice; and response to the item addressing whether the evaluation feedback provided information for professional development opportunities) an increase in the amount of teachers who disagree with the prompt is seen within the tables. While a small increase in the amount of disagreeing or neutral responses, this point aligns with the third and fourth assertions made from the qualitative data. Teachers see less of a link between their TES and their professional learning and instructional improvement. Particularly noteworthy, as seen in Table 7 (the largest number disagreeing or neutral), teachers are weaker in their agreement on the connection of the TES system to student learning. Given the stated goal of TES systems is to lead to ultimate increases in student practice and teacher learning, teachers responding more negatively in this connects well to the qualitative responses questioning the relevance and application of the instructional improvement programs received by teachers.

The four assertions from the qualitative data describe TES data that is more clearly being used for one-on-one professional development and instructional improvement as well as specific and collaborative feedback within the teacher evaluation process; while recognizing that a less clear link exists between the TES data and instructional improvement in small group, whole school, and formal professional development programs. This aligns with the quantitative results that state among different factor groups and survey items teacher generally agree that their teacher evaluation system is executed well, includes specific and collaborative feedback, and aligns with institutional curriculum and goals. However, the connection becomes less explicit when considering survey items connected with instructional improvement, professional development, and relevance to student learning. These assertions similarly appear when examining a sample of 23 high school teachers in line with the qualitative sample. This creates an overall picture of a TES system that shows some connection to instructional improvement, but demonstrates a far weaker connection when expanding beyond the one-on-one and formal evaluation process.

CHAPTER V

Conclusion

Implications of Assertions

There are several significant implications of the first assertion, “Administrators and teachers report data is used inform collaborative and specific feedback most frequently.” Qualitatively, this reflects that, throughout the targeted schools, teachers are primarily experiencing a TES system that in the formal process is being executed with integrity. Several studies suggest the conclusion that with a TES system that is conducted with integrity (including direct feedback, observation number, trained observers, collaborative post-conference, etc.), there can be potential correlation to gains in student achievement (Ho & Kane, 2013; Tyler & Taylor, 2012; Wilson & Berne, 1997). The qualitative responses supporting this assertion demonstrate that within the partner schools examined, this appears to be happening. This pattern is also seen in the quantitative results, in that teachers who responded to the survey responded positively and agree roughly by a two to one margin compared with neutral (and very disagree/strongly disagree). In the case of the first assertion, the qualitative and quantitative results align to demonstrate that within the schools in question, collaborative and specific feedback is executed in a manner in alignment with best practices.

The main conclusion to be drawn from the first assertion and its supporting evidence is that, in line with all the administrators interviewed, all the teachers in their qualitative responses and the majority of teachers in their survey responses are experiencing specific and collaborative feedback within their TES system. This is in line with best practices described within the literature review. In this respect, the data suggests administrators are utilizing TES data to make decisions within one-on-one contexts, and are utilizing the data to inform instructional

improvement decisions within one-on-one contexts free of barriers. This also supports the conclusion that administrators are largely applying the TES system as intended and in line with the training they received from the Rutgers SSI project. However, as seen in Appendix D Table 28, administrators are not engaging in the practice of a pre-conference, despite training and establishment in the literature as best practice.

Responding to the second assertion ““Teachers and administrators report the data is most used for individual instructional improvement, specifically informing PDP plans and future development goals,” has similar implications as the first. Administrators largely feel confident in their ability to execute the individual teacher evaluation process. They feel better changed and feel more confident in their ability to execute. They describe more examples of utilizing TES data within one-on-one interactions, whether that is in formal post conferences, informal follow up, PDP plans, or recommendations for individual growth. This is similarly born out within the quantitative data, with responses maintain the 2 to 1 (agree to neutral) ratio on items related to the one-on-one interactions and formal evaluation follow up process. This level of confidence fades moving into the next assertion.

The third assertion “Teachers and administrators report TES data is sometimes used to inform group and whole school level improvement, but barriers interfere with its use,” is in line with the rationale for the study and the expectations produced from the literature review. Teachers generally report in surveys that the formal instructional improvement they receive, particularly in group professional development settings, is not connected to TES data, student data, or seen by teachers as relevant to their practice (Curry, 2008; Darling-Hammond, 2008; Little, 2006; Thibodeau, 2008; Woods, 2007). As described within the quantitative section of the results, the number of teachers who agree or strongly agree that TES data is used to inform

professional development is lower than the other items examined. This is in line with the majority of teachers interviewed. Over half of the teachers expressed concerns over their small group or school level professional development and many explicitly made clear that they saw no link between their evaluation data and the instructional improvement they received. Even though there were some cases where such a link was established, in every school there was at least one teacher who expressed this level of reservation.

Administrators also expressed less confidence in their institutions' connection between TES data and instructional improvement programming. In Schools 2 and 4 all of the administrators expressed that they were uncomfortable personally in their ability to connect TES data to instructional improvement. In School 3, all administrators felt confident in their link (even though half of the teachers interviewed did not see this issue the same way), and in School 1, all administrators felt confident in what they could control instructional improvement, but felt as though some outside factors impeded in their ability to create a successful link as detailed within the qualitative results. In the areas of whole- school and group level instructional improvement, there appears to be a gap between best practices and execution within SSI project partner schools.

The fourth assertion addresses the question of "why?" The barriers of time, knowledge, and training of both administrators and teachers are interfering with data use for instructional improvement." This assertion is primarily derived from the qualitative results, in the reasons given for the lack of TES data and instructional improvement when such a link is described. Seen from both the teachers and administrators perspective, training of administrators emerges as the most significant and frequent reason provided by the majority teachers and administrators. This is followed by training of teachers in the ability to interpret TES data, also seen in both

teacher and administrator responses. Time is connected to the gap in resources, while funding was mentioned in one instance, time was mentioned in several; leading them to be combined within the assertion.

Knowledge of best practice is another issue that arises over and over again again throughout the interviews of administrators. Administrators often believe that they are executing best practice, but in doing so are taking actions (such as ineffectively facilitating PLCs) that are clouding the link between TES data and instructional improvement. If one believes they are establishing a clear link, but the link is not materializing, from the perspective of the teacher in practice no such link will exist. As such, the barrier of knowledge goes beyond specific training in the execution of PD or in Data-Based Decision-Making, to a long term resourcing and consultation need. Addressing these barriers form a significant portion of the recommendations to address the potentially unclear link between TES data and instructional improvement.

Recommendations to Address Assertions within Problem of Practice

Addressing the assertion of “Administrators and teachers report data is used inform collaborative and specific feedback most frequently,” the results potentially speak to the power of the initial training and onboarding process SSI project administrators undergo. The initial training and orientation process is essential for instilling core organizational skills within instructional leaders (Fullan, 2007; Odden, 2011; Reeves, 2009). In the current Rutgers SSI project training model, the majority of the initial orientation and onboarding occurs in the form of training on the CSS teacher evaluation method and Danielson Framework for Teaching. During this training, there is an extensive focus on modeling feedback, creating feedback, and applying the teacher evaluation system with integrity. Teachers report that administrators are

providing specific and collaborative feedback in line with the training and support provided by the larger SSI project.

The focus on specific and collaborative feedback within the training process has led to demonstrable satisfaction among teachers. In the summer institute for administrators (a one week long yearly training session for principals), there is one module currently on data-based decision making. In this material, the focus is within the one on one teacher focus and immediate observation/formal feedback system. Again, this focus is reflected in teacher responses that generally express satisfaction with the formal evaluation process and one on one process.

Another assertion is that “Teachers and administrators report the data is most used individual instructional improvement, specifically informing PDP plans and future development goals.” This assertion is in line with the focus of the training and support currently provided by the Rutgers SSI Project. Repeating this success with explicit training and support in linking TES data to instructional improvement is essential (Odden, 2011). This is an example of something being done well that can be utilized to inform organizational change within a broader context (Fullan, 2007).

Another recommendation comes from a “top-down” perspective. Administrators cited in several instances that they did not feel supported by those above them, whether they be Board of Trustee members, CEOs, or school leaders. The SSI project is in a unique position through the relationships developed and the ability to provide resources that it can potentially influence improvements in “buy-in” in these administrators over time (Fullan 2007, Reeves, 2009). Being able to both informally and formally communicate the importance of dedicating organizational weight, time, and training to linking TES data to instructional improvement is a worthy goal, and

should be executed by all members off the SSI project within their interactions with school leadership members in all sectors.

Most concerning to the Rutgers SSI project are the assertions of “Teachers and administrators report TES data is sometimes used to inform group and whole school level improvement, but barriers interfere with its use,” and “The barriers of time, knowledge, and training of both administrators and teachers are interfering with data use for instructional improvement.” These assertions point to gaps within the current training, resourcing, and abilities of all levels of administrators within partner schools to address. The SSI project, beyond providing support and training associated with human capital management strategies (taking the form of webinars, resources, workshops, and consultation) cannot as an outside partner fully address (beyond providing training on best practices) the barriers of time and resources. However, the SSI project has the ability and resources in this or future endeavors to provide more effective training on best practices in instructional improvement, that invariably lead to utilizing TES data to inform the types of programming provided (Darling-Hammond, 2008; Little, 2006). Designing and implementing tangible, research-based professional development and adult learning opportunities for administrators within the areas of executing professional development themselves for others is a key area for growth within the SSI project, and fits in line with large established facts in the literature describing best practices adult learning and instructional leadership (Brookfield, 1989; Franke et al., 2005; Little, 2006; Thibodeau, 2008). Explicit focus on training, onboarding, and preparation within best practices instructional improvement and linking TES to instructional improvement would be positive recommendations for the SSI project to address many of the assertions coming from this mixed methods study addressing a problem of practice.

There are three concrete recommendations for the Rutgers SSI project to address the implications of the assertions of this mixed methods study in light of the results and conclusions:

- 1) SSI project staff should continue to develop formal and informal relationships with school leaders, stakeholders, administrators, and teachers with the intention of promoting awareness of the lack of connection between TES data use and instructional improvement programs and highlighting the potential benefits of establishing a clear link.
- 2) The SSI Project should embark on formal training programs specifically addressing best practices in instructional improvement and in the explicit link between TES data use and instructional improvement that focuses on one-on-one, small group, and whole group interactions. These activities should be in line with successful current efforts in specific feedback, collaborative feedback, and the successful implementation of the TES system, especially within the on-boarding process.
- 3) The SSI project should model best practices within TES data use and instructional improvement. This includes through modeling, workshops, and peer observations within institutions. Given that there are actors within the SSI project (as seen in the qualitative and quantitative results) that re linking TES data use to instructional improvement in a best practice manner, distributing this knowledge throughout the SSI project is essential to full organizational change (Fullan, 2007).

Recommendations to Address Assertions Beyond the Problem of Practice

Current teacher evaluation reform efforts, including within the state of NJ, have within their stated goals that improved teacher evaluation systems will lead to improvements in teacher quality (National Council on Teacher Quality, 2012; Rotherham & Mitchel, 2014; Shulman, 2014). The results of this study portend that within this problem of practice, an uneven level of

success is found in linking TES data with the instructional improvement programs that in theory would lead to improvements in teacher quality. Given this disconnect, expanding focus on administrator training and eliminating barriers broadly are important to establishing potential gains from teacher evaluation reform.

The first assertion potentially establishes that an intervention program like the Rutgers SSI project can lead to specific feedback and collaborative feedback becoming norms of practice, with teachers largely satisfied with their TES system. These lessons are significant and have the potential to be considered beyond this one case of intervention. The recommendations produced for this problem of practice, with future confirmatory studies in expanded contexts, could potentially be applicable in other settings beyond the partner schools of the Rutgers SSI project.

Limitations of the Study

There are several limitations of this study that should be considered when discussing results, conclusions, and educational implications. All of the schools in question are charter schools. This impacts generalizability to public school districts, which typically have tighter requirements on how teacher evaluation data can be shared and applied.

All of the schools in question are a part of the same university/school partnership grant. When considering generalization, the schools in question have similar factors in training, teacher evaluation system implementation, and supports in best practices in professional development are similar across schools as they employ the same model. All schools are high poverty, high needs; and this fact should be considered when considering generalizability to other school districts. This characteristic should be considered when deciding the generalizability of the study beyond the immediate problem of practice addressed in this study.

Another limitation to consider, all of the qualitative work is specifically focused on high schools. The quantitative survey is of all project schools of all grade levels. This limits the ability to effectively link the quantitative and qualitative components of this mixed methods study. The professional development and instructional improvement realities of primary and middle schools may be different. This is an area for future study within the Rutgers University School Systems Improvement project.

The sampling constitutes major limitations. Quantitatively, few High School teachers responded to the survey, requiring a larger analysis base to draw conclusions and assertions. The sample of high school data were subsequently analyzed to ensure data applicability and connection to the qualitative sampling. The qualitative sample also reflects self-selection among teachers. Due to a limited response, the teachers who were interviewed were the only ones to respond to the interview request.

Another limitation to consider impacts the initial qualitative sample. A large part of the sample was chosen as a sample of convenience. Teachers in particular were only reachable during certain time periods and dates. It was not possible to conduct any randomization of the qualitative sample between grade level and subject areas of teacher. With administration, all targeted administrators were interviewed; not bringing this point into consideration. In future studies, randomizing the teacher qualitative sample would remove this limitation.

One final limitation to consider is that the data collected for this study is self-reported. Observations were not conducted as a part of this study. The data are self-reported via surveys and interviews and reflects the issues associated with self-reporting. Findings of this study should be considered in light of this limitation.

Recommendations for Future Research

Based upon the results of this study and the conclusions of the literature review, there are several key areas for future study that should be addressed. The studies which exist focus primarily on justifying the existence of teacher evaluation systems based on potential correlations to student outcomes (Ho & Kane, 2013; Taylor & Tyler, 2012; Wright et al., 1997). However, the base is not deep enough to say without any qualifications that a fully-implemented teacher evaluation system will lead to increased student learning. More study is needed to clarify the link between implementing a teacher evaluation system with integrity and seeing those gains eventually filter down into gains in student achievement. Establishing a clear link (if one exists) between implementing a teacher evaluation system in a best practices matter, and clear demonstrative gains in student achievement in large scale studies will be necessary for such work to continue within the context of broader teacher evaluation reform.

Another important area for future research is establishing whether teacher evaluation systems that demonstrate clear links to instructional improvement programs lead to clear gains in student achievement and improved teacher classroom practice. The literature that currently exists makes it clear that utilizing data within instructional improvement and professional development leads to instructional improvement that is beneficial to the teacher and demonstrates potential correlations to improved classroom practice (Darling-Hammond, 2008; Curry, 2008; Little, 2006; Woods, 2007). The current literature base does not include confirmatory studies that establish beyond question a direct cause and effect between utilizing TES data and gains within student achievement.

This mixed methods study demonstrates that within the schools examined, there is an uneven application of TES data to inform instructional improvement among teachers. Repeating

this study among different contexts, schools, states, and organizations is important to determine whether this state of affairs (and associated barriers) is pervasive across current American education. Determining whether the results of this study are generalizable and whether linking TES data to instructional improvement would demonstrate clear improvement in student outcomes is a topic worthy of future study.

References

- Bill and Melinda Gates Foundation (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET project's three-year study policy and practice brief*. Seattle: MET Project: .Bill & Melinda Gates Foundation.
- Brookfield, S. (1989).Overcoming resistance to learning. In *The skillful teacher: On technique, trust, and responsiveness in the classroom* (pp. 147- 162). San Francisco: Jossey-Bass Publishers.
- Brookfield, S. (1989). Evaluating learning and its facilitation In *The skillful teacher: On technique, trust, and responsiveness in the classroom* (pp. 261- 282). San Francisco: Jossey-Bass Publishers.
- Boudett, K. P., City, E., & Murnane, R. (2005). *Data wise: A step-by-step guide to using assessment results to improve teaching and learning*. Harvard Education Press: Cambridge, MA.
- Burke, B. M. (2013). Experiential Professional Development: A Model for Meaningful and Long-Lasting Change in Classrooms .*Journal Of Experiential Education*, 36(3), 247-263.
- Casabianca, J. M., McCaffrey, D. F., Gitomer, D. H., Bell, C. A., Hamre, B. K., & Pianta, R. C. (2013). Effect of observation mode on measures of secondary mathematics teaching. *Educational and Psychological Measurement*, 73(5), 757-783.
- Coe, R., Aloisi, C., Higgins, S., & Major, L. E. (2015). *Developing teachers: Improving professional development for teachers*. Seattle, WA: Sutton Trust.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.
- Creswell, J. W., & Plano-Clark, V. L (2007).Analyzing data in mixed methods research. In J. W.

- Creswell, & V. L. Plano-Clark. *Designing and conducting mixed methods research* (128-135). Thousand Oaks, CA: Sage.
- Cunningham, A. E., Zibulsky, J., Stanovich, K. E., & Stanovich, P. J. (2009). How teachers would spend their time teaching language arts: the mismatch between self-reported and best practices. *Journal of Learning Disabilities*, 42(5), 418-430.
- Curry, M. (2008). Critical friends groups: The possibilities and limitations embedded in teacher professional communities aimed at instructional improvement and school reform. *Teachers College Record*, 110(4), 733-774.
- Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., Orphanos, S., & National Staff, D. C. (2009). *Professional learning in the learning profession: A status report on teacher development in the united states and abroad*. National Staff Development Council.
- Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35-39.
- Danielson, C. (2011) *Danielson framework for teaching*. Teachscape Framework Proficiency Teaching System.
- Demonte, J. (2013) *High quality professional development for teachers*. Center for American Progress. Retrieved from <<http://www.americanprogress.org/wp-content/uploads/2013/07/DeMonteLearning4Teachers-1.pdf>> on November 2, 2014.
- Dedoose Version 7.0.23. (2016) web application for managing, analyzing, and presenting qualitative and mixed method research data. Los Angeles, CA: Socio-Cultural Research Consultants, LLC.
- Desimone, L. M., Smith, T. M., & Frisvold, D. E. (2010). Survey measures of classroom instruction: comparing student and teacher reports. *Educational Policy*, 24(2), 267-329.
- Doehring, P., & Volkmar, F. (2016). Knowledge gaps in ASD research: Short and long term

- implications for policy. *Journal Of Autism & Developmental Disorders*, 46(3), 733-736.
- Donaldson, M. L., & Papay, J. P. (2012). *Reforming teacher evaluation: one district's story*. Washington, D.C.: Center For American Progress.
- Doran, Amy S. (2015). TVAAS rankings and teachers' perceptions of data-driven professional learning in northeast Tennessee title I and non-title I elementary schools. *Electronic Theses and Dissertations*. Paper 2499. Retrieved from <<http://dc.etsu.edu/etd/2499>>
- Farrell, C. (2015). Designing school systems to encourage data use and instructional improvement: A comparison of school districts and charter management organizations. *Educational Administration Quarterly*, 51(3), 438-471.
- Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher*, 43(2), 100-107.
- Finnegan, R. S. (2013). Linking teacher self-efficacy to teacher evaluations. *Journal of Cross-Disciplinary Perspectives in Education*, 6, 18–25.
- Finnegan, R.S. (2016). "Teacher And principal perceptions Of a new evaluation program for teachers" dissertations, theses, and asters projects. *College of William and Mary Paper* 1463428391.
- Firestone, W. A., & Riehl, C. (2005). Introduction. In W. A. Firestone & C. Riehl (Ed.), *A new agenda for research in educational leadership* (pp. 1-11). New York, NY: Teacher College Press.
- Franke, M., Kazemi, E., Shih, J., Biagetti, S., & , D. (2005). Changing teachers' professional work in mathematics: One school's journey. In T. A. Romberg, T. P. Carpenter, & F. Dremock (Eds.), *Understanding mathematics and science matters* (pp.

- 209 –229). Mahwah, NJ: Erlbaum.
- Fullan, Michael. (2007). *The new meaning of educational change*. 4th Ed. Teachers College Press: New York, NY.
- Gable, R., Hester, P., Rock, M., & Hughes, K. (2009) .Back to basics: Rules, praise, ignoring, and reprimands revisited. *Intervention in School & Clinic*, 44(4), 195.
- Gallagher, H. A. (2004). Vaughn Elementary innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education*, 79, (4), 79-100.
- Garret, M. S., Porter, A., Desimone, L., Birman, B., & Kwang Suk Yoon (2001) What makes PD effective? *American Education Research Journal* 38(4), 915-945.
- Gist, Deborah. (2013). *Rhode Island teaching and learning: Year one report*. Rhode Island Department of Education. Retrieved from
<http://www.ride.ri.gov/Portals/0/Uploads/Documents/Teachers-and-Administrators-Excellent-Educators/Educator-Evaluation/Education-Eval-Main-Page/2013_Evaluation_Data_External_Report.pdf>
- Gold, Stephanie. (2015). Driven by data: How three districts are successfully using data, rather than gut feelings, to align staff development with school needs. *Technology & Learning*, 25(11), 6-10.
- Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability*, 26(1), 5-28.

- Hamilton, L. S., Stecher, B. M., Russell, J. L., Marsh, J. A., & Miles, J. (2008). Accountability and teaching practices: school-level actions and teacher responses. *16* (1). *Research in the Sociology of Education*, 31–66.
- Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal*, *51*(1), 73-112.
- Himley, M., & Carini, P. F. (2000). *From another angle: Children's strengths and school standards: The Prospect Center's descriptive review of the child*. New York: Teachers College Press.
- Ho, A. D., & Kane, T. J. (2013). *The reliability of classroom observations by school personnel*. (Research Paper) MET project. Tallahassee, Florida: Bill and Melinda Gates Foundation.
- Jackson, S.L. (2007). *Research methods: A modular approach*. Belmont, CA: Wadsworth Publishing.
- Kane, T. Taylor, E. Tyler, J., & Wooten, T. (2010). Identifying effective classroom practices by using student achievement data. Cambridge, MA: *NBER Paper Series*, National Bureau of Economic Research. Retrieved from <<http://www.danielsongroup.org/ckeditor/ckfinder/userfiles/files/IdentifyingEffectiveClassroomPractices.pdf>> on October 7th, 2014.
- Kimball, S. M., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, *45*(1), 34-70.
- Knowles, M.S., Holton, E.F., & Swanson, R.A. (1998) *The Adult learner*. Witman, MA: Butterworth Heinemann.

- Leung, D. P., & Kember, D. D. (2003). The relationship between approaches to learning and reflection upon practice. *Educational Psychology, 23*(1), 61-71.
- Little, J. W. (2006). *Professional community and professional development in the learning-centered school*. Washington, DC: National Education Association.
- McFarland, L., Saunders, R., & Allen, S. (2009). Learning and teaching positive guidance skills: Lessons from early childhood practicum students. *Journal Of Early Childhood Teacher Education, 29*(3), 204-221.
- Mahar, J., & Strobert, B. (2010). The use of 360-degree feedback compared to traditional evaluative feedback for the professional growth of teachers in K-12 education. *Planning and Changing, 41*(3-4), 147-160.
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey Bass.
- Morrison, J. (2009). Why teachers must be data experts. *Educational Leadership, 66*(4) 31-37.
- National Council on Teacher Quality (2012). *State of the states 2012: Teacher effectiveness Policies*. Washington D.C.: National Council On Teacher Quality.
- Neumerski, C. M. (2013). Rethinking instructional leadership, a review: What do we know about principal, teacher, and coach instructional leadership, and where should we go from here? *Educational Administration Quarterly, 49*(2), 310-347.
- New Jersey Department of Education.(2014) *AchieveNJ overview*. Retrieved October 2, 2014 from <<http://www.nj.gov/education/AchieveNJ/intro/guide.pdf>>
- Odden, A. R. (2011). *Strategic Management of Human Capital in Education: Improving Instructional Practice and Student Learning in Schools*. Routledge, Taylor & Francis Group

- Patton, M.Q. (2002). *Qualitative research and evaluation methods*. (Third Edition). Thousands Oaks, CA: Sage Publications.
- Putnam, R. & Borko, H. (2001). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher* 29(1), 4-15
- Rademaker, L. L. (2008). Using qualitative observations to enhance the evaluation of a K-12 faculty professional development program: What the numbers can't tell you. *Journal of Ethnographic & Qualitative Research*, 2(3), 205-211.
- Reddy, L. A., Fabiano, G. A., Dudek, C. M., & Hsu, L. (2013). Instructional and behavior management practices implemented by elementary general education teachers. *Journal of School Psychology*, 51(6), 683-700.
- Reddy, L. A., Fabiano, G., Dudek, C. M., & Hsu, L. (2013). Development and construct validity of the classroom strategies scale-observer form. *School Psychology Quarterly*, 28(4), 317-341.
- Reddy, L. , Dudek, C. ., Kettler, R. ., Peters, S. ., & Kurz, A. . (2016). Measuring educators' attitudes and beliefs about evaluation: Construct validity and reliability of the teacher evaluation experience Scale. *Educational Assessment*, 21(2), 120-134.
- Reese, W. J. (2005). America's public schools: From the common school to "No Child Left Behind". Baltimore: Johns Hopkins University Press.
- Reeves, D. B. (2009) *Leading change in your school: How to conquer myths, build commitment, and get results*. Association of Supervision and Curriculum Development: Alexandria, VA.
- Rotherham, A. J., & Mitchel, A. L. (2014). Genuine progress, greater challenges: A decade of teacher effectiveness reforms. Boston, MA: Bellwether Education Partners.
- Sallis, E. (1996). *Total Quality Management in education*. Second Edition.

School Systems Improvement Project (2014). *Internal memorandum regarding TES weights*

.Rutgers University School Systems Improvement Project. November 9th, 2014.

Schulman, P. (2014) *AchieveNJ 2014 update*. Trenton: New Jersey Department of Education.

Silberman, M. &Auerbach, C. (2006). *Active training: A handbook of techniques, designs, case examples, and tips* (3rd Edition). San Francisco: Pfeiffer

Sparks, S. (2013). Data based decision making: Data driven improvement and accountability. *Education week*, 33(10), 9-17.

Taylor, E. S., & Tyler, J. H. (2012). Can teacher evaluation improve teaching? *Education Next*, 12(4), 78-84.

Thibodeau, G. M. (2008). A content literacy collaborative study group: High school teachers take charge of their professional learning: Collaborative interactions are an important and effective professional learning option for secondary teachers attempting to integrate literacy strategies and content instruction. *Journal of Adolescent & Adult Literacy*, 52(1), 54-64.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). New York: Harper & Row.

Wagner, T., Sconyers, N., & Center on Families, C. L. (1996). "Seeing" the school reform elephant: Connecting policy makers, parents, practioners, and students.

Walker, H. M., Colvin, G., & Ramsey, E. (1995). *Antisocial behavior in school: Strategies and best practices*. Pacific Grove, CA: Brooks/Cole.

Wilson, S., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education*, 24, 173- 209.

Wood, D. R. (2007). Professional learning communities: Teachers, knowledge, and knowing.

Theory into Practice, 46(4), 281-290.

Wright, S. P., Horn, S., & Sanders, W. (1997). Teacher and classroom context effects on student

achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in*

Education, 11(1), 57-67.

Appendix A

Christian Mathews, Interview Protocol for Administrators and Potential Probing Questions

Hello (insert name of participant):

Thank you for taking the time to talk with me. My name is _____ and I will be conducting this interview. This interview will be recorded for research purposes.

The purpose of this interview is to investigate practices of instructional improvement and professional development within schools partnering with the Rutgers University School Systems Improvement Project. The interview will take approximately 30 minutes and will begin with a little background about yourself then move into discussing teacher evaluation, instructional improvement, and Professional Development.

EVALUATION AND INSTRUCTIONAL IMPROVEMENT

Describe what you do as an evaluator? Take me through the process from start to finish

- a. What resources do you use?
- b. Probe for what happens in pre and post conferences and what form of feedback teachers get?
- c. Could you give me a specific example of some of the feedback that you have given this year?
- d. Could you describe the type of feedback you give to teachers?
- e. What happens to the feedback after a post conference (if applicable)?
- f. How do you use data within your school?
- g. How is evaluation data used within your school?
- h. Did the teacher contribute during pre/post conference?

What kind of follow up occurs after a feedback session?

- a) What happens to CSS/Danielson reports following a feedback session?
- b) Are their formal follow up conversations? Informal?
- c) What do you look for in walk-through observations?
- d) How do teachers receive instructional improvement within your school?
- e) How do teachers respond and engage with instructional improvement programs?

What types of instructional improvement is offered in your school?

- a) Individual instructional improvement opportunities?
- b) Group instructional improvement opportunities?
- c) School-level instructional improvement?
- d) How are these programs decided/developed?
- e) What potential barriers exist towards implementation?

In what ways is evaluation data used in your school?

- a) Probe for professional development: Could you describe how you use the evaluation data to inform instructional improvement?
- b) Walk me through your process for using the evaluation data within your building”
- c) What experience do teachers have in engaging in classroom data?
- d) Probe for barriers

Now we are going to change gears, could you tell me how professional development is planned for the year within your school?

Can you describe the process at your school for developing or choosing professional development programs for your teaching staff?

Probe: What information do you use to make these decisions?

Thanks so much for your time in answering my questions.

Appendix B

Christian Mathews, Protocol for Teacher Focus Group and Potential Probing Questions

Hello (insert name of participant):

Thank you for taking the time to talk with me. My name is _____ and I will be conducting this Focus Group. This Focus Group will be recorded for research purposes.

The purpose of this group is to investigate practices of instructional improvement and professional development within schools partnering with the Rutgers University School Systems Improvement Project. The interview will take approximately 30 minutes and will begin with a little background about yourselves then move into discussing teacher evaluation, instructional improvement, and Professional Development.

EVALUATION AND INSTRUCTIONAL IMPROVEMENT

Describe teacher evaluation within your school? Take me through the process from start to finish

- i. What resources are used?
- j. Probe for what happens in pre and post conferences and what form of feedback teachers get?
- k. Could you give me a specific example of some of the feedback that you have been given this year (gauging comfort discussing specific examples)?
- l. What happens to the feedback after a post conference (if applicable)?
- m. How is data used within your school?
- n. How is evaluation data used within your school?
- o. What barriers exist in the teacher evaluation process in school?

What kind of follow up occurs after a feedback session?

- f) What happens to CSS/Danielson reports following a feedback session?
- g) Are their formal follow up conversations? Informal?
- h) What opportunities exist for you to improve your instruction..

In what ways is evaluation data used in your school?

- e) Probe for professional development: Could you describe the link between teacher evaluation, professional development, and instructional improvement within your school?
- f) What experience do you have in engaging in data from your evaluation?

Now we are going to change gears, could you tell me about the professional development offered in your school.

Can you describe the process at your school for developing or choosing professional development programs for the teaching staff?

Probe: Is there a link between the evaluation process and formal professional development programs

Is there a link between evaluation and the opportunities offered to improve your instruction?

Thanks so much for your time in answering my questions.

Appendix C

TEES-T Items (Reddy, Dudek, Kettler, Kurz, & Peters, 2016)

Evaluation System

- 1) Evaluation system was useful
- 2) Evaluation system communicated clear expectations for classroom teaching
- 3) Evaluation system helped to improve the quality of instruction
- 4) Evaluation system was comprehensive
- 5) Evaluation system helped to improve student learning .
- 6) Evaluation system assessed effectiveness
- 7) Evaluation system improved professional growth
- 8) Evaluation system represented my instructional ability
- 9) Evaluation system informed changes in classroom practice
- 10) Satisfied with the evaluation system .

Evaluation Feedback

- 11) Evaluation feedback was useful (EF-1) Group 3
- 12) Evaluation feedback was timely (EF-27)
- 13) Evaluation feedback was specific .(EF-25)
- 14) Evaluation feedback was constructive .(EF-5) Group 3
- 15) Evaluation feedback helped improve instructional effectiveness (ef-15) Group 1
- 16) Evaluation feedback represented instructional abilities (EF 23)
- 17) Evaluation feedback informed specific changes (ef-17) (Group 1)
- 18) Evaluation feedback aligned with grade level(s)
- 19) Evaluation feedback was aligned with subject(s) (ef-9) Group 3
- 20) Evaluation feedback was aligned with the school instructional improvement goals
- 21) Evaluation feedback was aligned with the school district goals
- 22) Evaluation feedback provided information for PD opportunities .
- 23) Satisfied with the evaluation feedback .

Evaluation Process

- 24) Evaluation process was collaborative
- 25) During the feedback meeting(s), teacher(s) able to share thoughts
- 26) During the feedback meeting(s), teacher(s) encouraged to share thoughts .
- 27) Teacher(s) self-reflections on the observed lesson(s) were included in evaluation .
- 28) Prior to observation(s), teacher(s) had the opportunity to discuss lesson plan and goals .
- 29) After the observation(s), teacher(s) had the opportunity to discuss lesson plan and goals .
- 30) During the feedback meeting(s), evaluator(s) collaborated with teacher(s) on identifying professional goals.
- 31) Satisfied with the discussion(s) of teacher(s) performance.

Motivation

- 32) Motivation to Change 34 Motivated to make changes to classroom practice .

- 33) Motivated to make changes that contributed to the achievement of district goals
- 34) Teacher evaluation system increased teacher(s) motivation to change classroom practice .
- 35) The school environment supported teacher(s) commitment to change classroom practice .
- 36) Teacher evaluation system provided PD opportunities that motivated teacher(s) to change classroom practice(s) .
- 37) The teacher evaluation system provided monetary incentives that motivated change to classroom practice(s)

Appendix D. Remaining Quantitative Survey Result Tables

Table 10. The evaluation feedback was useful.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	6	3.9
Neutral	36	23.2
Agree	97	62.6
Strongly Agree	16	10.3
Total	155	100

Table 11. The evaluation feedback was timely.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	3	1.9
Disagree	8	5.2
Neutral	29	18.7
Agree	92	59.4
Strongly Agree	23	14.8
Total	155	100

Table 12. The evaluation feedback was specific.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	1	.6

Disagree	6	3.9
Neutral	30	19.4
Agree	96	61.9
Strongly Agree	22	14.2
Total	155	100

Table 13. The evaluation feedback was constructive.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	8	5.2
Neutral	30	19.4
Agree	97	62.6
Strongly Agree	20	12.9
Total	155	100

Table 14. The evaluation feedback helped to improve my instructional effectiveness.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	9	5.8
Neutral	41	26.5
Agree	86	55.5
Strongly Agree	19	11.2
Total	155	100

Table 15. The evaluation feedback represented my instructional ability.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
-----------------	------------------	----------------

Strongly Disagree	0	0
Disagree	12	7.7
Neutral	55	32.4
Agree	73	47.1
Strongly Agree	15	9.7
Total	155	100

Table 16. The evaluation feedback informed specific changes in my classroom practice.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	4	2.6
Neutral	46	29.7
Agree	88	56.8
Strongly Agree	17	11
Total	155	100

Table 17. The evaluation feedback was aligned with the National Teaching Standards.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	1	.6
Disagree	3	3.9
Neutral	26	16.8
Agree	52	33.5
Strongly Agree	42	27.1
Total	155	100

Table 18. The evaluation feedback was aligned with Core Curriculum Content Standards.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	1	.6
Disagree	4	2.6
Neutral	31	20
Agree	52	33.5
Strongly Agree	46	29.7
No Response	21	13.5
Total	155	100

Table 19. The evaluation feedback was aligned with the grade level(s) I teach.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	4	2.6
Neutral	43	27.7
Agree	88	56.8
Strongly Agree	20	12.9
Total	155	100

Table 20. The evaluation feedback was aligned with the subject(s) that I teach.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	2	1.3
Disagree	9	5.8

Neutral	35	22.6
Agree	90	52.9
Strongly Agree	19	12.3
Total	155	100

Table 21. The evaluation feedback was aligned with the school instructional improvement goals.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	3	1.8
Neutral	36	21.2
Agree	100	64.5
Strongly Agree	16	10.3
Total	155	100

Table 22. The evaluation feedback was aligned with the school district goals

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	3	1.9
Neutral	43	27.7
Agree	89	57.4
Strongly Agree	20	12.9
Total	155	100

Table 23. I was satisfied with the feedback I received from my teacher evaluation.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
-----------------	------------------	----------------

Strongly Disagree	1	.6
Disagree	8	5.2
Neutral	39	25.2
Agree	89	57.4
Strongly Agree	18	11.6
Total	155	100

Table 24. The teacher evaluation process was collaborative.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	3	1.8
Disagree	9	5.4
Neutral	47	28.3
Agree	92	55.4
Strongly Agree	15	8.8
Total	155	100

Table 25. During the feedback meeting(s), I was able to share my thoughts.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	2	1.3
Disagree	2	1.3
Neutral	26	16.8
Agree	91	58.7
Strongly Agree	34	21.9
Total	155	100

Table 26. During the feedback meeting(s), I was encouraged to share my thoughts.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	2	1.3
Disagree	7	4.5
Neutral	27	17.4
Agree	90	58.1
Strongly Agree	29	18.7
Total	155	100

Table 27. My self-reflections on the observed lesson(s) were included in my evaluation

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	4	2.6
Disagree	12	7.7
Neutral	35	22.6
Agree	86	55.5
Strongly Agree	18	11.6
Total	155	100

Table 28. Prior to my observation(s), I had the opportunity to discuss my lesson plan and goals with my evaluator.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	10	6.5
Disagree	14	8.2
Neutral	31	20
Agree	75	48.4
Strongly Agree	25	14.7

Total	155	100
-------	-----	-----

Table 29. After my observation(s), I had the opportunity to discuss my lesson plan and goals with my evaluator.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	4	2.6
Disagree	8	5.2
Neutral	21	13.5
Agree	100	64.5
Strongly Agree	22	14.2
Total	155	100

Table 30. During the feedback meeting(s), my evaluator collaborated with me on identifying my professional goals.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	3	1.9
Disagree	9	5.8
Neutral	36	23.2
Agree	85	54.8
Strongly Agree	22	14.2
Total	155	100

Table 31. I was satisfied with the discussion(s) of my performance.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	2	1.3
Disagree	9	5.8

Neutral	31	20
Agree	93	60
Strongly Agree	20	12.9
Total	155	100

Table 32. The evaluation system was useful

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	1	.6
Disagree	5	3
Neutral	53	31.9
Agree	88	53
Strongly Agree	19	11.4
Total	155	100

Table 33. The evaluation system communicated clear expectations for classroom teaching.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	1	.6
Disagree	5	3
Neutral	38	22.9
Agree	103	62
Strongly Agree	19	11.4
Total	155	100

Table 34. The evaluation system was comprehensive.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
-----------------	------------------	----------------

Strongly Disagree	2	1.2
Disagree	8	4.8
Neutral	42	25.3
Agree	95	57.2
Strongly Agree	19	11.4
Total	155	100

Table 35. The evaluation system assessed my effectiveness as a teacher.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	14	9.0
Neutral	39	25.2
Agree	89	57.4
Strongly Agree	18	11.6
Total	155	100

Table 36 The evaluation system represented my instructional ability.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	1	.6
Disagree	15	9.7
Neutral	52	33.5
Agree	72	46.5

Strongly Agree	15	9.7
Total	155	100

Table 37. I was satisfied with the evaluation system

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	2	1.2
Disagree	9	5.4
Neutral	53	31.9
Agree	90	54.2
Strongly Agree	12	7.2
Total	155	100

Table 38 . I was motivated to make changes to my classroom practice.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	6	2.9
Neutral	28	18.1
Agree	96	61.9
Strongly Agree	25	14.7
Total	155	100

Table 39. I was motivated to make changes that contributed to the achievement of district goals.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	10	6.5

Neutral	38	24.5
Agree	82	52.9
Strongly Agree	25	16.1
Total	155	100

Table 40. The teacher evaluation system increased my motivation to change my classroom practice.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	0	0
Disagree	7	4.5
Neutral	39	25.2
Agree	87	56.1
Strongly Agree	22	14.2
Total	155	100

Table 41. The school environment supported my commitment to change my classroom practice.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	1	.6
Disagree	8	5.2
Neutral	36	23.2
Agree	87	56.1
Strongly Agree	23	13.5
Total	155	100

Table. 42 The school provided professional development opportunities that motivated me to change my classroom practice.

<i>Response</i>	<i>Frequency</i>	<i>Percent</i>
Strongly Disagree	4	2.4
Disagree	20	12.9
Neutral	50	32.3
Agree	62	40.0
Strongly Agree	19	12.3
Total	155	100