ASSISTANT PRINCIPALS' PERCEPTIONS OF STUDENT PROGRESS GOALS
AND THEIR IMPACT ON INSTRUCTIONAL LEADERSHIP

by

Matthew J. Ragone
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Education

Committee:

_____ Chair

_____

_____

_____ Program Director

_____ Dean, College of Education and Human
Development

Date: _____ Summer Semester 2017
George Mason University
Fairfax, VA

Assistant Principals' Perceptions of Student Progress Goals and Their Impact on
Instructional Leadership

A Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

by

Matthew J. Ragone
Education Specialist
Virginia Polytechnic Institute, 2003
Master of Curriculum and Instruction
George Mason University, 1993
Bachelor of Arts
Mary Washington College, 1990

Director: Diana D'Amico, Professor
College of Education and Human Development

Summer Semester 2017
George Mason University
Fairfax, VA

Dedication

This is dedicated to my patient and supportive family, Cindy, Kathryn, and Michael.

Acknowledgements

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abstract

ASSISTANT PRINCIPALS' PERCEPTIONS OF STUDENT PROGRESS GOALS
AND THEIR IMPACT ON INSTRUCTIONAL LEADERSHIP

Matthew J. Ragone, Ph.D.

George Mason University, 2017

Dissertation Director: Dr. Diana D'Amico

This study examines high school assistant principals' perceptions of a single,
large, suburban school district's student achievement-based teacher evaluation policy,
and how the policy interacts with their instructional leadership capacity. As the result of
accountability policies within No Child Left Behind (NCLB) and Race to the Top (RttT),
districts across the nation have adopted teacher evaluation policies that attempt to hold
teachers directly accountable for their student's achievement. Such policies, including the
one under study, require 40% of teachers' final ratings include some measure of student
progress. This district defined student achievement through teacher created, student
progress goals (SPGs). In order to examine how this policy is perceived and implemented
at the ground level, I interviewed 16 high school assistant principals from a variety of
schools in District A. I found that a majority of these assistant principals did not consider
SPGs essential to assessing teacher quality or for holding teachers accountable for student
achievement. I also found that assistant principals' unrestricted authority, and the

presence of school teams, such as collaborative teams (CTs), impacted SPG policy implementation to the extent that the policy was modified. Finally, I found that SPG policy had a positive impact on participants' instructional leadership capacity because it created opportunities for assistant principals to have data conversations with teachers who were able to monitor struggling students, including students of ESOL and special education subgroups.

Chapter One: Introduction

Teacher evaluation polices, developed in response to Race to the Top (RttT) and state waivers to No Child Left Behind (NCLB), are potentially problematic because they appear to interfere with school-based administrator instructional leadership. Twenty-one states were awarded RttT grants ("Department of Education Awards $200 Million," 2011), and 44 states received NCLB waivers ("Center on Education Policy," 2015), resulting in new student achievement-based teacher evaluation policies that require a significant link between teacher performance and student achievement. How high school assistant principals perceive and interpret these external forces, and how they attempt to reconcile them with existing or developing ideas on instructional leadership, has not been deeply explored. Local school board interpretations of national and state student achievement-based teacher evaluation policy have led to a variety of links to student achievement. Student progress goals (SPGs) are highly subjective tools that give assistant principals significant power over how to link teacher performance to student achievement. Understanding how high school assistant principals wield this authority is the purpose of this study.

In order to investigate a possible shift in assistant principal roles – from instructional leaders to those who are now required to hold teachers directly accountable for student achievement – I studied a large, suburban school district's teacher evaluation

policy. I was curious how high school assistant principals in the district of study, hereafter referred to as District A, perceived the policy requirement that required teachers to create SPGs, the core teacher accountability portion of District A's new teacher evaluation policy. I was also curious how this new requirement interacted with assistant principal's instructional leadership mission, and how their perceptions and actions to execute it impacted policy construction. This study reveals what happens at the ground level between assistant principal and teacher, thus leading to a better understanding of how assistant principals' instructional leadership is impacted by accountability portions of teacher evaluation policy.

SPGs, sometimes referred to in the literature as Student Learning Objectives (SLOs) (Lacireno-Paquet, Morgan, & Mello, 2014; Marion, DePascale, Domaleski, Gong, & Diaz-Bilello, 2012), are relatively new components of student achievement-based teacher evaluation policies. They require teachers, under assistant principal authority, to create measures of student progress in order to assess their effectiveness on student learning. High school assistant principals are of particular interest because these policies afford administrators authority to accept or reject SPGs, thereby allowing them to construct policy as it is implemented at the ground level ("Teacher Performance Evaluation Program Handbook," 2013).

Because of the latitude provided in the evaluation policy for assistant principals to implement SPGs, unless assistant principals are studied directly, we cannot know how administrators perceive the policy's accountability measures. Knowing their perceptions leads to a better understanding of how their perceptions impact policy implementation,

and how they influence instruction. This study reveals how assistant principals, those most directly involved with SPG policy implementation, understand and implement SPG policy. Knowing what assistant principals think has the potential to provide insight for educational practitioners, policy makers, and scholars as they work to improve teaching and learning.

Research Significance

A study of high school assistant principal perceptions of SPG policy extends and synthesizes two strands of education research – teacher evaluation policy and instructional leadership. Also, this study is framed around co-construction (Datnow, Hubbard, & Mehan, 1998), rather than the more common framework used in recent teacher evaluation literature, sense-making (Carraway & Young, 2015; Spillane, Reiser, & Reimer, 2002; Weick, 1995). Therefore, the goal of this study was to make a significant and unique research contribution in two fields of research, using a conceptual framework that accounts for the impact of external forces on internal policy implementation.

Teacher evaluation scholarship is currently dominated by debate on value-added methods (VAM) (Amrein-Beardsley, 2008; Berliner, 2014; Callister Everson, Feinauer, & Sudweeks, 2013; Darling-Hammond, Newton, & Wei, 2010; Good, 2014; Hanushek, 2011; Harris, Ingle, & Rutledge, 2014; Herlihy et al., 2014; Rothstein, 2010; Sanders, 1998); correlating multiple measures of teacher quality, i.e. observation ratings and VAM data (Cosner, Kimball, Barkowski, Carl, & Jones, 2015; Hansen, Lemke, & Sorensen, 2013); and general reviews to provide districts with research-based analysis of their

newly installed evaluation systems (Rigby, 2015; Riordan, Lacireno-Paquet, Shakman,

Bocala, & Chang, 2015). Research has also collected broad survey data that alludes to

evaluator frustration from lack of time, and support from central offices (Cosner et al.,

2015; Dodson, 2015; Maharaj, 2014; Riordan et al., 2015; Ruffini, Makkonen, Tejwani,

& Diaz, 2014). Yet, according to Kowalski and Dolph (2015), "principals' opinions about

assisting teachers to develop growth plans are largely unknown" (p. 8). Further, Honig

(2006) states the keys to policy success are the "interaction between policies, people and

places – the demands specific policies place on implementers; the participants in

implementation and their starting beliefs, knowledge, and other orientations toward

policy demands; and the places or contexts that help share what people can and will do"

(p. 2). Key among these forces in this study are assistant principals who are a vital piece

of SPG policy. Their perceptions and explanations of how they interact with SPGs

complements existing research on teacher evaluation policy implementation.

Instructional leadership research, the second strand of scholarship pertinent to this

study, has evolved to a point that it can now be extended into specific accountability

situations. Early instructional leadership research defines and instructs assistant principals

and principals on how to transform themselves from managers into instructional leaders

(Davis, Ellett, & Annunziata, 2002; Edmonds, 1979; Hallinger & Murphy, 1985; Heck,

1992; Weber, 1971). In the last few years, how instructional leadership is impacted by

accountability policies has been explored (Halverson, Grigg, Prichett, & Thomas, 2015;

Le Fevre & Robinson, 2015; Lochmiller, 2016; Neumerski, 2013; Petrides, Jimes, &

Karaglani, 2014; Terosky, 2014). A thorough examination of this literature in Chapter 2

will demonstrate that we still do not have enough empirical evidence on how specific accountability policies, such as SPGs, impact evaluator philosophy and practices of instructional leadership. High school assistant principals working in a district where SPGs are a significant part of the overall evaluation policy, need to be interviewed in order to fully understand their thinking as they try to reconcile holding teachers directly accountable for student achievement using SPGs, with their philosophy on instructional leadership.

A unique aspect of this study is the application of co-construction (Datnow et al., 1998) as a lens in which to view teacher evaluation policy implementation. Co-construction assumes that policy implementation is controlled by power actors, in this case high school assistant principals, and how their actions are influenced by external forces, in this case RttT and NCLB waivers that have imposed new evaluation policies on districts (Datnow et al., 1998). Co-construction incorporates many aspects of sense-making (Weick, 1995) in that it explains how assistant principals make sense of evaluation policy. However, co-construction also accounts for how external forces continue to act upon these implementers as the policy is carried out. In other words, the imposition of federal, student achievement-based teacher evaluation policy on states and districts, requires assistant principals to combine these external factors with their existing perceptions of instructional leadership and local contextual factors in their interpretation of policy. So, not only do they attempt to make sense of the policy, they use their power, and are influenced by the force of the policy, to construct a policy unique to their context. My study was designed to examine assistant principals' perceptions of this process.

Therefore, my first research question extends teacher evaluation policy literature by asking how SPGs are perceived by high school assistant principals. My second research question asks how assistant principal perceptions impact policy implementation through the lens of co-construction. And, my third research question extends instructional leadership literature by asking how assistant principal perceptions of SPGs impact their capacity to lead instruction.

Theoretical Framework

Figure 1 represents how the external and internal forces interact to bring about teacher evaluation policy implementation and how these forces impact instructional leadership. In my view, high school assistant principals are the lynchpin between new teacher evaluation policies and all the factors that govern policy implementation and instructional leadership. Who assistant principals are, what they believe about teaching and learning, and what external forces influence their work, have impact on how they implement teacher evaluation policy and how they lead instruction.

The job of school administration has been influenced by decades of research on instructional leadership (Edmonds, 1979; Hallinger & Murphy, 1985; Halverson et al., 2015; Heck, 1992; Le Fevre & Robinson, 2015; Leithwood, Seashore Louis, Anderson, & Wahlstrom, 2004; Lochmiller, 2016; Neumerski, 2013; Weber, 1971). Further, their values, backgrounds, and perceptions on teacher accountability and instructional leadership are important to the success or failure of school reform (Cuban, 1993). Figure 1 illustrates how assistant principals are vital to teacher evaluation policy and instructional leadership, and yet they have been largely left out of the policy discussion

(Cosner et al., 2015; Maharaj, 2014; Torff, 2005). What occurs after the policy is handed down to assistant principals raises questions about what is happening at the ground level and the overall impact on teaching and learning. Figure 1 also depicts the two-way street of policy development. For example, just as the evaluation system probably has an impact on the role of the high school assistant principal, it is my hypothesis that high school assistant principals have an impact on the evaluation policy itself (Datnow et al., 1998; Honig, 2006).

*Figure 1.*The conceptual framework for studying SPGs in student achievement based teacher evaluation policy.
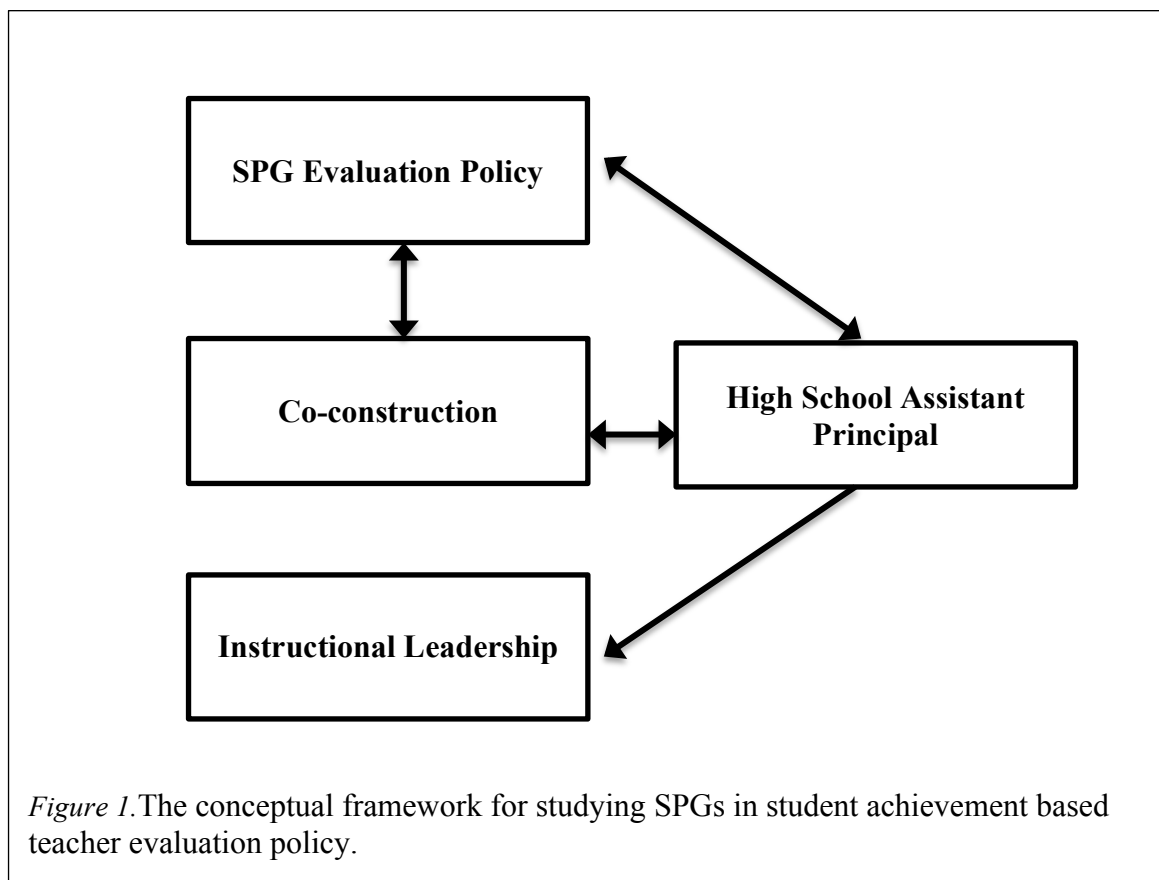
Policy Background

A comprehensive study of a large, suburban district's teacher evaluation policy requires background knowledge of several research and policy trends. In this section, I provide background on the teacher evaluation policy at the federal, state, and local levels. This brief summary provides context in which the assistant principals under study operate. Next, some background on District A's instructional leadership traditions is necessary in order to understand how its assistant principals attempt to reconcile instructional leadership roles with new teacher evaluation policies.

Federal policy. Recent federal action to reform teacher evaluation policy comes from public pressure to solve the perceived problem of teacher ineffectiveness in public education (Chetty et al., 2010; Hanushek, 2011; Lazear, 2003; Mulligan, 1999; Murnane, 2000), and research criticizing evaluation systems (Danielson, 2007; Goldhaber, Walch, & Gabele, 2014; Kane, Taylor, Tyler, & Wooten, 2011; Weisberg, Sexton, Mulhern, & Keeling, 2009). Some politicians have targeted education as the source and solution of American economic sluggishness (Bernanke, 2007; Viadero, 2008). At the national level, the problem is simply defined: scholars are in agreement that teachers are the most important in-school variable to student success (Aaronson, Barrow, & Sander, 2007; Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Mangiante, 2011; Praetorius, Lenske, & Helmke, 2012; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004; Stronge, Ward, & Grant, 2011), and teachers are not doing enough to improve student achievement (Kane, Rockoff, & Staiger, 2006; Rivkin et al., 2005). Further, previous

teacher evaluation policies have been ineffective in rooting out poor teachers or

improving teaching (Danielson, 2007; Goldhaber et al., 2014; Kane et al., 2011; Peterson,

2004; Weisberg et al., 2009). As a result political pressures based on this research, during

President Obama's administration, student achievement-based teacher evaluation systems

became federal policy through RttT (U.S. Department of Education, 2009) and waivers to

NCLB (Klein, 2012).

The national movement to create student achievement-based teacher evaluation

policy has been dominated by economist researcher influence on the national, political

agenda (Goldhaber & Brewer, 1999; Hanushek, 1971; Hanushek & Rivkin, 2010; Hoff,

2008; Kane et el., 2006; U.S. Department of Education, 2009). For example, the federal

regulation setting out *Race to the Top* procedures cites in footnotes only economists as

justification for creating strict evaluation measures for teachers (U.S. Department of

Education, 2009, p. 37806). President Obama's Department of Education chose to ignore

other scholars who object to linking teacher quality to student achievement. These

scholars say that there are too many variables beyond the teacher's control that determine

student success (Amrein-Beardsley, 2008; Berliner, 2014; Callister Everson et al., 2013;

Darling-Hammond et al., 2012; Ravitch, 2014). Nevertheless, as a result of federal

enticements, 43 states have developed new student achievement-based teacher evaluation

policies since 2008 (McGuinn, 2012).

State policy. The state under study here received its NCLB waiver in 2012

("[District A] ESEA Flexibility Map Page," 2015). Officially called the "ESEA

Flexibility Request," the waiver requires states to require local school boards to adopt

student achievement-based teacher evaluation policies that include student progress as a "significant factor" in determining teacher ratings ("ESEA Flexibility," 2012, p. 3). As a result, the state under study created a handbook for local districts entitled, *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* (2011). The vague language in the handbook suggests that the state will be flexible with districts with regard to how tightly they should hold teachers accountable for specific measures of student achievement.

These guidelines were developed in collaboration with many stakeholders including teachers and administrators. Guidance also came from a primary consultant, Dr. James Stronge ("Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers," 2011). The state policy reflects Stronge's previously published notions of teacher evaluation, including an emphasis on the use of multiple measures to counter subjective observational data and his belief that teachers should be held accountable for student learning (Stronge et al., 2011). The use of Dr. Stronge as a consultant is notable because it demonstrates the state's middle, political position regarding teacher evaluation. Stronge does not fully represent the economist camp so important to federal policy (U.S. Department of Education, 2009). However, his views do not reject linking teacher quality to student achievement that has been strongly opposed by many social scientists (Amrein-Beardsley, 2008; Berliner, 2014; Callister Everson et al., 2013; Darling-Hammond et al., 2012; Ravitch, 2014). The result of this middle position is a document that provides local districts with some flexibility in how strictly they apply teacher accountability measures.

The teacher evaluation policy in the state under study outlines seven performance standards. Standards 1-6 are standards-based benchmarks typical of previous evaluation policies. They include: Professional Knowledge, Instructional Planning, Instructional Delivery, Assessment of and for Student Learning, Learning Environment, and Professionalism. Standard 7 represents the student achievement-based component, "Student Academic Progress" ("Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers," 2011, p. 8). Though guidance from the NCLB waiver process is vague on defining "significance" of the student achievement factor, this state followed the national trend and determined that student progress would make up 40% of a teacher's final rating. The origin of 40% as the definition of "significance" comes from one requirement for RttT funding (U.S. Department of Education, 2009). No RttT or waiver state counts student achievement less than 40% ("Center on Education Policy," 2015).

The aim of this state's evaluation policy, as stated in the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* (2011), is to hold districts accountable for "incorporating student academic progress as a significant component of the evaluation while encouraging local flexibility in implementation" (p. 42). Districts may use their own measures of student progress, or a combination of these measures plus Student Growth Percentiles (SGP), as configured from End of Course (EOC) tests, provided from the state's department of education. However, because all teachers do not teach subjects with EOC tests at the end, less than 30% of teachers would have availability to this data if/when it is ready (Buckley & Marion, 2011). The policy

recommends, but does not require, that measures of student progress are "grounded in validated, quantitative, objective measures, using tools already available in the school" with an emphasis on "multiple measures" (p. 42). It also recommends schools should "consider individual teacher and schoolwide goals, and align performance measures to the goals" (p. 43).

Beyond the above guidance, the state policy is short on specifics or support for local districts. Districts are left to develop processes for collecting progress data, and a method for having goals approved. The state guidelines do not provide specific examples for SPGs or other measures of student achievement beyond standardized test data. There is also no method or plan for the state to follow up to hold districts accountable for policy implementation, nor is there a state-wide data collection system to collect, analyze, and provide feedback on SPG data. Further, aside from the link to the evaluation policy and supporting documents on the state's department of education website, the state has no structure for district support – no contact information for implementation clarification, and no mention of professional development or policy instruction seminars ("Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers," 2011).

After the publication of the *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* (2011), districts began creating new district teacher evaluation policies. By 2012, most districts in the state of study were piloting new policies. In District A, local policy makers took the more non-restrictive approach to Standard 7, opting not to require any linkage between teachers and EOC scores.

Local policy. In creating new student achievement-based teacher evaluation policy, local districts took guidance from *Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers* (2011). District A also employed Dr. Stronge, who was the primary author of their *Teacher Performance Evaluation Program Handbook* (2013). The local policy reiterates the state's general philosophy of multiple measures and linking teacher performance to student achievement. It specifies processes of collecting student progress data, and how to score teachers. However, the options for SPG data collection are so inclusive – everything from norm-referenced tests to teacher-created rubrics are permitted – it is hard to imagine any student progress measure conceived by teachers that would fall outside the policy's limits ("Teacher Performance Evaluation Program Handbook," 2013). The vagueness of the state and local policies allows for significant evaluator control over policy implementation with respect to SPGs. It is this power assistant principals have over the policy that makes a study to understand their perceptions of Standard 7 important to gaining insight as to how assistant principals use evaluation policy to improve instruction, or even if improving instruction through the policy is possible.

District A's *Teacher Performance Evaluation Program Handbook* (2013) stresses that connecting teacher performance to student learning is a high priority. For example, the policy states, "All written feedback serving as a Data Source must communicate the impact on student learning" (p. 6). In other words, Standard 7 is not the only required connection between teachers and student learning. Also, the local policy echoes state policy in making quantitative, validated growth measures already in use within the

district as a first priority data source. The handbook constantly reminds evaluators of the importance of connecting teaching to learning.

In practice, however, District A's evaluation policy is ambiguous regarding the value and importance of connecting teacher performance to student achievement. The list of suggested data sources for measuring student progress is as follows:

- Criterion-referenced tests

- Norm-referenced tests

- Standardized achievement tests

- School-adopted tests created by content experts [Collaborative Teams (CTs)]

- Performance-based measures – i.e. portfolios, rubrics, performance (p. 12)

The policy's Appendix C lists even more varied examples by subject. Some of these include, IB Exams, "Developed Pre and Post Tests," primary source analysis, textbook published assessments (p. v). These data sources indicate policy approval for teacher-created pre- and post-tests without guidance on how the data should be collected or analyzed. Despite the call to use quantitative, validated district data to measure student growth, it also states that "[District A] uses goal setting as the main measure to document student progress" (p. 10).

District A's evaluation policy does offer some guidance on what assistant principals should look for in student progress goals. Goals should be in what they refer to as the SMARTR format: Strategic, Measurable, Attainable, Results-oriented, Time-bound, and Rigorous. The single, high school example in the handbook is as follows: "For the current instructional period, 100% of students in my period 3 will make

measurable progress in biology. Each student will improve his or her pre-assessment raw score by 25 points on the post-assessment" (p. 11). The local policy offers assistant principals many options with which to advise teachers regarding data collection for student progress goals. Clearly, policy makers have made room for administrator discretion in practice; while at the same time maintaining their philosophical commitment to the connection between teacher performance and student learning.

How other districts in the state chose to interpret the state guidelines is beyond the scope of this study. However, one brief example will further illustrate District A's approach. Some districts held teachers more strictly accountable to EOC data, and created more restrictive measures of student progress where EOC data was not available. For example, City Public School's (pseudonym) evaluation handbook spends a significant amount of space discussing SGPs, while District A does not even mention them ("City Public Schools Teacher Performance Evaluation System," 2012). Again, due to the fact that all teachers do not teach courses with EOC tests, the City Public School's policy stops short of mandating the use of SGP for all teachers. But, the list of data sources other than SGPs is so much more restrictive in City Public Schools than in District A that evaluators will not be able to avoid using them in certain cases ("City Public Schools Teacher Performance Evaluation System," 2012; "Teacher Performance Evaluation Program Handbook," 2013). That District A has avoided SGPs altogether is telling for two reasons: First, it demonstrates the flexibility of the state policy. Second, it demonstrates District A's liberal take on the policy, and perhaps signals to assistant principals that linking teacher behavior to student achievement is not a serious priority.

District A's instructional leadership. A brief description on how assistant principals in District A perceive themselves as instructional leaders is necessary here to set the stage for this study. Some research indicates veteran administrators are not as equipped to be instructional leaders (Petrides et al., 2014). Indeed, in my own personal experience as an administrator in District A, veteran administrators, those who were in the job since the late 1980s and early 1990s, typically characterize themselves as managers – doling out discipline, checking attendance, fielding complaints from student and parents, and managing resources. Assistant principals coming up since 2000, have been more instructionally minded – professionals with years of teaching experience who feel confident discussing techniques with teachers, meeting with CTs, and mining student achievement data for CTs to analyze (personal experience).

Assistant principals in District A are probably more on the cutting edge of instructional leadership than in most districts around the nation. As a veteran administrator of District A, I can attest that the characterization of assistant principals in the literature as bureaucratic and managerial is not completely accurate in District A (Greenfield, 1985; Koru, 1993; Militello, Fusarelli, Mattingly, & Warren, 2015; Petrides et al., 2014). District A has over 90 high school assistant principals who evaluate teachers as one of their primary duties ("[District A] Home Page," 2015). Over the last decade, it has promoted instructional leadership in a variety of ways. First and most prominently, the district expects all its schools to be Professional Learning Communities (PLC). Second, the last two teacher evaluation policy iterations have increased classroom observation requirements. And, all teacher evaluation duties are assigned to assistant

16

principals. This development alone provides assistant principals with prime opportunities for instructional leadership. Further, District A has had a reputation for focusing on instruction from the superintendent on down (Duke, 2005). Certain structures and rhetoric have created a tradition of instructional leadership that has become ingrained in its culture. For example, the current unified vision from the superintendent to each individual school is to focus on instructional strategies to increase literacy ("[District A] Home Page," 2015).

District A adopted the PLC model in the early 2000s. Administrators and teachers were sent to many PLC conferences (personal experience). The tenets of PLC are based squarely in instructional leadership in that school administrators need to define vision, mission, and values around instruction; assist teachers in unpacking the essential information students need to know; and creating structures and strategies for teachers so they can provide opportunities for students to learn, and then re-learn, what is essential (DuFour, 2004). As a result of adopting the PLC culture, all high schools were required to organize teachers into CTs. Every subject matter team, i.e. Algebra I, English 9, Biology, is required to meet regularly, set meeting norms, set goals, collect data on student learning, analyze data, and reteach concepts not mastered by students. In District A, it is standard operating procedure for assistant principals to monitor, supervise, and support CTs. Typical high schools have 3 or 4 assistant principals, each supervising anywhere between 5 to 8 CTs. Based on their role of CT oversight, and assuming they take this job responsibility seriously, assistant principals are, in daily practice, instructional leaders.

District A assistant principals are also instructional leaders based on their work as teacher evaluators, apart from the student achievement-based portions of the system. In other words, even before the advent of SPGs, the requirement that assistant principals observe and document teacher lessons, put them in instruction leadership roles. The addition of Standard 7 does not lessen evaluator time in the classroom – in fact, the new evaluation system increases observation requirements. Again, this assumes assistant principals take their jobs as instructional leaders seriously, have the experience and confidence to suggest strategies to teachers, and otherwise support them as they try to improve their craft. The assumption, rather than having actual evidence, that all assistant principals act in this manner, makes it imperative that we examine assistant principal perceptions on student achievement-based teacher evaluation systems that might have the potential to distract them from their instructional leadership roles.

Research Questions

This study examines high school assistant principals in a single district, their perceptions of the district's student achievement-based teacher evaluation policy, how these perceptions impact the way they co-construct evaluation policy, and how these perceptions impact their instructional leadership philosophy and practice. Therefore, my research questions are as follows:

1. What are high school assistant principal perceptions of student progress goals in student achievement-based teacher evaluation policies?

2. How do these perceptions impact how assistant principals construct evaluation policy?

3. How do these perceptions impact assistant principals' capacity for instructional leadership?

As high school assistant principals make decisions about teacher-created student progress goals they contribute to the construction of the student achievement-based teacher evaluation policy. For example, their decisions on accepting or rejecting teacher-created SPGs will inform research on ground-level connections between instruction and student achievement, whether or not student achievement can be directly controlled by teacher behavior, and whether these instructional leaders think it is possible to hold teachers directly accountable for their students' achievement. Asking assistant principals what they perceive and why, asking them to explain their thought processes, and allowing them the opportunity to speak about instructional vision, has revealed new and interesting ideas about leading teaching and learning.

Definition of Terms

The following terms are relevant to this study:

Student achievement-based teacher evaluation policy – Teacher evaluation policies that are based in "significant" part on student achievement data (U.S. Department of Education, 2009, p. 9). The federal government does not specify what exactly constitutes student achievement data. The state and District A define these measures broadly to include standardized tests, teacher created assessments, norm-referenced and/or criterion-referenced tests, or school/teacher created performance rubrics ("Teacher Performance Evaluation Program Handbook," 2013).

Student progress goals (SPGs) – Standard 7 of District A's teacher evaluation policy that requires teachers to develop a single goal to measure their current students' academic progress ("Teacher Performance Evaluation Program Handbook," 2013).

Evaluation – A process that ascribes merit to the results of an observation and the collection of data pertinent to employees (Cizek, 2000).

Implementation – "transition period during which organizational members . . . become committed to innovation" (Klein & Sorra, 1996).

Value-added – Statistical model that claims to assess how much value, or the general overall affect, a teacher adds to student learning when other factors are controlled (Lavigne & Good, 2013; Sanders, 1998).

Chapter Two: Literature Review

Existing literature around teacher evaluation policy, instructional leadership, roles of assistant principals, and co-construction conceptual framework, leave space for a study on how assistant principals perceive certain aspects of teacher evaluation policy today. The goal of this study is to fully understand how assistant principals perceive and implement student achievement-based teacher evaluation policy, in the context of their instructional leadership roles. The literature review that follows illustrates how existing scholarship has left me curious about high school assistant principal perceptions on student achievement-based teacher evaluation policy, how their perceptions impact their instructional leadership, and how exactly they go about constructing policy as it is implemented.

This literature review shows that more empirical data is needed on how assistant principals reconcile federally imposed accountability notions, such as holding teachers directly accountable for their students' achievement, with their own notions of instructional leadership. In this chapter I focus on the two largest strands of scholarship, teacher evaluation policy and instructional leadership literature in order to demonstrate how my research questions complement the literature. I also review the literature around assistant principals and co-construction, as they are also relevant to my research

questions. Each of these sections describes what we already know and outlines what is left to know.

Teacher Evaluation Policy

In order to best illustrate how federal policy has impacted teacher evaluation literature, I have divided this research into two eras: pre-RttT, and post-RttT. Studies prior to RttT are consistently critical of policy effectiveness (Danielson, 2007; Goldhaber et al., 2014; Kane et al., 2011; Peterson, 2004; Spring, 2014; Weisberg et al., 2009). Studies post-RttT are consistently critical of the hasty nature in which these systems were enacted (Cosner et al., 2015; Dodson, 2015; Donaldson et al., 2014; Harris et al., 2014; Herlihy et al., 2014; Kowalski & Dolph, 2015; McGuinn, 2012; Ramirez, Clouse, & Davis, 2014; Riordan et al., 2015; Ruffini et al., 2014; Supovitz, 2012; Whitehurst, Chingos, & Lindquist, 2014; Young, Range, Hvidston, & Mette, 2015). Both eras inform this study in different ways. Pre-RttT scholarship is valuable because it helps frame the purpose of evaluations and exposes evaluation policy weaknesses. These critiques now need to be applied to student achievement-based evaluation policy. Post-RttT studies offer excellent guides for my study because they begin to examine student achievement-based components of evaluation policy. However, they do not isolate assistant principal perceptions of these components and their impact on instructional leadership. And, probably because post-RttT evaluation systems have not been around long enough to take root, these studies are clouded by the frustrations educators have over the hastily constructed policy without training or support. My study design highlights specific aspects of the post-RttT literature.

Pre-RttT. Research criticizing evaluation systems prior to RttT is prevalent. Critics point out that conflicting purposes of teacher evaluations – summative and formative – results in policy confusion (Cuban, 1993; Firestone, 2014; Goldrick, 2002; Hazi & Rucinski, 2009; Maslow & Kelley, 2012; Milanowski, 2006; Range, Scherz, Holt, & Young, 2011; Wood & Pohland, 1983). Studies also denounce pre-RttT systems for their failure to remove ineffective teachers (Danielson, 2007; Goldhaber et al., 2014; Kane et al., 2011; Peterson, 2004; Spring, 2014; Weisberg et al., 2009). Now that most states have installed new systems, concerns raised over purpose and ineffectiveness need to be reexamined.

Purpose of evaluation systems. The development and implementation of student achievement-based teacher evaluation policy is not likely to resolve scholarly debate over the purpose of evaluating teachers. However, review of the literature helps demonstrate the importance of assistant principal perceptions of evaluation systems. Economists have weighed in with their analysis that the purpose of teacher evaluation systems is to judge teacher effectiveness on student achievement (Frase & Streshly 1994; Hanushek & Rivkin, 2010; Lefgren & Sims, 2012), and to improve the teacher labor market by removing ineffective teachers (Goldhaber & Brewer, 1997; Hanushek & Rivkin, 2010; Hanushek, 2011; Heneman & Milanowski, 2003; Kimball & Milanowski, 2009). Social scientists, on the other hand, work from the assumption that evaluations are designed to improve teacher performance (Darling-Hammond, 1986; Darling-Hammond et al., 2010; DeMitchell, DeMitchell, & Gagnon, 2012; Firestone, 2014; Newton, Darling-Hammond, Haertel, & Thomas, 2010; Wise, Darling-Hammond, McLaughlin, & Bernstein, 1985).

This is just one dimension of the debate that has appeared in VAM literature (Hanushek & Rivkin, 2010; Lavigne, 2014; Newton et al., 2010; Sanders, 1998) and literature on in-school and out-of-school variables to student success (Ballou, Sanders, & Wright, 2004; Berliner, 2014; Goldhaber & Brewer, 1999; Lavigne, 2014; Milanowski, 2004).

Milanowski (2006) best exemplifies the economist view that evaluations should be summative, arguing to apply market-based strategies to the teacher labor force. He states that professional development most suited to individual teacher needs can only come through informal conversations with administrators apart from evaluation settings. Similarly, Frase and Streshly (1994) state that teacher evaluations are not in-line with professional development needs because professional development requires feedback from administrators untrained or unwilling to provide it. Whether or not assistant principals agree with these conclusions would be interesting to know.

Darling-Hammond (1986) has long represented the social scientist view that teacher evaluations have great potential to identify and target professional development to specific areas of teacher weakness. She claims that evaluation systems became bureaucratic when policy makers began to search for more objective measures with which to evaluate out ineffective teachers. Darling-Hammond concludes, "It is the bureaucratic conception of teacher evaluation . . . that limit its validity for assessing teacher performance and its utility for improving teaching" (p. 532). Research supporting the use of evaluations for formative purposes holds that potential outcomes are preferable to summative purposes that might lead to educational malpractice (DeMitchell et al., 2012), or that might unfairly judge teachers based on invalid methods (Newton et al.,

2010). Evaluator perceptions of formative evaluation policy are better understood than their perceptions of summative practices because formative strategies are intertwined with instructional leadership. I will address instructional leadership research in the next section.

The above scholarly commentary on teacher evaluation purpose is valuable, but recently the discussion has become unproductive. Economists and social scientists have solidified their positions and polarized the debate. Some scholars have even turned to personal attacks (Kupermintz, 2002; Ravitch, 2014; Sanders, 2004). Therefore, another direction for the debate is to ask what assistant principals think is the purpose of teacher evaluation. Their perceptions probably will not change economist or social scientist opinions, but since assistant principals are the ones actually implementing policy, it seems a logical direction for future research. This study posed questions to assistant principals asking what they perceived as the purpose of the evaluation policy, with particular attention on SPGs.

Failure of systems to remove ineffective teachers. Related to purpose, is the apparent failure of evaluation systems to achieve the summative purpose described above. With the advent of new accountability-based evaluation systems the next question is: Will assistant principals, who are in control of implementing student progress goals, perceive the new policy a failure if SPGs do not help remove ineffective teachers? In the era prior to RttT, research criticizing evaluation systems had been piling up since the *National Commission of Excellence in Education* (1983) report, commonly known as *A Nation at Risk*. The most damaging and oft-cited critique of pre-RttT evaluation systems

was a brief article by Weisberg et al. (2009), entitled *The Widget Effect* (Google Scholar counts 424 citations of this work). However, there are many other more comprehensive assessments of these systems. This line of research explains how we got to where we are with new student achievement-based teacher evaluation policy.

Weisberg et al. conclude from a four state survey that 99% of all teachers were rated satisfactory with a two-tiered rating system, and 94% of all teachers were rated in the top 2 categories in a multi-tiered rating system. In all systems, less than 1% were rated unsatisfactory (p. 33). Goldhaber et al. (2014) expanded on this point to argue that with no variation in how teachers are rated, there is no potential to reward the best teachers, make decisions on tenure, or to target professional development. These studies represent the market-based view of teacher evaluation policy that economists claim has been an effective strategy for the business world and should therefore be a successful approach in education as well. Other studies raise concerns about evaluation systems' perfunctory and bureaucratic processes.

Sullivan's (2012) history of teacher evaluation systems best explains traditional concerns of teacher evaluation ineffectiveness:

Teachers were evaluated by principals who sat stone-faced in the back of classrooms with a clipboard. Their subjective judgment, which could include seemingly arbitrary considerations, was the final word. They could deduct points for things as trivial as the straightness of window blinds or the angles of the push-pins on the bulletin board, and they could also take into account things like a teacher's attractiveness, manner of dress, or personal charm (p. 144).

Prior to RttT, standards-based systems prevented some of these arbitrary judgments. However, research indicates the perfunctory nature of the process continued up until post-RttT systems were put in place (Danielson, 2007; Peterson, 2004). These critiques and others also conclude that evaluator observations are not based on valid or reliable methods (Danielson, 2007; Goldhaber et al., 2014; Kane et al., 2011; Peterson, 2004). These critiques are good guides for evaluating student achievement-based systems, but research has yet to apply this analysis to new systems. A next step would be an analysis of assistant principal perceptions to see how their practices and philosophies on evaluation purpose and apparent ineffectiveness align with new systems. In short, do assistant principals believe new systems have addressed old concerns?

Post-RttT. Nearly every state has created new evaluation systems as a result of either RttT or waivers to NCLB (Collins & Amrein-Beardsley, 2013; McGuinn, 2012). Since 2009, when RttT states began to come on-line, researchers have flocked to districts and states, sometimes at education officials' request (Riordan et al., 2015; Ruffini et al., 2014), to study these systems and their effectiveness (Dodson, 2015; McGuinn, 2012). Generally, this research does not support conclusions about school administrators' ability to lead instruction. Principal perceptions of evaluations are typically obscured by frustrations over lack of training and support, and by the increased time it takes to conduct post-RttT evaluations (Dodson, 2015; Kowalski & Dolph, 2015; Riordan et al., 2015; Ruffini et al., 2014). Getting past administrator discontent with technical issues, such as navigating new computer programs to collect teacher assessment data, managing time to conduct more observations, and technical support from districts and states, is a

large challenge for any study of administrator perceptions. In the post-RttT era, studies have only scratched the surface of how administrators perceive these new systems.

State and district surveys. Many post-RttT studies are, appropriately, assessments of entire evaluation systems. Arizona (Ruffini et al., 2014), Connecticut (Donaldson et al., 2014); Colorado (Ramirez et al., 2014), Kentucky (Dodson, 2015), Missouri (Killian, 2010), New Hampshire (Riordan et al., 2015), and Ohio (Kowalski & Dolph, 2015), have all benefited from analysis of their new systems. One consistent conclusion from this research is how policies were enacted hurriedly in order to meet RttT deadlines. These actions resulted in miscommunication, confusion over policy goals, and lack of support and training for evaluators (Donaldson et al., 2014; Kowalski & Dolph, 2015; Riordan et al., 2015; Ruffini et al., 2014). Another consistent conclusion were administrator concerns over the time required to conduct evaluations effectively (Cosner et al., 2015; Dodson, 2015; Kowalski & Dolph, 2015; Riordan et al., 2015; Ruffini et al., 2014). Most new systems require more observations than pre-RttT policies, making it difficult for administrators to manage.

These studies are useful for this study because allusions to administrator discontent with evaluation systems leaves me wanting to know more about their concerns. Research focusing on individual systems must continue because districts likely benefit from the analysis. However, in order to advance teacher evaluation scholarship, eventually studies need to get beyond administrator discontent with technical issues and delve into their thoughts on evaluation systems' philosophical underpinnings. Scholarship

on teacher evaluation policy will benefit from a study of ground-level evaluators, such as assistant principals, and how they perceive accountability portions of evaluation policies.

   School-based administrator perceptions. As the research on state and district teacher evaluation policy indicates, administrators are frustrated with the lack of training, support, and time to implement new policies (Cosner et al., 2015; Dodson, 2015; Kowalski & Dolph, 2015; Riordan et al., 2015; Ruffini et al., 2014). There are, however, some hints that principal perceptions of accountability components within new evaluation policies are of particular concern to school administrators (Harris et al., 2014; Kowalski & Dolph, 2015; Young et al., 2015). Yet, these and other studies recognize there is still not enough empirical evidence to explain specifics or the nature of these concerns (Kowalski & Dolph, 2015; Maharaj, 2014; Ovando & Ramirez Jr., 2007).

   Post-RttT studies indicate principals prefer evaluation systems used for formative purposes over systems that hold teachers strictly accountable for student achievement (Harris et al., 2014; Kowalski & Dolph, 2015; Young et al., 2015). In a state-wide survey, Young et al. (2015) reported generally positive principal perceptions of systems that could help them provide valuable feedback to teachers. Kowalski and Dolph (2015), in a study of Ohio's new state policy, collected principal comments from surveys stating their concerns that too much weight was given to VAM scores on teacher evaluations. And, Harris et al. (2014) conclude that VAM evaluation systems likely conflict with "traits that school principals currently value highly" (p. 74). The implication here is that holding teachers accountable for student achievement is not aligned to principals' notions on how

to improve teacher quality or lead instruction. However, these studies were broad surveys and principals' notions are only speculative at this point.

What is missing from the literature is a deeper understanding of principal perceptions of specific accountability portions of evaluation systems. Recent studies acknowledge this missing avenue of research. For example, Kowalski and Dolph (2015) specifically note the absence of data on principal perceptions in helping teachers develop student growth plans. Likewise, Ovando and Ramirez Jr. (2007) and Maharaj (2014) call for more research to study school administrator perceptions of evaluation policy. Maharaj states, "minimal research attention has been directed to the perspective of those who are actually tasked with conducting the [evaluation] process" (p. 3-4).

Derrington (2014) studied a post-RttT evaluation policy in a Southwestern state. Like most post-RttT studies, she concluded the policy lacked support and training during implementation. However, she also indicated there are more questions to be asked about how student achievement-based systems relate to instructional leadership. Most principals in the study agreed that they became better instructional leaders because the policy required them to be in classrooms more than previous systems. "However," Derrington writes, "principals struggled to balance their belief that all teachers can grow and improve with a system of accountability that rates and ranks teachers" (p. 128). We need to know more about this struggle if we are to understand the purpose and value of SPGs. The purpose of my study is to follow up on Derrington's analysis with a study of high school assistant principals and how their instructional leadership is impacted by

accountability portions of evaluation policy, as opposed to components of the policy that are clearly designed to support instruction.

The intention of the above teacher evaluation policy studies is in-line with this study. However, these studies conclude that the rushed nature of policy development and implementation have created unsupportive atmospheres in districts. Most studies end their assessments with recommendations to add training and support for evaluators (Cosner et al., 2015; Dodson, 2015; Kowalski & Dolph, 2015; Riordan et al., 2015; Ruffini et al., 2014). Yet, they were not intended to go deeper to study assistant principal perceptions on the student achievement components of policy. These perceptions are critical because in order to understand how policy is co-constructed, we need to understand how evaluators perceive the policy and how their perceptions impact instructional leadership.

Instructional Leadership

In modern public schools, research guides school-based administrators to be instructional leaders as opposed to building administrators (Edmonds, 1979; Hallinger & Murphy, 1985; Heck, 1992; Youngs & King, 2002). With the advent of student achievement-based teacher evaluation policy, revisiting instructional leadership at the high school-level is necessary because its policy components likely have broad impact on instructional leadership. A few scholars have responded to this need (Murphy, Hallinger, & Heck, 2013; Rigby, 2015). Yet, more specific attention to assistant principal perceptions of their own instructional leadership in the context of specific accountability measures, like SPGs, is necessary as well. In this section I describe the evolution of

instructional leadership research, discuss how instructional leadership scholarship interacts with teacher evaluation policy, and summarize how modern instructional leadership research has been impacted by the accountability era. In this way, I demonstrate what we already know and what is left to learn about the impact of student achievement-based teacher evaluation policy and instructional leadership.

Evolution of instructional leadership. The role of the modern, instructional leader has developed over decades as public schools have evolved. Rousmaniere's (2013) *The Principal's Office* describes how the building principal developed from a kind of head teacher in the early 20th Century to a mid-level bureaucrat with a separate office around mid-century. Mostly job descriptions entailed ordering supplies, hiring staff, and building maintenance (Rousmaniere, 2013; Sullivan, 2012). Tyack and Cuban (1995) cite examples of school leaders struggling to overcome resistance to change. Though, with savvy and intellect, leaders were able to install co-educational programs, in-door plumbing, and other innovations throughout the century. Very little of what principals did had much to do with instruction until recently (Cuban, 1993; Rousmaniere, 2013; Tyack & Cuban, 1995).

Instructional leadership in the 1970s and 1980s was still far from a practicing reality. Though, scholars were working to change the definition of the school-based administrator. Burns's (1978) seminal work on leadership applied to educators as well as politicians. According to Burns, transformational leadership raises up leaders and followers together by moral means, with a clear, shared vision. *Effective Schools Research* follows Burns with more specific attention to school-based administrators

(Bracey, 1984; Edmonds, 1979; Hallinger, 1992). In their attempt to identify characteristics of effective schools, these authors consistently identified characteristics of school-wide instructional leadership – a vision of instruction, collaboration on instructional decisions, and a culture where all decisions are made with instruction in mind.

During the 1980s and 1990s, education theorists identified characteristics of instructional leadership in more specific terms (Andrews & Soder, 1987; Bossert, Dwyer, Rowan, & Lee, 1982; Cruickshank, 1990; Hallinger & Murphy, 1985; Heck, 1992). Cruickshank (1990) synthesized dozens of studies and their conclusions on what makes schools effective. From this analysis, common characteristics are clear: leaders that care about instruction, principal visibility in the classroom, continuous monitoring of student learning, setting high expectations for learning, and a shared emphasis by all staff on student achievement (pp. 23-37). With characteristics defined, research turned toward executing more nuanced instructional leadership models. However, by 1988, few studies had yet to establish any kind of causal link between instructional leadership and academic outcomes in education (Murphy, 1988).

Murphy's (1988) analysis of instructional leadership literature, just as it was becoming incorporated into policy, is critical of scholarship that had yet to define instructional leadership as something that could be measured. Studies up to that point did not demonstrate a causal link between so-called instructional leadership behaviors and student or school outcomes. A method to measure schools or school administrator capacity for instructional leadership had also not yet been devised. Murphy also alludes

33

to the myriad variables that impact instructional leadership capacity and outcomes on student learning. "Future work," he states, "will need to examine empirically how environmental and organizational contexts shape the instructional leadership activities of administrators" (p. 127). My study attempts to take up Murphy's suggestion by studying one particular organizational variable – SPGs.

Modern instructional leadership. In the first decade of the 21st Century instructional leadership had evolved to include characteristics such as collaboration (DuFour, 2004), problem identification (Cuban, 2001), and facilitation of instruction as opposed to leading it (Fullan, 2001). Discourse on instructional leadership has taken on many new terms – authentic leadership (Evans, 1996), symbolic leadership (Hoyle, 2002), transformational leadership (Leithwood et al., 2004), shared instructional leadership (Marks and Printy, 2003), and leadership density (Smith & Ellett, 2002). Each label characterizes instructional leadership in nuanced ways, but they all claim to have positive impacts on teaching and learning. They all also claim that these strategies will transform schools into learning institutions as opposed to teaching institutions (Rigby, 2015). However, despite the positive impacts of instructional leadership advanced by this work, there has been nothing in the scholarship to directly relate instructional leadership practices to improved teacher quality or increased student achievement because measures of teacher quality and student achievement have been so elusive (Murphy et al., 2013). Also, instructional leadership, however it is defined, cannot alone eliminate the persistent achievement gaps (Clotfelter, Ladd, & Vigdor, 2010; Hanushek & Rivkin, 2009), or improved American students' performance on international tests (Levin, 2012).

In some cases, student achievement-based teacher evaluation policy research intersects with modern instructional leadership literature. Murphy et al. (2013) conducted a deep analysis of recent studies on instructional leadership and school improvement. They note, "teacher evaluation as an explanatory or ancillary variable of interest in [a wide variety of instructional leadership studies] of school improvement was conspicuous by its absence" (p. 350). In another study, Ovando and Ramirez Jr. (2007) found that new evaluation systems positively impacted principals' ability to lead instruction in their schools because of increased observation requirements. Further, Rigby (2015) studied how messaging teacher evaluation policy through principal professional development was "the key lever to how principals made sense of their roles as instructional leaders" (p. 387). These studies are important first steps in examining teacher evaluation policy in the accountability era. The next logical step would be to include assistant principal roles in this research.

Assistant Principal Literature

Until recently there has been very little literature focused solely on the high school assistant principal. Perhaps because of the lack of studies, scholars are not yet in agreement on the instructional value of assistant principals. Some studies, even very early on, claim there is great potential for assistant principals to lead instruction through observation and teacher evaluation (Bates & Shank, 1983; Calabrese, 1991; Celikten, 2001; Greenfield, 1985; Koru, 1993; Militello et al., 2015). Though, when studies of instructional leadership or teacher evaluation are conducted, the principal is the focus of the study (Colby, Bradshaw, & Joyner, 2002; Cosner, et al., 2015; Peterson, 2004). The

dearth of scholarship on the assistant principal as instructional leaders, in conjunction with scholars' calls for more study of ground-level actors' perceptions of student achievement-based teacher evaluation policies (Harris et al., 2014; Kowalski & Dolph, 2015; Maharaj, 2014), highlights the importance of any study that would shed light on assistant principals' instructional perceptions.

In this section, I organize assistant principal research into two waves. The first wave from the 1980s and early 1990s characterizes assistant principals as taxed, bureaucratic, and overwhelmed by tedius tasks (Bates & Shank, 1983; Calabrese, 1991; Greenfield, 1985; Koru, 1993). The second wave envisions the assistant principal role in the context of emerging ideas of instructional leadership (Celikten, 2001; Militello et al., 2015). The second wave is highlighted by a comprehensive collection of literature by Shoho, Barnett, and Tooms (2012) that recognizes the lack of study on these central actors, analyzes their modern duties in the context of the accountability era, and compares assistant principal job descriptions between the 1990s and the first decade of the 21st Century. In summary, this body of work describes assistant principals as evolving toward what they most desire to be – instructional leaders, but who are still burdened and beholden to bureaucratic district structures, their principals, and a focus on accountability. Therefore, even in these later studies, the role of the assistant principal is controlled to a large extent by school context and the principal's leadership style (Celikten, 2001; Militello et al., 2015). In general, literature characterizes the assistant principal as a job with great potential, but confined by daily work.

Early assistant principal literature. Assistant principals were first recognized in the literature in the 1980s as the growing complexity of public schooling and bureaucracy from earlier decades began to over-burden the principal (Busch & MacNeil, 2012). These relatively new roles were in need of examination. Most literature of this time attempts to define the roles of the assistant principal and yearns for more substantive roles for them (Greenfield, 1985).

Greenfield (1985) is most prominent during this first wave. His call to "reconceptualize the roles of the assistant principal to be more instructionally focused, represents a common theme that lasts into the next century. Despite the desire of both scholars and practitioners for assistant principals to be more involved in academic and instructional decisions, the realities of the job continued to distract assistant principals from doing just that. Glanz (1994) surveyed sitting assistant principals and asked them to rank their duties based on how much time they spent doing them. Then, he asked them to rank the same duties according to what they thought was most important. The result is two lists diametrically opposed to each other, reflecting a serious incongruence between what assistant principals do and what they think they should do. Instructional leadership, for example, ranks 20th on the actual duty list, and 5th on the wish list. Student discipline, on the other hand, ranks 1st on the actual duty list, and 21st on the wish list (p. 285).

Two important assumptions are made in this first wave of assistant principal literature that requires reexamination in the modern context. First, that time spent on duties equals value of the duty. For example, even in today's public high schools, dealing with student discipline is the number one job of the assistant principal (Sun, 2012).

However, the assumption from earlier research is that their time could be better spent on instruction, ignoring the fact that instruction matters very little if student discipline is not addressed – if not for the sake of the misbehaving student, than for the attentive students who might be distracted by that misbehavior. The second assumption is that the highest ranking duties on Glanz's list are managerial in nature as opposed to instructional. In the modern public high school, assistant principals' perception of their duty is more important than the simple definition of it. For example, lunch duty (#2 in 1994), and parent conferences (#5 in 1994), could very well promote instruction or build positive relationships that might improve the general instructional environment. Assistant principals might see lunch duty as an opportunity to engage students, talk with them, and advise them on any number of academic topics. Certainly, depending on the assistant principal in question, lunch duty might remain a mundane and purely managerial duty.

Central to my study is a deep examination of how assistant principals perceive their duties, not just how much time they spend on them or how the title of the duty might be perceived by others. For example, one assistant principal in this study described her work with SPGs as managerial while others saw at least some instructional value in the process. Modern literature on assistant principals comes closer to recognizing the value of assistant principal attitudes toward duties, but it still has not gone into depth on their perceptions of the impact they might have on instruction.

Modern assistant principal literature. Recently, a second wave of literature on assistant principals has emerged that asks essentially the same questions as the previous wave: What do assistant principals do? And, what is their role regarding instructional

leadership? Shoho et al. (2012) compiled several studies examining assistant principals in the context of the modern era – academic accountability, social justice, ethical decision making, and their role as instructional leaders. Like the first wave, this literature paints two pictures – what assistant principals should be doing and what they want to do vs. what they actually do. The main difference between the first and second waves, however, is contextual. These studies highlight and mirror modern societal issues and portray assistant principals as being on the front lines of meaningful issues, issues of social justice, ethics, and teaching students right from wrong.

Rintoul (2012) describes student discipline and parent conferences as opportunities for assistant principals to engage in important substantive, albeit challenging, work. Where the first wave would have described these duties as managerial (Glanz, 1994; Greenfield, 1985), Rintoul's qualitative study illustrate these leaders struggling with teaching students, and in some cases entire communities, how to get along in the real world. Sometimes, assistant principals tell stories of how they balance what is good for the student vs. what is good for the school community. Sometimes, they work to bring stakeholders together to help them resolve larger community issues, such as race relations. The predicaments related during these interviews described some of the most difficult issues of modern society, issues that political and social leaders often have trouble resolving. They are not necessarily instructional, but they are far from managerial.

More recent studies have examined assistant principal roles more closely. Sun (2012) redistributed Glanz's survey to modern assistant principals and found that

instructional leadership and teacher evaluations rose significantly on the list of what

assistant principals actually do (pp. 162-163). In total, what assistant principals think they

should be doing is still far from what they actually do. However, two patterns emerged

that illustrate an evolution in thinking of the value of their work. First, assistant principals

now perceive their number one job as instructional leadership. In 1994, it was fifth.

Teacher evaluation was ranked fourth in 1994, in 2012 it ranked second. This marked rise

in perception of their own duties is perhaps in response to the instructional leadership

scholarship promoting a new ideology for school-based administrators described above

(Hallinger & Murphy, 1985; Halverson et al., 2015; Heck, 1992; Leithwood et al., 2004;

Lochmiller, 2016; Marshall, 2005; Neumerski, 2013). The second pattern is that assistant

principals actually spend more time on instructional leadership and teacher evaluations,

ranking sixth and fourth, respectively on the modern list. This pattern suggests that

school structures, principals, and school boards, have provided these opportunities for

assistant principals, recognizing their value in such endeavors.

Other modern literature on assistant principals acknowledges their potential

importance on instructional leadership. Oleszewski, Shoho, and Barnett (2012)

categorized assistant principal duties into student management, instructional leadership,

and personnel management. But, they say these roles are loosely defined and which role

they assume depends on the assistant principals' career trajectory. They suggest

redefining the assistant principal's role to be a more significant part of the school's

mission. Also, Celikten (2001) notes assistant principals can and should be instructional

leaders, but it depends mainly on clearly defined roles and the principal's support. Also,

Militello et al. (2015) state, "the principalship is fundamentally important to the success or failure of a school, and by extension, the role of the assistant principal is vital as well" (p. 195). Studies also suggest that at the high school level in particular, assistant principals have a more unique job considering the number of teachers and variety of subjects they supervise (Cosner et al., 2015; Dodson, 2015; Kowalski & Dolph, 2015).

Literature on instructional leadership and teacher evaluation, in most cases, justifies the study of principals over assistant principals because they assume principal authority over instruction and evaluation cannot be influenced by assistants (Colby et al., 2002; Cosner, et al., 2015; Peterson, 2004). Cosner et al. (2015) best represents this justification to study only the principal: "Assistant/associate principals, where they are present in schools, will likely assist in the teacher evaluation process . . . But it is the school principal who must lead and oversee the change process at the school level" (p. 78). This attitude fails to account for assistant principals at the high school level who do more than assist with evaluations. It also assumes the principal is in direct and constant control of instruction and evaluation. This is not at all what is happening in the district under study here where assistant principals conduct evaluations from start to finish and principals are not even required to sign off on final evaluations (personal experience).

The lack of scholarship on the role of the assistant principal, particularly in the context of new student achievement-based teacher evaluation policy, highlights the need for further study. Within the small body of work, scholars hint at the important potential these actors could have on the instructional environment. Additionally, as post-RttT evaluation policies require more classroom observations and data collection, their role in

teacher evaluations will only increase. Understanding assistant principal perceptions of accountability standards within student achievement-based teacher evaluation policies would shed light on aspects of the teaching and learning process. Plus, a study that focuses solely on these powerful, ground-level actors would contribute empirical data to what is presently limited scholarship.

Co-Construction

This study utilizes Datnow's (2000) theory of co-construction as a theoretical framework in order to understand how high school assistant principals implement student achievement-based teacher evaluation policy. I chose this lens because I believe it has the best chance to capture external and internal forces governing policy development and implementation. Sense-making is more frequently used as a framework to study education policy implementation (Weick, 1995). Co-construction and sense-making share many of the same ideas. However, co-construction's additional focus on external forces and power positions best fits a study that aims to understand how federally imposed teacher evaluation policy is reconciled by high school assistant principals. In this section, I review the literature of co-construction and sense-making in order to set the stage for describing my research design in the following chapter.

In order to illustrate co-construction more accurately, I compare it to sense-making. Weick (1995) developed the idea that after policy is developed and handed down to those who would use it; those actors have to somehow make sense of the policy. Crafting policy to fit the myriad contextual situations is beyond the capacity or desire of any policy maker. Further, policy developer availability after policy is constructed, to

support individual contexts it not realistic. Therefore, policy users must make decisions, sometimes in isolation and without policy developer support, in order to make the policy work in their situation.

Sense-making has been frequently used in recent literature on education policy (Carraway & Young, 2015; Coburn, 2005; Louis, Febey, & Schroeder, 2005; Spillane et al., 2002; Young & Lewis, 2015). For example, Spillane et al., (2002) come to interesting conclusions regarding teachers' cognitive structures that are based on their prior knowledge, experiences, and institutional knowledge. Similarly, sense-making has been used to help explain how school principals interpret instructional programs (Carraway & Young, 2015; Coburn, 2005; Louis et al., 2005). Carraway and Young (2015) study the sense-making approach principals used to understand why some components of a new teacher professional development program were implemented with fidelity and other components were not. For the purposes of studies that propose to define perceptions of administrators, and how these opinions might impact policy implementation, studies like these are highly relevant because they explain how actors' perceptions factor into how policy is implemented. However, in situations where there are more forces at work than the actors that implement policy, another dimension is required.

Like sense-making, co-construction helps explain administrator perceptions of policy by factoring in the forces that might influence their perceptions (Datnow et al., 1998). What co-construction adds is a "conditional matrix" (p. 2) that includes external social and political forces, as well as local contextual factors that might influence policy implementation. It also accounts for what Datnow (2000) refers to as "micropolitics" (p.

359) – the daily, often contentious relationship dramas that occur between people in offices and schools. Co-construction does not assume reform policies, such as new teacher evaluation policies, are accepted by all educators uniformly or blindly. It also does not assume all teachers and all administrators agree on the value or purpose of the evaluation process. Finally, it does not assume external factors are a one-time imposition on local contexts. Policy makers are a constant source of support or frustration just as local implementers are constantly resisting, or accepting, or adapting to change. Mixed into that dynamic are any number of micro-situations, personalities, or local events that might support or disrupt change.

Co-construction is an ideal tool to analyze reform movements, such as the one that occurred in San Diego from 1998-2002. Hubbard, Stein, and Mehan (2006) chronicled the San Diego reform movement by examining the external forces and politics that was the movement's eventual undoing. Hubbard et al. focused on ground-level actors' interpretation of external reforms and how their perceptions, local contexts, and previous notions of teaching and learning, led to their resistance of a top-down literacy program. This study did more than interpret how administrators made sense of policy; it conceptualized policy implementation within a matrix of conditions. San Diego's reform failure was not caused by the new superintendent, or anyone on his staff. The failure was explained by Hubbard et al. using a complex web of interactions between the community, the superintendent, building principals, teachers, parent communities, and students.

Similarly, in a study on externally developed school reform models, Datnow (2000) did not assume a top down implementation process. "Interrelationships among the

reform promoters and the reform adopters" are key to her analysis (p. 358). Her case studies illustrated that power and politics played large roles in reform policy adoption. Further, different local contextual factors weighed heavily on policy implementation in each situation. Co-construction is an appropriate model for research on how assistant principals perceive teacher evaluation policy because, like Datnow's study, assistant principal perceptions reveal micro-political influences on how they implement policy, as well as the factors that help them to make sense of the externally imposed policy.

Chapter Three: Research Design

This study is a qualitative examination of high school assistant principals'

perceptions of the SPG portion of a student achievement-based teacher evaluation policy,

and how these perceptions impact their capacity to lead instruction. This study extends

the literature described in Chapter 2 on teacher evaluation, instructional leadership, and

the role assistant principals play in policy implementation. There are a number of studies

that touch on evaluator perceptions of student achievement-based teacher evaluation

policy. However, these are broad survey tools (Dodson, 2015; Kowalski & Dolph, 2015),

or large-scale evaluations of entire systems conducted at the request of states or local

districts (Riordan et al., 2015; Ruffini et al., 2014), or smaller scale, yet still global

assessments of entire evaluation systems (Killian, 2010; Ramirez et al., 2014).

Perceptions held by high school assistant principals, those closest to teacher evaluation

policy implementation, "are largely unknown" (Kowalski & Dolph, 2015, p. 8). In this

chapter, I first present District A in detail as the study's setting. Then, I detail the

research design, procedures, and data collection. Lastly, I review validity threats I had to

navigate in order to design an effective study.

Setting

The setting for this study was a large, suburban county I refer to as "District A."

District A is a large school district with thousands of teachers who participate in the

evaluation system every year under one of two plans – formative or summative. *Formative* refers to informing teacher behaviors, in the same way teachers use formative assessments to identify areas of remediation for students. *Summative* refers to a summary assessment of teacher quality. Ideally, teachers in their formative years are receiving guidance or support to improve their teaching. However, there is no system to track what is learned, or how teachers are supported. Only during *summative* evaluation years are teachers actually rated. Every third year, teachers are put under *summative* review where assistant principals assist teachers in creating SPGs, conduct observations, and collect other data in order to calculate a single, final score. Teachers in their *formative* years still create SPGs with their assistant principals. However, these plans are not scored ("Teacher Performance Evaluation Program Handbook," 2013).

District A schools and teachers. District A has 24 high schools with nearly 100 evaluating administrators ("[District A] Home Page," 2015). Therefore, on average, each assistant principal is responsible for approximately 44 teachers in either their summative or formative years. Assistant principals chosen for this study had between 9 and 23 summative evaluations, with an average of 16.75 evaluations (see Figure B3). The resulting workload for assistant principals is a challenge, but not a unique one to District A. Teacher evaluation research consistently finds that school administrators are overworked and overloaded (Cosner et al., 2015; Kowalski & Dolph, 2015; Young et al., 2015). Key to this study was to look beyond the workload challenges and examine how assistant principals perceived the philosophy behind policy, or what the policy intends;

then, to discover how these perceptions enhanced or distracted from evaluators' instructional leadership.

Historically, District A has enjoyed an excellent reputation based on public perception in the housing market. ("2015 Best Counties to Live In," 2015; "These Are The 10 Best Counties In America," 2014; Toscano, 2014). Housing marketers create rankings based on imperfect and unscientific measures to feed public perception of school systems' reputation. The perception, not reality, is the pertinent point here. The result is that it raises the levels of expectation in all facets of the school district. Some District A high schools even find themselves in the national media for academic excellence. Yet, within the high end academic environments, there is also a great deal of academic and economic diversity among District A schools. Table 1 illustrates this economic diversity. Further, in prior years, several District A high schools had received academic warnings based on state tests. This study was designed to include participants who represent this broad range of schools.

District A assistant principal roles. The role of assistant principals in District A are typical across schools. All District A schools assign assistant principals to oversee teacher evaluation and supervision by departments, e.g. Social Studies, English, etc. All assistant principals in the selection pool evaluate teachers as one of their primary assigned roles. Some principals assign discipline or attendance duties to other pseudo-administrative positions, such as a Dean of Students or an administrative intern. Other schools have an assistant principal exclusively designated to supervise Special Education,

or specialized career and technical programs. These specialized administrators were only included in the participant pool if one of their assigned tasks was to evaluate teachers.

District A assistant principals differ in their assigned duties from school to school. Apart from evaluating teachers of differing departments, they have different students to supervise – some schools divide the supervision of students by alphabet, some by grade-level. Assistant principals have different special projects to manage, or different technical or managerial tasks such as creating bell schedules or maintaining the school website. However, since 2012, the first year of the new student achievement-based teacher evaluation policy, evaluating teachers has become one of the most time consuming and important roles for every assistant principal invited to this study. The co-construction theoretical framework acknowledges and accounts for these internal factors that might influence how assistant principals perceive student progress goals and their impact on instructional leadership.

Socio-economic conditions. Table 1 illustrates the percentages of students who qualify for the federal Free Reduced Lunch (FRL) program (The data omit District A's magnet and alternative high schools). The difference in District A high schools is vital to the method of this study in that one of the important parameters for participant selection was to draw assistant principals from a range of schools based on socio-economics.

Table 1

District A High Schools – Free and Reduced Lunch Averages

High School Code	Free Reduced Average
HS 10	63.79
HS 23	56.87
HS 20	56.55
HS 1	55.09
HS 17	53.13
Nat Avg	48.10
HS 21	42.32
HS 18	39.36
HS 8	37.61
State Avg	36.70
HS 3	31.99
HS 13	30.07
Dist Avg	27.48
HS 22	26.75
HS 12	25.76
HS 24	25.13
HS 11	19.12
HS 7	17.50
HS 19	17.38
HS 5	16.33
HS 14	13.29
HS 6	12.67
HS 9	11.03
HS 15	10.75
HS 4	9.77
HS 2	9.24
HS 16	2.17

Note. National average available at https://nces.ed.gov/programs/digest/d12/tables/dt12_046.asp. State average available at http://doe.[state].gov/support/nutrition/statistics/free_reduced_eligibility/2015-2016/divisions/frpe_div_report_sy2015-16.pdf . School statistics from District A school profiles available at http://commweb.[districta].edu/directory/. ("Digest of Education Statistics, 2012," 2012)

According to the literature, school and community context are paramount to policy implementation (Fuhrman, Clune, & Elmore, 1991; Honig, 2006). FRL rates cannot describe all the possible contextual nuances of school and community context, but it did help guide participant selection so that I was able to interview assistant principals from a variety of socio-economic settings.

Table 1 shows that 5 District A high schools are above the national FRL average and 7 are above the state average. Ten schools are above the district average and 14 schools are below it. Certainly there are poorer districts in the nation, but the FRL statistic illustrates the depressed pockets within a generally affluent, suburban county. Because socio-economic conditions might factor into how assistant principals perceive their instructional leadership roles, it was important not to exclude assistant principals serving poorer neighborhoods. In order to obtain a good pool of participants representing low SES communities, I defined *Low SES Schools* as those that have FRL rates above the district average. The 10 schools above the district average provided enough opportunity to include assistant principals from low SES communities. Even though HS 13, for example, is above the national average and might not be considered "Low SES" by comparison, this study compares assistant principal perceptions as compared to other assistant principals within District A, not assistant principals nationally. I detail the participant selection process below.

Assistant principal years of experience. The setting of District A can also be characterized as having a well experienced corps of school-based administrators. Of the nearly 100 assistant principals, only 28 are in their first 5 years on the job. Years of

administrator experience is important to the design of my study because, according to

Petrides et al. (2014), veteran administrators tend to view themselves more as managers

than instructional leaders. In order to have a true representation of the district's

evaluation policy and its impact on instructional leadership, it is important to obtain

perspectives of both veteran and less experienced administrators. As with SES, an exact

equal representation is not necessary, so long as both perspectives are included.

Years of experience is also important because District A's new evaluation system

was implemented in 2012. Assistant principals with fewer than 5 years of experience will

only know this new policy. For example, when asked to analyze how they are

implementing the policy and what impact it is having on their instructional leadership

capacity, those with more than 5 years of experience were able to compare the old system

with the new system. Table 2 details existing District A assistant principals by categories.

I put assistant principals in the *High Experienced* category if they have been on the job

for more than 5 years, and in the *Low Experience* category if they have been an assistant

principal for 5 years or less.

Table 2 also describes the entire District A assistant principal corps by gender and

those who represent low or high SES school communities. A detailed analysis of how

gender impacts SPG perceptions or instructional leadership is beyond the scope of this

study. However, the descriptive information is vital in order to best describe the

participant pool, and might be important as others might consider future research

questions.

Table 2

District A Assistant Principals – Descriptive Statistics.

Category	# of AP
ALL	98
Women	55
Men	43
Minority	24
White	74
Low Exp < 5 yrs.	27
High Exp > 5 yrs.	71
Low SES	42
High SES	56

Note. Data compiled from individual school websites and personal knowledge of assistant principals and schools.

Research Design

I have based this research design on my goal of understanding what high school assistant principals are thinking when they implement *Standard 7: Student Progress Goals* of the district's evaluation policy ("Teacher Performance Evaluation Program Handbook," 2013). Assuming District A assistant principals are similar to those described in the literature (Celikten, 2001; Militello et al., 2015; Petrides et al., 2014), this design is intended to get beyond their frustrations with technical aspects of the evaluation system, e.g. not enough time, lack of support, etc., and to understand their perceptions of the philosophical underpinnings of the policy with regard to teaching and learning. In this way I hope to contribute valuable knowledge not just on how assistant principals implement policy, but how they might use their understanding to impact

instruction. Discovering how SPG policy impacts their capacity to lead instruction is the ultimate goal.

In this section, I explain my research design. Included here are procedures, detailed descriptions of participants, data collection methods, and data analysis methods. Throughout the design phase of this study, I employed Maxwell's (2013) interactive research design method. This method emphasizes flexibility between goals, conceptual framework, research questions, validity, and methods (see Figure 2). The interactive research design is intended to be recursive whereby all components are subject to change and revision based on how the process proceeds. According to Maxwell, in order to fully explore the phenomena, it is necessary for the researcher to continuously reexamine all components of the interactive research design.

Figure 2 illustrates the goals of this study. These are related to purpose and significance of the study described in Chapter 1. I believe an understanding of assistant principals' perceptions of student achievement-based teacher evaluation policy contributes valuable information to several fields of research. In addition, I am confident the study will contribute knowledge and expertise to the entire District A education community when the results are shared with them. Regardless of the study's conclusions, interviewing and conversation with assistant principals about their evaluation and instructional practices was productive. We shared knowledge, reflected on evaluation policy practices, and discussed instructional leadership ideas and philosophies. The practice itself seemed beneficial for participants as well as for me.

Goals
- Contribute empirical evidence to teacher evaluation, instructional leadership, & assistant principal literature.
- Provide district leaders with evidence to inform evaluation & instructional leadership practices.
- Provide an opportunity for assistant principals to reflect on practices and share ideas.

Conceptual Framework
- APs have power/authority to construct SPG policy.
- AP's IL is impacted by SPG policy regardless of previous notions of IL or how they implement policy.
- SPGs are ineffective tools in measuring teacher quality, student achievement, and removing ineffective teachers.

Research Questions
1. What are high school assistant principals' perceptions of student progress goals?
2. How do assistant principals perceptions contribute to policy co-construction?
3. How do these perceptions impact assistant principal capacity for instructional leadership?

Methods
Qualitative Study:
- Semi-structured interviews of high school assistant principals.
- Participant selection from variety of schools and with varying years of experience.

Validity
- Checks on researcher bias through researcher memos, transparency of biases.
- Checks on participant insincerity of answers due to researcher's position as principal in the system of study.

Figure 2. Interactive Research Design. Adapted from Maxwell, J. A. (2013). Qualitative *research design: An interactive approach* (3rd ed.). SAGE Publications, Inc.

The conceptual framework is what I believe is currently happening at the ground level of policy implementation. My hypothesis is that effective evaluation policy is determined by how the policy is implemented, and that high school assistant principals possess powerful control over policy implementation. The methods, described in more detail below, are how I answered the research questions, achieved the goals of the study, and described what was happening. Threats to validity were particularly interactive in that the threats themselves were constantly presenting themselves, and methods to address them were subsequently conceived. Finally, my research questions kept me centered on what I wanted to know – assistant principal perceptions of student progress goals, how these perceptions impact their capacity for instructional leadership, and how assistant principals co-construct evaluation policy as they implement it. Next, I describe the elements of the study in detail.

Participants. Participants for this study were assistant principals employed at 24 traditional high schools in District A. "Traditional" refers to typical, state-accredited schools. I excluded the district's technology magnet school and the four alternative high schools serving students with behavioral or emotional challenges, because these institutions create very different experiences for administrators than those at traditional high schools. For this same reason, "traditional" also excludes academy administrators. These are large vocational programs within high schools that draw students from all over the district and are managed by administrators equivalent to assistant principals. The potential pool of participants in District A included 98 evaluating assistant principals (see Table 2). To avoid direct conflicts of interest, I excluded several assistant principals with

whom I had previous relationships. Relationships included current and former

subordinates, former colleagues, and assistant principals at the school where my children

attend. I also excluded an assistant principal who had applied, but was not selected, to be

an administrator at my current school.

Socio-economic communities and years of administrator experience were key

variables in participant selection because my assumption prior to interviews was that

these variables, in the context of this study, would have major implications on assistant

principal perceptions. Research indicates that years of administrator experience determine

how assistant principals perceive their instructional leadership roles (Glanz, 1994; Sun,

2012). Socio-economic status of the communities served by high schools also has impact

school administrator roles (Celikten, 2001; Militello et al., 2015).

Three other participant descriptors are worth noting: gender, ethnicity, and school.

Gender likely has an impact on instructional leadership and perceptions of policy.

However, it was my assumption entering into the study that years of experience and

school SES were more important than gender, and that gender would have little to do

with assistant principal perceptions in this case. Nonetheless, on the chance that my

assumption would be wrong, I monitored gender during the selection process to ensure

the participant list wasn't too heavily weighted male or female. By the terms of my

agreement with District A to conduct this study, I could not ask participants for their

ethnicity. I did not assume ethnicity would have anything to do with assistant principal

perceptions in any case.

Finally, I was careful to not invite more than two assistant principals per District A school. Co-construction theory hypothesizes that any number of school-level factors might impact policy implementation. Therefore, more than two participants from a single school might over-represent practices at that school. Practices, or elements of school culture, such as PLC commitment, or alignment of SPGs to school improvement goals, as it turns out, were very important to how the SPG was co-constructed. Additionally, interviewing more than two assistant principals at a single school would have been an undue burden to that school.

Participant invitation and selection. After exclusions were applied, as explained above, I invited 83 assistant principals to participate via email (see Appendix E). Response to my invitation was strong. I sent all 83 emails on October 18, 2016 to District A assistant principals. By 2:00 p.m. the next day, I had received 27 replies expressing interest in participating, and six more emails of interest after the first week. The high level of interest was surprising and may be an indication of the importance of this study to District A administrators.

To obtain the cross section of assistant principals according to the variables I considered important – years of administrator experience and school SES – I organized interested participants into 4 categories illustrated by Table 3. Each interested assistant principal was given a randomly generated pseudonym. Because I wanted to have an equal number of participants from each of the 4 categories – High SES/Low Experience, Low SES/Low Experience, High SES/High Experience, and Low SES/High Experience – and

because District A restricted the total number of participants to 19 during the IRB

process, 16 was the ideal maximum number of participants I could select for interview.

Table 3

Participant Selection Grid

	High SES	Low SES
Low Years of Experience <5 years	Deborah Ballard, HS 4 Kathryn Campbell, HS 6 Samantha Carney, HS 14 Diane Collins, HS 9 Doris Cook, HS 2 Ann Evans, HS 5 Tim Hughes, HS 7 Maria Miller, HS 12 Betsy White, HS 15 (N = 9)	Eric Alexander, HS 23 Michael Cox, HS 21 Melissa King, HS 21 Mark Sanders, HS 3 Deborah Walker, HS 1 (N = 5)
High Years of Experience >5 years	Jeffery Butler, HS 7 Joyce Edwards, HS 9 Jose Foster, HS 12 Raymond Griffin, HS 12 Anna Hill, HS 22 Kimberly Lee, HS 5 Amy Lopez, HS 22 Julie Morris, HS 7 Amanda Nelson, HS 2 Heather Rogers, HS 7 Brenda Wright, HS 16 (N = 11)	Shirley Allen, HS 20 Polly Baker, HS 3 Ronald Coleman, HS 1 Carol Garcia, HS 18 Anthony Jenkins, HS 18 Patricia Johnson, HS 10 Paul Kelly, HS 3 Ruth Mercado, HS 18 (N = 8)

Note. HS refers to High School. District A schools were assigned random numbers. An assistant principal was selected at random from each quadrant until 4 from each category were selected, equaling total 16 participants. All names are pseudonyms (N = 33).

To start the selection process from the pool of interested assistant principals, I put the names of all interested assistant principals (33) into 4 hats according to the quadrants illustrated in Table 3. I drew a single name from each hat, and then repeated the process until I obtained 4 names per quadrant. After each selection, I was cognizant of how many male and female assistant principals I had picked, and how many from each school. When I selected a third assistant principal from the same school, I discarded that name and selected another name from the same hat. The male/female ratio remained mostly equal throughout the process, so I did not have to alter the selection process to create an acceptable balance of male/female. The total population of District A assistant principals, before exclusions (98), skews female, 56% female to 44% male (see Table 2). At the end of the selection process, I ended up with 7 males and 9 females. Table 4 shows the list of final participants. Appendix C details the final participants years of experience, departments they supervised, the number of summative teacher evaluations they were responsible for during the 2016-2017 school year, as well as other demographic information collected at the time of the interview.

Table 4

Final randomly selected list of participants.

	High SES	Low SES
Low Years of Experience	Deborah Ballard, HS 4 Kathryn Campbell, HS 6 Tim Hughes, HS 7 Betsy White, 15 (N = 4)	Eric Alexander, HS 23 Michael Cox, HS 21 Mark Sanders, HS 3 Deborah Walker, HS 1 (N = 4)
High Years of Experience	Jose Foster, HS 12 Anna Hill, HS 22 Amanda Nelson, HS 2 Heather Rogers, HS 7 (N = 4)	Ronald Coleman, HS 1 Anthony Jenkins, HS 18 Patricia Johnson, HS 10 Ruth Mercado, HS 18 (N = 4)

Note. HS refers to High School. District A schools were assigned random numbers. Four participants were chosen from each quadrant (Table 3). Once 2 participants were selected from a single high school, further selections from that school were eliminated. Gender selection was monitored throughout the process, but remained equitable (N = 16).

Interviews. Interviews were conducted between October 24 and November 11, 2016 (see Table 5). Interview times and locations were arranged considering participant comfort and convenience as the highest priorities. It was convenient for most assistant principals to meet in their school offices. Deborah Walker requested to meet at District A headquarters for her convenience. And, Ruth Mercado requested a phone interview after our original time fell through. The phone interview originally concerned me because face-to-face interviews were important to the research design. However, I did not notice any difference in the quality of her responses, nor in her demeanor or apparent sincerity

in answering questions. I remained especially cognizant of her answers throughout the

analysis phase of the study and still feel confident her participation is valid.

Table 5

Interview timeline and locations.

Date	Time	Participant	Location
October 24, 2016	1:00 p.m.	Jose Foster	HS 12
October 24, 2016	4:00 p.m.	Ruth Mercado	Phone
October 28, 2016	8:30 a.m.	Michael Cox	HS 21
October 28, 2016	10:00 a.m.	Ronald Coleman	HS 1
October 28, 2016	11:30 a.m.	Heather Rogers	HS 7
October 28, 2016	1:00 p.m.	Anna Hill	HS 22
October 28, 2016	2:30 p.m.	Betsy White	HS 15
October 31, 2016	11:30 a.m.	Timothy Hughes	HS 7
October 31, 2016	1:00 p.m.	Amanda Nelson	HS 2
October 31, 2016	4:00 p.m.	Patricia Johnson	HS 10
November 7, 2016	10:00 a.m.	Anthony Jenkins	HS 18
November 7, 2016	2:30 p.m.	Eric Alexander	HS 23
November 8, 2016	8:30 a.m.	Mark Sanders	HS 3
November 8, 2016	11:30 a.m.	Deborah Walker	District Office
November 8, 2016	1:00 p.m.	Kathryn Campbell	HS 6
November 11, 2016	8:30 a.m.	Deborah Ballard	HS 4

Note. HS refers to High School. District A schools were assigned random numbers. November 7 & 8 were teacher workdays. I took personal leave for other interview days/times.

Prior to each interview I presented each participant with a disclosure statement

explaining participant rights (see Appendix E). I then presented another document

explaining the intended purpose of the study, what happens to research documents and

recordings when I am through, and how my results will be reported out. Finally, I presented them with the "Permission to Audio Record" document and asked them to sign. No participants had questions or expressed any concerns about the process, or about confidentiality. In Ruth Mercado's case, I explained all of the above, received verbal consents, and then obtained signed consent forms later via scanned files through email.

Data collection. Each assistant principal was interviewed using a semi-structured interview guide (see Appendix B). I chose a semi-structured format for these interviews because my goal was to fully understand what assistant principals perceive about evaluation philosophy, process, and the policy's impact on instructional leadership. Because I did not have the luxury of long-term involvement, I attempted to craft my questions so they might elicit thoughtful and complete answers. Interviews needed to have flexibility so I could follow up on answers in order to fully understand what assistant principals were thinking.

Some descriptive information, such as gender, years of administrator experience, and subjects supervised, were obtained prior to interviews from publically available information on school websites. I verified this information at the beginning of each interview and asked additional demographic questions. My descriptive questions were as follows (see Appendix C for a displays of demographic information):

1. How long have you been an assistant principal?

2. How many teacher evaluations are you responsible for this year?

3. What departments do you supervise?

4. What was your teaching experience prior to becoming and administrator? Years of teaching experience, and subjects taught?

I recorded each interview using a voice recording application on my personal iPad. The goal was to keep interviews between 30-45 minutes. Weiss (1994) states that the "reasonable expectation" (p. 56) is for interviews to last between 90 and 120 minutes. My agreement with District A was to keep interviews brief so as not to overly burden the assistant principal or the school. Two interviews went longer than 40 minutes. All others were between 26 and 37 minutes with an average of 33.8 minutes.

The interview guide (see Appendix B) was designed around my 3 research questions. The first set of questions asked assistant principals about their perceptions of SPG policy; the second set asked about policy implementation, or their co-construction of the policy; and the last set of questions pertained to their instructional leadership and how the policy impacted their capacity to lead instruction with their teachers. Due to the semi-structured nature of the research design, follow up questions were improvised occasionally in addition to the follow-up questions that appear on the interview guide. I found myself rewording some questions for clarity after the first few interviews. In general, though, I stuck to the interview guide as it was written.

Before closing the interview, I asked participants if there was anything else they would like to share before I turned off the audio recording. Few had anything additional to add. I also asked permission to contact them if I had follow up questions. And, cognizant of any potential regrets or nervousness about their participation, I reminded

them of their right to pull out of the study if at any time they might have second thoughts. At the time of this writing, no participants have requested to be removed from the study.

Data analysis. I personally reviewed and transcribed all interviews by listening and typing comments verbatim into Microsoft Word files. As a way to organize all research materials, researcher memos, and interview transcriptions I loaded everything into NVivo software. This software is an effective tool for information management. However, it does not take the place of a reflective researcher. Theme development, coding, and recoding data was continuous as I listened to and transcribed interviews, read, reread, and thought through participant comments. In order to further reflect, especially at times where the data seemed overwhelming, I wrote memos.

I resisted suggestions or opportunities to have interviews transcribed professionally. Transcribing helped me build a narrative. Listening and typing at times was tedious. However, the several modalities employed in the exercise of transcribing, I believe, created indelible memories of participant responses. In subsequent readings of transcripts, I found myself hearing participant voices as I read. Their tone, mood, attitudes, very clearly demonstrated their passions or frustrations with certain aspects of the evaluation policy and instructional leadership. A method I used to help solidify interview memories was to immediately play back interviews in the car as I traveled to the next interview, or listen to interviews immediately when I got home. Also, I tried to schedule free time after each day of interviews so I could begin transcribing as soon as possible. For the same reasons as above, I found this helped organize the material and to

create firm memories of the interviews. I also found the tedium of transcribing complementary to the philosophical and abstract thoughts that came to me as I listened.

After transcription, I employed two different strategies of data analysis based on Maxwell and Miller (2008). First, categorization helped organize the data so I began to see patterns. Second, narrative strategies helped me get back to the bigger picture of what was happening. The first categorizing step was to develop codes and themes from transcribed interviews (see Appendix D). NVivo software was eventually valuable to this process. However, I wrote several memos prior to coding in NVivo listing possible or potential codes. Several examples from this initial list were evident from commonly repeated words or phrases from participants as I transcribed interviews: checking boxes, goal rigor, skills vs. content, assistant principal power/authority, SPG validity, general attitudes of SPGs, data conversations, just to name a few. Eventually, these were expanded, then condensed, and reorganized several times, before the list of themes and codes were finalized as displayed in Appendix D. NVivo was utilized to help compare and contrast different codes, organize codes, and helped me look for patterns between demographic information and participant responses.

In another attempt to categorize my data, I created a matrix to align participant responses to my research questions. The matrix included broad themes such as – overall perceptions of SPG, SPG validity, SPG rigor, theory vs. practice, SPG impact on instructional leadership. Next to each theme, I pulled relevant quotations from each participant. This process helped me to see patterns, similarities, and differences in the data. At this stage, initial findings were beginning to become evident.

Despite the emergence of initial findings, the larger narrative was obscured by the large quantities of data. So, the next step was to utilize several reflective and narrative strategies described in the literature (Maxwell & Miller, 2008). I began by re-reading each transcribed interview for certain themes. For example, as I read an interview, I constantly asked, "What is his/her perception of how SPG impacts instructional leadership?" Or, "Does he/she believe SPG works to hold teachers accountable for student achievement?" Then, I wrote memos for each question. The combination of listening and reflecting helped me to see the big picture of what I thought was happening with SPG policy. Chapter 4 is a full, detailed accounting of this analysis.

Validity Threats

This study presented some interesting and unique validity issues. First and foremost, I am a high school principal studying the evaluation system in his own district. This is admittedly problematic and required thoughtful consideration about how to present my conclusions. Second, my dual role as principal and researcher also raised questions about the honesty of participant responses. Third, as an educator, I have a bias about policies that attempt to hold teachers directly accountable for student achievement. In the past, I have expressed to colleagues my serious concerns about the trend toward accountability as exemplified in NCLB. Finally, research indicates that participants are distracted or biased by technical flaws to the system. This study attempts to get beyond participant distractions and assess their core beliefs of SPG policy and instructional leadership. These threats to validity are serious enough to require solutions built into my research design. In the end, I believe I have controlled for each of these threats such that

the resulting findings are valid and make valuable contributions to evaluation policy and instructional leadership research and practice.

Reactivity. My presence as a researcher in District A, where I also work as a high school principal, is an interesting problem. A benefit to my role as researcher in District A is that I am in a position to understand the system, and the context in which it exits, better than outside researchers. The relationships I have established with schools and administrators directly or indirectly involved in the study do not necessarily undermine my objectivity to potential research conclusions unless I close my mind to alternative explanations. Therefore, rather than hide my role, throughout the analysis and in writing conclusions, I have constantly reminded myself, and my readers, who I am and what I am thinking. Transparency has been the check to this validity issue.

Yet, the main concern here was not just how I perceived my role, it was also how participants reacted to me, a principal, asking them questions about their job as assistant principals. There are 24 high school principals in District A. There are nearly 100 assistant principals, most of whom want to be one of those 24. I have been a teacher or administrator in District A since 1992, my entire career. I have been a teacher at four high schools, an administrator at three of them, and a principal at two. These experiences have put me in contact with many of the people I wanted to interview. I assumed they had previous notions or perceptions – positive, negative, or indifferent – about me. Understanding these perceptions could not be ascertained directly. However, by reexamining participant answers to questions, and reflecting on whether prior perceptions

were influencing answers or not, I felt I have maintained an honest analysis of participant responses.

Studying the people with whom I work raised questions about the truthfulness of participant answers. Ethnographic literature on reactivity stresses the importance of researcher acknowledgment of potentially reactive situations, not eliminating them (Glesne, 2011; Maxwell, 2013). According to Maxwell (2013), "eliminating the *actual* influence of the researcher is impossible" (p. 125. Emphasis in the original). In this study, I did not pretend those I interviewed would not be somehow impacted by my role as principal. Yet, in any study, the researcher is a "powerful and inescapable influence" (p. 125). What was most important was that I accepted the situation and constantly reflected on how it impacted participant responses. I controlled biasing my conclusions through constant reflection and attention to this potential threat to validity.

Checks to the reactivity validity threat were employed in developing interview questions; use of probing, follow up questions; and use of reflexivity strategies during the analysis and reporting phases of the study. Also, because I have been employed in District A for my entire career, I have the advantage of long-term involvement with the systems and culture of District A. Long-term involvement does not apply to the participant perceptions. However, as the results will demonstrate, District A's PLC culture, and knowledge of past evaluation systems, were key to a complete understanding of participant perceptions.

Weiss (1994) indicates that participants may shade the truth to makes themselves look better if they choose to do it, or to even lie outright. But, this can be checked by

asking participants for examples in addition to asking about general events. Some participants in this study were clearly very proud of the work they were sharing with me. They could have indeed embellished to the point of distorting the truth. However, because I asked them for examples and used follow up questions frequently, it would have been difficult for participants to maintain a distorted truth. For example, I asked participants for specific examples of good and bad SPGs. Another example was that I asked them for specific examples for how they were using SPGs to evaluate ineffective teachers.

Therefore, in this study, participant answers to open ended questions, and answers to my follow up questions helped to reveal, I believe, assistant principals' true perceptions of SPG policy, their implementation of it, and how it impacts their capacity for instructional leadership. The semi-structured nature of the interview guide was key here. It allowed for extended discussion of topics so that participants had ample opportunities to expose their beliefs with rich, detailed examples. In summary, even though the potential existed to obtain disingenuous replies, participants in this study were very forthcoming due to the opportunities created by the research design.

The reactivity validity threat was also controlled through comparison within and between interviews, as explained by Maxwell (2013). To do this, I re-read interviews for different things during the narrative phase of data analysis. At times I would re-read interviews critically looking for ambiguous statements and comparing them to other statements later in the interview. This exercise revealed different perceptions. For example, Anthony Jenkins stated at the beginning of the interview, "Oh, I fully support

the policy." Yet, his subsequent responses, including examples he shared, showed a lack of support for the policy. Conversely, Ruth Mercado expressed negative perceptions of SPG throughout most of her interview. However, at one point she provided examples of how the policy was indeed beneficial in holding teachers accountable for student achievement.

Researcher/participant relationship. Interview questions were crafted to focus on the work of assistant principals and how the evaluation policy impacts their work. The questions were not intended to judge them personally, their work ethic, for example. However, questions ran the risk of uncovering evidence that some assistant principals have not been conducting evaluations correctly. During interviews, it became clear that assistant principals, regardless of how they said they implemented the policy, still had strong opinions about it. In my view, participants had honest and frank answers of the evaluation policy. I made serious attempts to turn the interview into an interactive, meaningful discussion for them. I did this by asking follow up questions to their responses, even if they were not on the interview guide. In some cases, we digressed into interesting ideas and strategies assistant principals were engaged in outside implementing evaluations or leading instruction. The freedom they expressed to digress in interviews, I believe, disarmed participants and made them feel they weren't being judged or evaluated on the quality of their work.

My role as principal, again, weighs heavily on the researcher/participant relationship concern. I rationalized and checked for this validity threat in several ways. First, those assistant principals who have been reluctant to complete evaluations, or felt

insecure about their responsibility, probably did not volunteer to discuss evaluations with a principal in the first place. Therefore, I assumed the willing participants were those evaluators who were either proud of their work, or had strong feelings about evaluation policy and were anxious to share. The results demonstrate a broad range of opinions about SPG policy such that I am confident not all participants accepted the interview as an opportunity to complain or to brag. Even those who expressed intense dislike for the policy had good things to say about it. A few participants even seemed to reevaluate the policy during the interview. The results indicate that the open-ended nature of the semi-structured interview design was an ideal tool to check the researcher/participant validity threat.

Member checking was another way to reduce the researcher/participant validity threat. I spoke with Amanda Nelson and Betsy White in person to clarify points they had made. The conversations gave them an opportunity to elaborate as well as made them feel like they had provided complete answers. I also exchanged an email with Ruth Mercado to verify information that was garbled in the recording. Most of information gathered during interviews was enough to get a sense of participant perceptions. I feel confident here because at no time during data analysis did I never feel like I had wished I had asked different questions.

Researcher bias. Since the advent of NCLB accountability policies, I have come to believe that the worst policy decision in the history of education has been an attempt to hold teachers directly accountable for student achievement. I believe value-added methods are flawed statistical devices and that the premise on which they are founded is

flawed as well. These very strong political stances of mine threatened to bias any conclusions of a study on student achievement-based teacher evaluation policy. To check this threat, I constantly reevaluated my political believes during the data analysis portion of the study. This reassessment could not have happened without an open-minded approach to my research design.

I am not confident I have controlled researcher bias completely in this research study. I am not sure there is a way to disassociate myself from my beliefs and passions during a semi-structured interview, nor do I believe I should. However, results will demonstrate that I did the best I could to be transparent about my beliefs with participants and throughout the analysis and reporting phases of the study. I confronted my bias through reflecting memos, and in interpretation of the data. The results show that this process gave me the opportunity to re-evaluate my strong beliefs about teacher evaluation policies that attempt to hold teachers directly accountable for student achievement. As a result of this study, I too have benefited from the empirical evidence collected.

Again, reflexivity exercises helped to control my researcher bias. According to Glesne (2011), researchers engage in reflexivity to address concerns over validity of data "by inquiring into either their own biases . . . or into the appropriateness of their research methodology and methods, including concerns regarding data collected, interpretations made, and representations produced" (p. 151). These inquiries are best made in researcher memos and transparency in reporting conclusions. For example, I read interviews looking for negative perceptions, or flaws in SPG. Then, I re-read the same interviews looking for positive perceptions. In other words, instead of eliminating

presumptions I might have had about participant perceptions, I included different presumptions, thereby uncovering ideas that might have been clouded by my personal bias.

Participant distractions. A final validity threat was the possibility that assistant principals might be distracted in a number of ways. First and mainly, research indicates that school administrators are frustrated with technical aspects of new, student achievement-based teacher evaluation policies (Cosner et al., 2015; Dodson, 2015; Kowalski & Dolph, 2015; Riordan et al., 2015; Ruffini et al., 2014). These technical aspects include lack of time, training, and support for conducting new policy. Also, District A assistant principals are dealing with a new computer software program used to collect evaluation data, their discontent with which is well known. Assistant principal discontent with technical problems is not the focus of this study. What makes this study useful to research is collecting data on how assistant principals feel about the philosophical underpinnings of the evaluation system and its impact on teaching and learning. Therefore, the threat was that technical issues would distract participant responses.

The primary check to technical distractions was thoughtful development of the interview guide, and strategic redirecting of participants to the main points of the questions during the interview. I anticipated which questions might have the potential to distract assistant principals with technical issues and was ready to redirect them. Pilot interviews helped in this planning. As it turned out, participants were not at all interested in discussing technical distractions. Some mentioned the tedious nature of record keeping

in passing. But, they quickly moved on, knowing that our objective was to discuss the philosophy behind the policy and its impact on instructional leadership capacity. I was impressed with the high level of thinking and thoughtfulness with which each participant focused on the problem at hand.

The second distraction type I term *literal distractions*. Fire drills, a serious discipline issue, parent phone calls, or a call from the principal, had the potential to distract participants during interviews from giving 100% attention to answering questions. Getting assistant principals alone for a quiet 35 minutes, and at times where they did not feel overwhelmed or too tired to be thoughtful, was a challenge. This required careful thought to the interview schedule and a little bit of luck.

In scheduling interviews, I reasoned through my experience as an administrator, that Mondays – before issues began to pile up, and Fridays – when they were likely to have resolved issues of the week, were the best times to catch assistant principals at their most reflective. As Table 5 shows, I conducted the first 10 interviews on Monday, October 24, Friday, October 28, and Monday, October 31, 2016. I scheduled an hour for each interview so participants would not feel rushed. All interviews were conducted during the day so I did not have to ask them to come early or stay late. This did risk in-school distractions. But, I figured it was better not to add time to their day. Also, I benefited from the timing of having two teacher workdays on November 7th and 8th. I conducted the next 5 interviews on these days with no students in the building and relatively light duty for participants. I was lucky to not have had any major distractions during interviews.

The entire process of checking validity threats has been a matter of integrity. The process means nothing if I am not willing to be critical or suspicious of my results. My task was to establish credibility through honest and open reflection about reactivity, the researcher/participant relationship, and researcher bias. Careful reflection, constant searching for alternative explanations, comparing participant responses, and developing categorizing and narrative strategies, all served to check validity threats and keep me honest. Further, though I am confident my study represents participant perceptions in their truest light, I remain open to other interpretations that may result from future study or continued analysis of this data. For now, though, I believe the integrity of the research design, and the processes employed above, have uncovered perceptions that otherwise might have remained lost in the data.

Chapter Four: Research Findings

Studying high school assistant principals' perceptions of a large, suburban student achievement-based teacher evaluation policy, and its impact on instructional leadership addresses important research and policy questions about teaching and learning during the accountability era. As outlined in Chapter Two: Literature Review, research on how assistant principals implement teacher evaluation policy, and how such accountability policies impact instructional leadership, is limited. This study contributes needed empirical evidence to existing research on assistant principals, student achievement-based teacher evaluation policy, and instructional leadership.

Analysis of assistant principal responses to questions about SPG policy and instructional leadership resulted in three major findings: First, these assistant principals did not perceive SPG policy as essential to assessing teacher quality or to holding them accountable for student learning. Second, assistant principal power and authority and goals created by school teams, such as CTs or school improvement teams, had influence on how they co-constructed SPG policy. And third, SPGs had a positive impact on instructional leadership by targeting struggling students and by creating opportunities for data discussions with teachers about students. In this chapter, I report these findings in detail. First, a brief restatement of the research problem and questions is in order.

The research problem addressed by this study is that recent student achievement-based teacher evaluation policies, e.g. SPG policy, brought about by NCLB and RttT, appears to interfere with the development of assistant principals as instructional leaders. These policy conditions, and that research has largely ignored assistant principals' roles as evaluation policy implementers, were my main interests in designing this study.

Therefore, in order to better understand how SPG policy interacts with assistant principals' capacity for instructional leadership at the ground level, I set out to answer the following questions:

1. What are high school assistant principals' perceptions of SPGs?
2. How do assistant principals' perceptions contribute to SPG policy co-construction?
3. How do these perceptions impact assistant principal capacity for instructional leadership?

The results detailed below answer these questions and open up discussion about District A assistant principal capacity to lead teaching and learning. I begin this chapter by laying out the study's context by broadly defining participant perceptions and examining their descriptive characteristics. This analysis is important in order to fully understand responses and how characteristics might have impacted answers. The rest of the chapter is organized in the order of the research questions stated above as each relates directly to a major finding.

Setting Context

In this section, I broadly categorize participant responses and review participant characteristics in order to set context for the deeper analysis that follows. The categories I developed below are not intended to define each participant completely, just as the descriptive statistics discussed below do not define them. However, when taken together, along with the responses I report in subsequent sections, the categories and the demographic data help to reveal complete pictures of who was interviewed for this study.

General perceptions. On the whole, the quality of reflection and thoughtfulness demonstrated by each of the 16 participants was impressive. No matter what their perceptions, each of them clearly cared about the integrity of the process. Their answers reflected those of professionals who have thought deeply about this evaluation system for some time.

As participants shared their philosophies and experiences with SPG policy and instructional leadership, I began to see three distinct patterns of attitudes (see Table 6). The labels – *unsupportive*, *supportive*, and *supportive but uncertain* – equate generally to negative, positive, and uncertain. The majority of participants, 10, were in the *unsupportive* group who stated their dislike for the policy flatly, but also did not support the policy because they largely ignored it to assess teacher quality. Instead, they used other means, particularly the other six standards, to assess teachers. The second group consisted of two participants, Michael Cox and Eric Alexander. They were fully supportive of SPG, appreciating it in theory as well as in practice. Each claimed to use SPG systematically to coach and guide their teachers to improve student learning. The

third group lay somewhat in the middle. Anthony Jenkins, Patricia Johnson, Mark Sanders, and Deborah Walker stated their support for SPG policy in theory, but reported it was so misunderstood by teachers and their fellow administrators that they did not believe it was being used effectively. Or, in Mark's case, the 40% weighting applied to SPG was so flawed that it rendered the policy ineffective. Their uncertainty for SPG lay in policy implementation, as much as in the policy itself.

Table 6

General perception categories (N=16).

Category	Participants (Years of Administrative Experience)
Unsupportive	Deborah Ballard (5), Kathryn Campbell (7), Ronald Coleman (6), Jose Foster (16), Anna Hill (5), Tim Hughes (2), Ruth Mercado (6), Amanda Nelson (5), Heather Rogers (14), Betsy White (2) (N=10).
Supportive	Michael Cox (2), Eric Alexander (6) (N=2).
Supportive/Uncertain	Anthony Jenkins (16), Patricia Johnson (5), Mark Sanders (4), Deborah Walker (1) (N=4).

Unsupportive. The majority of participants did not support the SPG portion of District A's teacher evaluation policy. They spoke about their apprehensions as well as

80

demonstrated their lack of support by not using the policy to assess teachers or to hold

teachers accountable for student learning. By contrast, there was unanimous support and

approval from all participants, regardless of these categories, for the way the other six

standards were organized and measured. Betsy White made this distinction in the

following conversation:

> BW: As an administrator, I want to see three things that the teachers are strong in
>
> – relationships, pedagogy [with] adaptable strategies, and knowing their content.
>
> MR: Does [SPG] help you with any of those three things?
>
> BW: Um, to be quite honest, no.
>
> MR: You do those three things without [SPG]?
>
> BW: Yeah.

Then, Betsy listed the other standards and how they related to what she had identified as

quality teacher characteristics, "Standard 1, of course, that's the content area . . .

instructional delivery, learning environment. I think that is important to me." Timothy

Hughes was discouraged by the policy:

> It just doesn't seem fair if a kid walks through the door, and he's already at 99%
>
> of the mastery that he needs to be at, how is that a reflection of the teaching that
>
> takes place between September and June? It just doesn't seem right.

Timothy's attitude was one extreme, but most participants in the unsupportive category

felt similarly unconnected to the goals of SPG. Ruth Mercado sums up this common

opinion, "I think we still need to figure out a better way because I don't know that the

[SPG] is even helping our bottom line." Of 16 participants, 10 felt unconnected with SPG

to the point they felt it ineffective. However, two participants were in stark contrast to those in the *unsupportive* category.

Supportive. Michael Cox and Eric Alexander expressed full support for SPG in theory and in practice. Michael's first interview response was this, "I think the policy is good in terms of how they framed it. It's concise. It's specific. And, as a product, it really spurs on conversations, dialogues, discussions between the evaluator and those being evaluated." Michael's comment expresses support for the philosophical underpinnings of the policy, as well as an appreciation for how the policy is working for him. He also completely bought in to the policy.

MC: [SPG] helps me do a better job.

MR: And you believe in it.

MC: I do. I do. It's flexible, yet guiding at the same time.

Michael's efficacy with the SPG policy made it work for him. Eric Alexander expressed a similar attitude, but hinted at implementation concerns, "I like the focus being on students learning. And, I think this was much better than the old evaluation system. And, I think if it's done with some good fidelity, it can be done very well." Despite his hint that it could be implemented without fidelity, the policy worked effectively for Eric, "It gives teachers an objective view of [student progress]. It gives us ways of looking at students and making sure that students are learning what they're supposed to learn."

According to both Michael and Eric SPG is an impactful policy that works to support them as they try to improve teaching and learning at their schools. Their support

for and efficacy with SPG is what makes the policy work for them. However, neither Michael nor Eric provided examples of how SPG was actually improving teaching and learning. Each discussed how they worked within the system and intertwined SPG with other school systems, such as CT goals. When asked directly whether or not they saw SPG improving teacher quality, they both said they thought it did. But, they talked in terms of improving CTs more than improving individual teachers. Michael illustrates this point, "The majority of [SPGs] were crafted by the CT. So, [SPG] really fuels that quality CT goal. That's been one of the best products or results of having [SPG]." Therefore, even with *supportive* participants, there existed some question as to the value of SPG on assessing individual teacher performance.

Supportive but uncertain. Four participants described their attitudes and perceptions of SPG policy in more uncertain terms. Each of them expressed support for the policy in theory. However, in practice they did not see it used effectively. Anthony Jenkins's first response was, "Oh, I'm in favor of the policy. I think the idea of having some type of an evaluation system [that] is related to a teacher's effectiveness is beneficial." Yet, Anthony was preoccupied with his teachers' inability to create effective, rigorous SPGs. In his interview, the conversation turned back to specific examples of unrigorous goals five separate times regardless of the question I was asking.

Patricia Johnson and Deborah Walker had similar apprehensions about policy implementation. They felt that teachers did not understand how to create SPGs and also did not trust that their fellow administrators had the expertise to guide teachers toward effective goals. Patricia talked about teachers needing to be trained on how to write

quality SPGs, "I find that there are very few teachers [who] understand what it is to analyze data over time of an assessment." With regard to other administrators, Patricia hinted she was better equipped at guiding teachers, "because this is the first year I've had this department, the conversations are different this year. Being that I have the data background, I'm able to answer more the purpose of why we would do this." According to Patricia, the policy was not working, but it should be working. Similarly, Deborah expressed support for SPG, but was clearly skeptical of its implementation. "I believe the [SPG] is adequate if, in fact, the evaluation process is carried out the way it's supposed to be done, with the time and fidelity of implementation." Then, she went on to say, "I just put on homecoming. The month of September is a big, dark, black blur, of what the hell happened." Both Deborah and Patricia spent most of their interview time expressing confidence in their own abilities to utilize SPG effectively, but were equally concerned about how it could be implemented consistently throughout the district.

Mark Sanders also expressed support for SPG, but was apprehensive about it for a different reason. Because SPG is weighted so heavily at 40% of teachers' final ratings, Mark did not trust final ratings matched his idea of teacher quality. Many other participants in the *unsupportive* group expressed this same feeling. However, Mark's support for the policy in theory sets him apart from that group. Mark summed up the *supportive/uncertain* philosophy this way, "I appreciate that student progress is a part of our evaluation. I think that was a positive step in the right direction. I find myself uncomfortable with the weight of 40% in the teacher evaluation." Mark felt the weight skewed teacher ratings, but generally thought the process – developing SPGs with

teachers, looking at data with them, and talking about why students were achieving or not – was valuable and worthwhile enough to support SPG policy.

When combined, the first and third categories, *unsupportive* and *supportive/uncertain* respectively, total 14 of 16 participants. Each of these groups is similar in their distrust for SPG policy. Participants in the first group were distrustful of the policy itself, its motives, and the philosophy behind it. Participants in the third group were distrustful of those that use the policy, or a single component of it. Therefore, all but two participants in this study had negative perceptions about District A's SPG policy.

Descriptive Factors

The selection process, as detailed in Chapter Three, narrowed down a potential pool of 83 assistant principals to 16 interested participants. The resulting 16 interviewees varied in age, gender, administrative experience, years of teaching experience, subjects taught as a teacher, departments supervised, number of evaluations being conducted, and school SES (see Appendix C). By agreement with District A, these were the only descriptive statistics I was permitted to collect. Descriptive elements are important in order to set context. This data might also be useful for future studies in District A.

Two factors, years of administrator experience and school SES are particularly important because I had hypothesized these characteristics might have significant impact on how assistant principals perceived SPG policy and instructional leadership. It was surprising to me that these factors did not result in any patterns. The analysis below explains why.

Years of administrator experience. Research indicates that veteran administrators characterize their roles as more managerial, as opposed to instructional (Glanz, 1994; Sun, 2012). This evidence predicts distinctly different response to questions about teaching and learning. However, this study provides no evidence that veteran participants perceived SPG policy or instructional leadership differently from less experienced participants. General perceptions of the policy as I have categorized them – *unsupportive*, *supportive*, *supportive but uncertain* – do not correlate to years of administrator experience (see Table 6). Notions of instructional leadership are somewhat different among the most veteran participants, but not as these ideas relate to how the SPG policy impacts instructional leadership. In the sections below, I elaborate on perceptions of SPG policy and explain the categories in Table 6 in more detail. The purpose of this section is to explain how years of administrator experience does not seem to relate to these perceptions.

Of the 16 participants, nine had been assistant principals for five years or less. Therefore, they only have worked under District A's 2012 SPG policy (see Figure B1). Seven participants were administrators under previous policies that did not include attempts to link teacher behavior to student achievement. Three participants had over 10 year's experience – Jose Foster and Anthony Jenkins each with 16 years, and Heather Rogers with 14. Their extensive experience allowed for comparisons with a broader range of past policies. Anthony Jenkins talked about his experience with past evaluation systems and liked this one better, "because it gives us kind of that common language to focus on over the course of the school year." Jose Foster's attitude was more cynical, "I

also don't know if it's something that's here to stay. I think it's one of those things. The longer you've been around, in any system, stuff gets regurgitated with a different name on it." Yet, he was still able to see benefits to the new system, "The whole program helps you break things down for teachers." Therefore, even participants who have been around long enough to be distrustful of new policies did not seem to automatically dismiss SPG as dead on arrival.

These responses are not different from less experienced participants. Ruth Mercado's perception was nearly identical to Anthony's, "I like the rubric. I think it's common language for everybody to use in the district." Also, with only six years as an administrator, Ruth Mercado was the only one to characterize evaluations as managerial, "When I'm face to face in a summative evaluation meeting with the teacher going over [SPG] results . . . I see myself as a manager in that role." Further, Betsy White, in her second year as an assistant principal, had similar apprehensions about the policy as the veterans. In referring to the data collected by teachers for their SPGs she stated, "I think there is more to a student's growth than just the percentage." Veteran perspectives were broader and more historically based, but their perceptions of the policy's value or usefulness were not necessarily different from the others.

School SES. Another non-factor in determining participant perceptions of SPG was school SES. Participants in low SES schools answered questions similarly to participants at high SES schools. Participants acknowledged some challenges resulting from working with poorer communities. But, these were quickly dismissed as unimportant, thereby rejecting low SES as an excuse for not doing the work they strongly

felt needed to be done in any school community. It is notable that those participants who had spent their entire careers at high SES schools, perceived administrators in lower SES schools to have a tougher time implementing SPG policy. The eight participants from low SES schools, on the other hand, were very clear that expectations for SPG policy implementation did not and should not be altered because of SES conditions.

Two participants worked in both low and high SES schools in their careers, Patricia Johnson and Ronald Coleman. They both acknowledged the challenges of working with highly mobile populations. Yet, they made adjustments without lowering expectations. Patricia's comment best reflects these adjustments as she helped design SPGs with teachers:

> I would say, here [in] our school, we have some particular challenges with the fact that we have a very mobile student population. So, I can have some teachers that start with, say, a sample size of 20, and that sample size can diminish to half of that simply because of attendance, or kids moving to a different school, or a different area. So, I encourage them [to begin] with a bigger sample size.

Ronald expressed frustration at the large increase of attendance issues from his previous, high SES school, to his new, low SES school. "We had 167 kids assigned to Wednesday detention. We have close to 90 kids for Friday after school detention today. Those notices have to go out, those parents have to be contacted." However, the different nature of his work at his new school, made no difference in the way he applied the SPG policy. In response to my question about whether or not he saw a difference in how the SPG policy was applied at his old school vs. his new school, he said, "No. Actually, no.

88

My expectations are the same." Ronald, and others in the low SES group, talked about making evaluations a priority. Ronald stated, "If I get a referral from a teacher for a kid cussing in class, and I know I have to go into a classroom, I'm going to go to the classroom." SES is not necessarily correlative to discipline. However, these examples demonstrate the ability of participants to prioritize instruction over other tasks, even tasks that might appear more urgent.

It was interesting to note that participants who had only been at high SES schools had the perception that low SES is associated with more distractions to instructional leadership for assistant principals. In answering a follow up question, "Do you think socioeconomic status of the community the school is in might impact capacity for instructional leadership?" Betsy White replied:

Absolutely. I feel very fortunate that I can do more, I think, then other schools because our discipline is not the same. I mean, they have, I don't know how many discipline issues a day? And, I can sometimes go a day without one.

Having never worked in a low SES school, however, her perception is based only on what she thought happened there.

In another example, Anna Hill stated, "If you have lower socioeconomics and you have higher discipline it gets really, really hard to put your priorities with the instruction in pulling things together." But then, she went on to describe how she was behind in her evaluations and had to send a note to her teachers saying "due to bad decision-making of several students" her observation write-ups would be late. Eric Alexander, who just moved from one low SES school to another low SES school, had a unique perspective.

At [this school], instructional leadership is the most important thing. That's where

I'm trying to put my focus. If I was at my previous school, it would be discipline.

I've dealt with 10 referrals this year. That's like a day at [my previous school]. It's

not comparable. I mean, it's like, I'm amazed. And you would think the population

is not that different. Socioeconomically they are about the same. If not, we are

more on the poverty line than [my previous school]. But, like, referrals, it's like

they don't exist here.

It is not clear what factors made Eric's experiences at his two schools so different. He

wasn't sure either. Both had similar demographics. Both had new principals. Perhaps it

had more to do with how teachers treated discipline at the classroom level, or some

nuance of school culture. Whatever the case, the difference could not be explained by

SES.

Other descriptive statistics. In order to provide full context for this study, I

collected a number of other statistics including: gender, participants' teaching

experiences, subjects supervised by participants, and number of evaluations each

participant was required to conduct for this school year. The participant pool represented

broad ranges within each of these factors. Gender was somewhat equal, nine females to

seven males. Participants also had a broad range of teaching experiences. Table B2 shows

years of teaching experience ranges from 5 to 23 years; subjects taught include several

different disciplines; several participants were assistant principals at other District A high

schools; and two worked as instructional coordinators at the central district office. In my

research design, I made no assumptions about how these descriptive statistics might

90

impact outcomes. And, my analysis suggests how participants perceived SPG policy are not defined by these statistics. However, these non-factors are important to note in order to provide context as well as to provide District A with data for future study.

The number of teacher evaluations conducted had the potential to impact participant perceptions due to the distractions of a heavy workload. For example, it was reasonable to guess that the heavier the workload the more negative their perceptions of SPG policy. However, responses did not bear this out. Heather Rogers and Mark Sanders had the least amount, nine and 11 evaluations respectively. Deborah Walker, Ronald Coleman, Amanda Nelson, and Patricia Johnson each had over 20 (see Table B3). I saw no patterns in responses to questions of SPG perceptions or instructional leadership between these groups. Those participants that did talk about workload, talked about them similarly. Patricia Johnson, with 20 evaluations, said, "Never had under 20. And, that can be a challenge, along with certain discipline issues that arise at the beginning of the year." Mark Sanders, with 11 evaluations, said, "I think the constraints of an administrator's schedule . . . can get in the way of modeling with our staff what we want them to do." Yet, generally, participants did not complain about the number of evaluations they had. Similar to reactions to SES, they did not use it as an excuse.

That this study was conducted in late October, just as the first marking period was ending, might have mitigated the number of evaluations data point to some degree. In other words, it would be interesting to see if those participants with over 20 evaluations would feel similarly at the end of the school year. On the other hand, the evaluation

system is in its fifth year, and all but Deborah Walker have been through the entire cycle several times. Therefore, these administrators should know what to expect.

Participants in this study supervised a variety of departments. Figure B2 shows the breadth of departments represented. (District A administrators typically supervise more than one department). Special Education and ESOL teachers have notably different challenges with regard to SPGs and measuring the achievement of students who have more obstacles than teachers in other departments. So, assistant principals who supervise these teachers might have had differences in how they implemented SPG with them. Indeed, as I report in a later section, ESOL and special education teachers were highlighted in participant responses as needing different guidance. However, this was true for all participants regardless of the departments they supervised. Therefore, descriptive statistics of participants, the schools they work in, and the teachers they work with had little to do with the way they perceived SPG policy, its implementation, or how it impacted instructional leadership capacity. More research will be necessary to delve into why assistant principals perceive SPG the way they do. These descriptive statistics could be instrumental in those future studies. This study, on the other hand, was more concerned with how assistant principal perceptions impacted teaching and learning. I discuss these perceptions next.

Teacher Quality and Accountability

One of three major findings of this study is that most participants did not perceive District A's SPG policy essential to assessing teacher quality or to holding them accountable for student learning. This is significant because the stated purpose of District

A's policy is to do just that. It is also significant because this finding provides one example of how student achievement-based teacher evaluation policies implemented across the nation, as the result of NCLB and RttT, is not working as intended. Finally, this finding is significant because the empirical evidence collected from this study runs counter to literature suggesting that holding teachers accountable for student achievement will improve teacher quality, remove ineffective teachers, and improve student achievement (Goldhaber et al., 2014; Hanushek, 2011; Milanowski, 2006; Sanders, 1998). Participants in this study, even those most in favor of District A's policy, were fairly uniform in their perceptions of the policy's impact on teaching and learning. In this section, I analyze participant responses to questions of SPG validity and reliability, analyze participant perceptions of whether the policy is able to hold teachers accountable for student learning, and assess the impact of the policy on ineffective teachers.

SPG validity and weight. The stated goals of the 2012 District A teacher evaluation policy are to improve teacher quality and raise student achievement ("Teacher Performance Evaluation Program Handbook," 2013). However, most participants in this study did not consider SPGs essential to meeting these goals. The main reason expressed for the failure of the policy to achieve its goals is that the policy did not match their perceptions of teacher quality as measured by the policy's other six standards or their existing perceptions of teachers. Almost all participants agreed that the policy made a good attempt to hold teachers accountable for student achievement, but because teachers selected small subsets of students to measure, SPGs could not hold teachers accountable for all students, all of the time.

Validity. That SPG is not a valid measure of teacher effectiveness was a running theme throughout most participant comments. Question #10 of the interview guide asks, "Do you think there is a strong correlation between SPG ratings and teacher quality?" After the first interview, I revised this question for better clarity to: "Do you think there is a correlation between the SPG ratings teachers get and ratings they receive on the other six standards?" Ten participants stated flatly that there was no correlation. Two did not respond directly to the question. Four replied "yes" (see Table 7). However, these answers do not tell the whole story. They each had different reasons and several followed up with contradictory or qualifying statements.

The majority of participants who did not see SPG as valid saw no alternatives to teachers who could manipulate the data to benefit their chances for a higher rating. Amanda Nelson characterized this as "playing teacher."

> The teachers who know how to play teacher, are going to be able to come up with the data that will prove that they are quote/unquote effective, because they know the numbers as well as I do. So, they're going to make that work. And, I can't always catch that. Teachers who know how to play teacher . . . and you get one other highly effective [on the other six standards], you're highly effective overall. And, I can't do anything about that. And, I completely disagree with how some of those outcomes come out.

Of the four participants who answered, "yes" to question #10, several had contradictory qualifications. Patricia Johnson, who answered "yes" initially, described herself several times as a "data person." As such, she was more confident in her ability as

94

an evaluator to counter teacher manipulation of the SPG. But, she expressed concern that most assistant principals would not be able to. "I think it's important to look at students through [the SPG]. I'm just not sure it's being looked at in a way that's impacting student achievement validly." So, in her situation, which she considered ideal, SPG was valid. But, she did not believe the ideal was true in all assistant principal/teacher situations.

The two participants in the *supportive* category (Table 6), Michael Cox and Eric Alexander, initially said there was a correlation. However, Eric had reservations similar to Amanda that teachers could game the system, "If a teacher crafts a well-written [SPG], they know that they're going to make it. So, if they're savvy enough, they can definitely reach effective." Michael Cox also responded "yes," at first. But, he indicated that he hadn't thought to check whether or not SPG ratings matched a teacher's overall rating: "I have to double check that, because how would I explain your *developing* or *needs improvement* in [the SPG standard], but I've got you on the verge of *effective* or *highly effective* overall?" Therefore, even those participants most supportive of SPG still questioned SPG validity.

Michael and Eric were not as concerned with validity as the others. They perceived the SPG system, like many other systems, as imperfect but workable. Eric stated simply, "But, there is no way around that, in my opinion, the way the system is set up right now." In other words he trusted that the system afforded them enough control to overcome whatever the system could not do. Michael's general attitude was similar to Erics, even if he had not thought to exactly align SPG with the other standards.

Some participants believed SPG might incidentally align with their assessments of teacher quality, but considered it neither the first nor the most reliable indicator of teacher quality. Assistant principals used the other six standards, as well as other evidence not related to the evaluation process, to make their own assessments of teacher quality. Amanda Nelson best summed up the value of the other six standards:

AN: I think the other ones do a better job at the teacher quality piece of it.

MR: What other ones? The other standards?

AN: Yes, the other standards. So, the instructional delivery, instructional

planning, those to me when I'm going in and observing, I get much more teacher

quality.

Other participants saw SPG as an ineffective tool for measuring teacher quality because they felt it could not measure student achievement reliably. Anna Hill stated, "A strong correlation? No. Not all the time. That's not consistent at all." Consistency, or reliability, was an issue because of concerns over inconsistent expectations of goal rigor. Participants noted frequently that expectations for goal rigor were unclear and creating rigorous, effective goals was a constant challenge that clearly impacted SPG validity. I discuss rigor in more depth below.

Weight. Many participants also expressed concern over how the 40% weight of SPG impacted the validity of teacher quality ratings. Eleven of the 16 participants believed 40% was too much weight for a single standard (see Table 7). All 16 expressed confidence that the other six standards could assess teacher effectiveness, and would have liked to see those carry more weight. Jose Foster expressed this common concern, "I

96

think 40% was way too much. Because, you break down 60% of all the other things that you're doing, whether it be planning instruction, all the other the standards really, I think it is skewed." Five participants accepted 40% as a proper weight for SPG. However, two of these had concerns about the potential effectiveness of the SPG to hold teachers accountable accurately for student learning. One of these, Anna Hill, stated:

I'm not so much concerned about the 40% as I am as how it's structured and we have an authentic way of measuring student achievement that doesn't allow for so much open-endedness of how you determine what your goal is going to be.

In other words, if she had faith the SPG could do what it was designed to do, she would be okay with 40%.

On the other hand, several participants believed in their ability to control teachers' SPGs to the extent that they would result in accurate ratings. Patricia Johnson struggled with teaching her teachers to understand how to write valid SPGs, but was committed to working with them until they could come up with a quality goal that had a good chance of validly assessing their ability. She said sometimes this would take three or four meetings, but it was worth it. Eric Alexander thought 40% was appropriate, and said, "I wouldn't mind it going to 50 or 60%." His determination to link teachers to student achievement is rooted in his educational philosophy. He said, "That's why we're here. We have to. I mean, that's the most important thing. That's the job." Other participants were less inclined to judge teachers so directly on student achievement. Eric, on the other hand, believed so strongly in the link between teachers and student achievement that he

thought any system attempting to do this was on the right track and trusted his ability to

work around policy imperfections.

Table 7

Perceptions of SPG Validity and Weight

Participant	Valid?	Validity Comment	40% Weight
Eric Alexander	Yes	Can be manipulated	50 or 60%
Ronald Coleman	No		Ok
Michael Cox	Yes	Had not checked	Ok
Anna Hill	No		Ok
Deborah Walker	Yes	Relationships are more important	Ok
Deborah Ballard	No		Too high
Kathryn Campbell	No		Too high
Jose Foster	No answer		Too high
Timothy Hughes	No		Too high
Anthony Jenkins	No		Too high
Patricia Johnson	Yes	If SPGs are written correctly	Too high
Ruth Mercado	No		Too high
Amanda Nelson	No		Too high
Heather Rogers	No answer		Too high
Mark Sanders	No		Too high
Betsy White	No		Too high

Validity or reliability was a concern for almost every participant regardless of

how they responded to the correlation question. For most, their concern was strong

enough to dismiss SPG completely. Almost every participant agreed that some measure

of student progress was necessary in evaluating teachers, but either the policy or its

implementation was suspect. Eric Alexander and Michael Cox, on the other hand, saw it

differently. They were less concerned with comparing SPG ratings with other standards because they trusted all the standards would come together to paint an accurate picture of teacher quality. Unlike the rest of the participants, because they trusted SPG, they did not overanalyze the ratings.

In general, these results indicated that assistant principal perceptions of teacher effectiveness seems to come from somewhere other than student achievement data. Perhaps from the teacher's reputation, the absence of student or parent complaints, from thinking about them in the context of the other six standards, or from classroom observations – all of which are independent of the SPG process. Participants who did not see a correlation between their understanding of teacher quality and SPG ratings did not question their assessment. Instead, they chose to ignore SPG outcomes. Ronald Coleman stated flatly, "I'm not going to let a good teacher have a bad evaluation because of [the SPG]." Kathryn Campbell said, "I mean, if I see strong teaching in the other six categories, and then something doesn't fall on the [SPG], I don't take that as a major indicator." For most participants, once their minds were made up about a teacher, the SPG was not powerful enough to unseat their previously held perceptions of teacher quality.

These findings raise several additional questions about assessing teacher quality. Is teacher quality best measured by linking teacher behavior to student achievement? Or, is it best measured by assistant principals' expertise in judging them based on observations or a variety of other evidence? Also, how do assistant principals account for teachers who game the system by "playing teacher" vs. those who are not savvy enough

to do so. In other words, do teachers' final evaluation ratings reflect teacher quality or teacher savvy? I raise these points for extended discussion in Chapter Five.

Holding teachers accountable for student achievement. In addition to the concern for SPG validity by the majority of participants, the ability of the policy to hold teachers accountable also came into question. Though, there was less agreement among participants on this point. Some thought SPG was effective in getting teachers to at least consider their accountability to student learning. Others believed strongly that SPG was ineffective because it only measured some of the students some of the time. In general, even though participants said they agreed that holding teachers accountable for student achievement was important, most could not figure out how to do this fairly.

Each of the 16 assistant principals interviewed reported they thought teachers ought to be held accountable for student learning to some degree. Each of them qualified that belief by saying other stakeholders were responsible as well. Betsy White's statement represented this common thinking, "The student is responsible. I think it's a partnership. You know, not just with the teacher and the student, but the families, the administration, any resource person. I think everyone's accountable for that." Additionally, participants understood that the SPG was designed to hold teachers accountable, though they questioned its ability to do what it was intended to do. Kathryn Campbell stated:

> I think [the SPG] is ineffective in doing that [holding teachers accountable]. I
> understand that that was the purpose. But, if I'm working with a teacher that is not
> doing what they need to be doing, I'm pulling their assessments. I'm pulling their

grade reports. I'm getting student interviews. I'm getting parents. All of that tells

me much more whether that teacher is being effective than their [SPG]. I mean,

the [SPG] can align with that sometimes. But, not always.

Timothy Hughes also understood the SPG's purpose was to hold teachers accountable for

student achievement. However, he did not consider the SPG as a basis for doing this

either. Instead, he believed "student engagement" would lead to higher student

achievement, and that was best measured by observing students in the classroom. "I

believe that the data and all the numbers are just going to be a byproduct of those kids

that are engaged."

Similarly, Ruth Mercado doubted the ability of SPG to hold teachers accountable

by itself. "I don't know if the current system is the best way to even monitor

accountability." She considered SPG as only a "quick snapshot" of student achievement.

And, at her school, snapshots were based on strand data from state EOC tests, which she

did not trust as true measures of student achievement. Others expressed similar concerns

about sporadic assessments. Deborah Ballard noted, "For the most part, we're only going

to talk about [SPG] in September, in January, in May. We really spend the rest of the year

focusing on the six informal observations, and the strategies, and what's happening in the

classroom." According to these participants, teacher accountability for student

achievement occurs by assessing observed teacher performance independently of SPG.

On the other hand, participants reported that the mere existence of the SPG policy

did seem to hold teachers accountable for student learning, or at a minimum, kept them

thinking that student achievement was key to their job. Several noted that, no matter what

the results of SPG outcomes, the process was vital in getting teachers to understand the importance of collecting and reviewing student achievement data. Patricia Johnson described the process as more important than the product.

> I think it's very important for teachers to look at data throughout the year. I would say that that's the positive thing. With this goal, they are looking at data, real time data, as we go. So, I would say that I think that's helpful. It's more or less forced the conversation.

Eric Alexander described how the SPG provided a new way for teachers to look at the impact of their teaching that they did not have before under previous systems. "It gives teachers an objective view of [student learning]. Student progress helps them . . . and gives us common language, and ways of making sure that students are learning what they are supposed to be learning." Michael Cox also had a positive view of the SPG as it related to teacher accountability, "It increases and improves the level of accountability with teachers in a way that is not punitive." Assistant principal buy-in to the process, even if they were concerned about the product, was vital to holding teachers accountable for student learning, any learning. Participants who saw SPG as an opportunity to engage teachers in conversation about student learning, were making the best use of SPG.

Holding teachers accountable for student achievement has been a major expectation of accountability policies that have come on line during the accountability era. District A designed the SPG standard to give flexibility to teachers in measuring their students' achievement. Reasons for this flexibility are unstated by the policy, but one may assume policy makers were not keen on requiring the use of the state's EOC tests as

a sole measure of student achievement. Therefore, as long as plans are approved by the assistant principal, teachers are free to use any measure of their own creation. Suggested measures are listed in the policy (see p. 14 for an analysis of these suggestions). Therefore, by design, data plans in District A can be as numerous as the number of teachers under evaluation. If assistant principals do not trust SPG to assess student achievement accurately, they have to rely on other measures, or their observations of teachers and students, to hold them accountable in other ways.

SPG impact on ineffective teachers. More specific to the questions of SPG validity, weight, and teacher accountability is whether or not SPG policy is able to deal with ineffective teachers. This is a key point because one of the motives for creating student achievement-based teacher evaluation policies was *The Widget Effect* criticism – that evaluation systems had been unable to rid public schools of ineffective teachers (Weisberg et al., 2009). If District A's version of a student achievement-based teacher evaluation policy fails to do this, it could be strong evidence against retaining such policies in the future. Participants in this study had mixed opinions about the ability for the SPG to assist them in evaluating out ineffective teachers.

Several participants said they had used teachers' SPG results as one of several tools to either apply pressure on them to improve or to leave. Michael Cox stated, "Sometimes pressure is more expeditious. And, [SPG] can be utilized as a tool for that. It's not nice. It's not always pretty. But, it's best for kids." Eric Alexander said that reviewing one teacher's SPG results "helped her recognize her limitations as a teacher." Eric was confident he had enough evidence to evaluate this teacher out, but she resigned

before the process to remove her was completed. Deborah Walker was, at the time of her interview, evaluating two teachers that, in her words, "aren't very good with kids." She described the evaluation process with these two teachers this way:

> I'm using this process to make sure that their goal is tight to the instruction they are doing in the classroom. What I notice in their assessment pieces that are tied to their progress goal, the instruction that I'm seeing isn't allowing them to help the kids make progress. So, do I think [the SPG] drives instruction? It's a pretty damn good passenger if not the driver.

> Others said SPG was just one piece of the evidence needed to evaluate teachers

out. Patricia Johnson said, "it helps the conversation." Kathryn Campbell suggested that SPG, despite its 40% weight, is not enough, "To get an ineffective teacher out, you'd need a lot more than that." Trust in the SPG policy to come up with the same results as the other six standards, or the same results as the assistant principals' notions of their teachers' quality was clearly an issue with some participants.

> A few participants stated that SPG hurt their ability to remove ineffective

teachers. According to Jose Foster,

> I think the data can be manipulated. And, I don't think you can really manipulate as much planning and assessment, the day-to-day teaching. You know, when you go in the class you can't really fake it as much, as opposed to data that can be manipulated.

Mark Sanders raised the point that outside variables may be something a teacher could claim as an excuse for failing to achieve a satisfactory evaluation rating:

As far as evaluating teachers out? That, I think, is a little bit harder. I think

because there are so many holes you can poke in a student progress goal. And, if

at the end of the day their student progress goal isn't met, there's a lot more things

that they could, if they really want to, push to try to defend why it didn't get to

where it needed to be. But, if they didn't have lesson plans, it's harder to defend

that.

Mark's point here is especially important and highlights concerns raised in recent

literature (DeMitchell et al., 2012; Pullin, 2013). He hints at the possibility for a teacher

to be able to grieve his or her evaluation and have it overturned. If teachers are able to

demonstrate that 40% of their evaluation is invalid because their SPG is based on factors

they cannot control – poor attendance, students who chose not to follow through on

assignments, poor studying conditions at home – this could be detrimental to student

achievement-based policies. Teachers would certainly be able to draw on literature that

claims most variables governing student failure or success occurs outside the classroom

(Amrein-Beardsley, 2008; Berliner, 2014). How would District A's SPG policy hold up

in court? This is a question I raise for discussion in Chapter Five.

Policy Co-Construction

The second major finding of this study is that assistant principals' power and

authority, and other school goals, are significant factors in co-constructing SPG policy.

Datnow et al. (1998), in studies using co-construction theory, cite many external and

internal forces as having significant impact on how policy is implemented at the ground

level. According to co-construction theory, these forces are so significant that policy is

actually constructed as it is implemented. In this study, though, participants did not

acknowledge any other factors other than their own power and authority and how tightly

they have aligned SPGs to existing school goals. That is not to say that these forces do

not exist, only that assistant principals did not recognize them as important to policy

implementation. Instead, participants used their positional authority and worked

alongside teams, such as CTs and school improvement teams, to construct SPGs with

teachers, and sometimes for teachers.

Assistant principals' power and authority. A clear theme throughout all 16

interviews was how much authority each assistant principal exhibited over SPG policy

implementation. The weight of their authority was significant enough to allow for policy

co-construction, or how the policy was changed by its implementation. District A's

teacher evaluation policy states that "each teacher creates an evaluator-approved goal"

("Teacher Performance Evaluation Program Handbook," p. 10). Clearly, according to the

policy, the onus for goal creation is on the teacher. Assistant principals are under no

policy obligation to help teachers create the goal, or to set more specific parameters

outside the policy in creating them. However, most participants in this study went beyond

their policy obligations. Each of them defined their own parameters for creating rigorous

SPGs, many required some teachers to meet with them several times to review initial

goals, and many participants applied their own subjective judgment on final SPG ratings.

Other than the presence of school teams, which I discuss next, no participants

acknowledged any other power or authority as having more control over the SPG process

than themselves, not even their principals.

Approving SPGs. Participants exhibited significant oversight in approving or rejecting SPGs. In District A, this process if far from a rubber stamp activity. Every participant reported having to work with teachers to revise and resubmit SPGs. Anna Hill stated, "I have 16 people I'm evaluating, and I would say all 16 was a matter of my going back and forth and asking questions." Ruth Mercado reported that about 25% of her teachers had to rethink their goals. Participants noted there had been some improvement in teacher understanding of writing quality SPGs since the policy was instituted in 2012. However, assistant principals monitoring of goals was still a top priority for most participants.

Deborah Ballard's response to the question, "To what extent do you have power and authority to implement the SPG?" was somewhat different. She replied:

> Not much. I have the authority to sit down with teachers and help them develop their goal, and work with them in that respect. But, you know, the goal, bottom line is their choice. Right? By the policy, they can pick whatever they want. In the end, I can only guide them.

Yet, even Deborah required teachers to write goals a certain way. She refused to allow teachers to set goals that would not reach all students, "I won't let them write 80%. I make them write 100%. Because, my thing is, what about those other two kids?" Deborah's response is illustrative of how stated policy is trumped by policy co-construction. She perceived the stated policy gave full power to the teachers. Yet, she added new elements that effectively changed the policy.

Skills vs. content. Another element added to the policy by participants was the idea that SPGs ought to measure skills as opposed to content. It is not clear where this idea came from, but it was noted by almost all participants. Defining SPGs by skill, or prohibiting content goals, is not in the written policy. Because there has been no district-wide training on how to create quality SPGs, it seems this idea was spontaneously and independently developed at many schools. The idea is that skills are a more rigorous measure than content. In other words, teachers can easily meet SPGs if they assess a student's knowledge of European Geography in September, before any of their students have had any opportunity to memorize European capitals. Then, in February, teachers would apply the same assessment and expect a great deal of growth from all students. However, over the years assistant principals have questioned the value of these content-based assessments. Amanda Nelson described it best:

> They're sitting in your class all year. Like, I would hope they would know more
> government by the end of the year. That's just what you're doing. So, what else
> can we do to track [learning]? So, I wouldn't approve any content goals. It had to
> be a skill goal. And then, having it be something that's actually going to make an
> impact.

Similarly, Anthony Jenkins cites an example of students taking a pre-assessment early in the year when students are getting between zero and 15% before they have been introduced to any content. "You better hope that by the time you get to midyear everybody's up in the 50, 60, 70% range, which is showing a great amount of student progress. Do we really need to do that?" Requiring skill goals are related to assistant

principals attempt to increase goal rigor. Defining goal rigor has been a constant challenge for both teachers and administrators.

Goal rigor. The authority necessarily applied by assistant principals over goal creation is due in part to confusion over the definition of *rigor*. District A's *Teacher Performance Evaluation Program Handbook* defines a rigorous goal as having "an appropriate level of rigor to demonstrate mastery of learning/program objective" (p. 11). This vague language has created confusion and some frustration among assistant principals as they try to first understand what rigor means and then teach their teachers how to apply it to SPGs. Participants spent a significant amount of interview time expressing frustration over unrigorous goals.

Ronald Coleman defines the rigor problem. He stated, "Again, rigorous, if you ask five people what that means you're going to get five different answers." But, he and other participants would rather teachers err on the side of too much rigor. For example, Betsy White required a teacher to change her SPG so that the expectation for the teacher's special education students would be the same as her general education students. She did not accept the teacher's attitude that the goal was too hard for them. Ronald also described several examples of asking his teachers to "push it. I want you to push it." The lack of goal rigor was a more common problem with participants than too much rigor.

Subjective final ratings. Another way participants exhibited their power and authority was in considering final evaluation ratings for teachers. Ronald Coleman and Kathryn Campbell comments from the above section on validity (p. 99), are relevant here as well as examples of power and authority: Ronald Coleman stated, "I'm not going to let

a good teacher have a bad evaluation because of [the SPG]." Kathryn Campbell said, "I don't take [SPG] as a major indicator." Deborah Ballard and Ruth Mercado suggested they could manipulate SPG ratings by requiring other measures of progress. Deborah began this process at the mid-year review. If the SPG goal was not on track to pass, she would counsel her teachers to begin putting together a plan to supplement evidence of student progress. Even Michael Cox said if SPG data turned out to be invalid at the end of the year he would not throw out the goal, but "it would be an asterisk." Participants exercised their authority to not allow poor SPG results to effect teachers' final ratings if they knew from other sources or perceptions that teachers were effective, or better.

Through the SPG approval process, creating a skill-based requirement, establishing requirements for rigor, and exercising control over final ratings, assistant principals co-constructed SPG policy. They created parameters policy makers had not thought to include, thereby extending their authority beyond the written policy. Why they all did this remains unclear. Perhaps it is the vacuum of power that exists in the space between the policy and its implementation that assistant principals felt obligated to fill. Anthony Jenkins alluded to this, "We're leaving a lot of the evaluation process up to the person who is sitting in there doing the evaluation, which is the administrator doing 90% of it." In other words, if assistant principals are not going to take SPG seriously or implement it with fidelity, no one else is there to do it.

Another possibility for why participants went far beyond policy dimensions was they felt it was their moral and professional obligation to do so. Indeed, many participants perceived making SPG policy meaningful was best for kids. In explaining why she has

teachers revise SPGs until they meet her high rigorous standards Anna Hill stated, "We're looking at every kid making some improvement and hoping the minimum is going to be mastery." Also, several participants referred to their ethical obligations in carrying out the policy as Michael Cox did, "It's that kind of ethical even moral drive in you." Likewise, I asked Anthony Jenkins in a follow up question, "Are we depending on your integrity as an administrator to make sure it's a valuable experience for everybody?" He replied, "Absolutely. That goes back to my initial thought that the evaluation system is totally left up to the assistant principal." Though, how assistant principals make the process valuable is not always done in isolation. Most participants said they weave SPG creation into other instructional systems already in place at schools. School teams were the other most impactful factor in policy co-construction.

School teams. The other main element in co-constructing policy, other than assistant principal authority, is the existence of school teams. There are several teams at different levels in typical District A schools: CTs, or collaborative content teams, i.e. Algebra 1 or Chemistry teams; administrative CTs; and other school instructional teams, sometimes referred to as Instructional Councils, or school improvement plan teams. Depending on how tight schools are to their instructional mission through these teams, depends on how closely these goals are aligned to teacher SPGs. Each of these are discussed below, followed by more discussion on how school culture in general impacts SPG co-construction.

CTs. One of the most significant factors influencing SPG policy co-construction is how SPGs link to CT goals. As I described in Chapter One: Introduction, District A has

committed itself to PLC culture since the early 2000s, based on the DuFour (2004) model. This culture precedes the current teacher evaluation policy. There has been little discussion among administrators in how to reconcile PLC tenets with evaluation policy mandates. However, schools have independently adapted SPG policy to interconnect them with broader instructional goals. Participants in this study used a variety of approaches to reconcile evaluations and school goals. Most participants worked in schools that have decided to encourage teachers to use their CT goals in creating SPGs. Because the policy states SPGs are teacher created, this does not always happen at these schools. Only three participants say they do not require a link, and one participant discouraged teachers from using their CT goal as their SPG. The significance of team influence on SPG policy is dramatic when it comes to assessing individual teacher performance. SPGs could turn out to be more of a measure of how the team is doing as opposed to how the teacher is doing. Therefore, SPG policy is co-constructed by the absence or presence of the CT/SPG link because it determines the content of the goal, how it will be measured, and who it will measure.

The rationale for the decision to link CT goals to teacher SPGs varied among participants. In some cases, participants wanted to lessen the workload on teachers or make it easier for the assistant principal to manage. For example, Ruth Mercado stated, "they're going to have a team goal anyway, and they have to have their own SPG, and put it into [the district's online evaluation database], all they have to do is modify [the CT goal]." Ronald Coleman wished the district would mandate the link because he believed

CT goals were stronger. He expressed frustration with getting some teachers on board with CT culture.

> I have a teacher, she was like this is what I want to do with kids in my class, And, I was like, 'No, you've got to do what the team is doing.' And, the conversation turned with, you know, you need to participate more in the CT plans and the meetings.

For Ronald, CT fidelity was more important than SPG fidelity. His comment also demonstrates the power of a heavily weighted SPG to positively influence other school objectives. In the above example, Ronald used the 40% weight as incentive for a teacher to get with the CT culture. He did not use it as an incentive to improve that individual teacher's student achievement. Betsy White implied that the link ensured SPGs were good goals, "if they're on a team, they are kind of forced to be more realistic and come up with a decent goal." On the other hand, Mark Sanders highlighted how CT/SPG alignment skewed teacher performance ratings:

> Some of our weaker teachers end up with better goals, because the team has helped push them in that direction. Some of our stronger teachers probably settle for what the team goal is, when in many ways some of them, if left to their own devices would probably come up with something a little bit more unique or rigorous.

Mark's characterization suggests that teachers may limit risk-taking, or have opportunities to be creative stifled by the team approach. In these examples, participants have constructed an evaluation policy based on group accountability to learning, as

opposed to individual accountability. Despite the obvious positive trade-offs for working collaboratively in teams, framing SPGs this way runs counter to the purpose of teacher evaluation as an individual assessment.

Administrative CTs. In most District A high schools, administrative teams organize themselves into their own CTs. Administrative CTs usually include all assistant principals and two directors – Guidance Director and Activities Director. Usually, Administrative CTs do not include the building principal. Administrative CTs typically meet once a week, or every other week. Like other CTs, their goal is to consistently apply practices so that teachers and students receive the same level of service from each assistant principal or director. As it pertains to evaluations, these CT meetings are vital to SPG policy co-construction because ideas, procedures, and expectations are significantly powerful to redefine the policy.

Nearly all participants referred to their administrative team goals as an important driving influence of policy implementation. Mark Sanders took a lead role on his administrative CT by serving as his school's evaluation coordinator. Because his school has a large administrative team, the desire was to be on the same page with how they implemented the process, and a way to hand down school-wide goals as a "backdoor way to shape [SPGs]." For example, in order to facilitate cross-curricular learning, with Mark's leadership, all administrators required skills-based goals, as opposed to content goals, because skills transfer across curriculum better than content. Also, Mark's administrative CT agreed to commit to aligning SPGs with CT goals.

Patricia Johnson's administrative team constructed talking points in order to be consistent. This is mostly to ensure technical consistency such as meeting deadlines and holding each other accountable for conducting the same number of observations for all teachers. But, some talking points addressed expectations for creating SPGs, and consistency on rating teachers, i.e. what does it mean to be *effective* vs. *highly effective* on SPGs. Even though policy documents outline definitions for *effective* and *highly effective*, Patricia's team felt the need to extend and elaborate on these definitions, thereby altering the policy to fit local context.

Despite weekly administrative meetings, and supposed commitments to expectations, participants did not always know what happened during evaluation meetings between teacher and administrator behind closed doors. Amanda Nelson confided that it might not all be happening as planned.

> We try to stay very unified to a certain degree in terms of what our expectations are, and what we expect, essentially what we're approving and not approving. But, when I'm in my office and somebody else is in their office, I think it gets a little challenging. I know I play a much larger role in writing goals then some of the other administrators.

In any case, administrative CTs are at least attempting to hold themselves accountable for consistent application of SPG policy. Other teams impacted co-construction similarly.

School improvement teams. A final team that had influence on SPG co-construction was the existence of school improvement teams in most schools. School improvement, or instructional teams, are organized with the purpose of completing

annual school improvement plans required by the district. They might also be convened to apply school-wide instructional goals the principal has identified as vital to the mission of the school. In either case, in effective schools, the mission of the school trickles down to the classroom level, often through SPGs.

Deborah Ballard reported that her school, HS 4, was committed to two school-wide goals: raising the bar, and closing the gap. This refers to raising achievement for all students on all assessments, and closing the minority achievement gap, respectively. Therefore, all CT goals, and subsequently, all SPGs, had to be aligned to these two school goals. The engine for executing HS 4's school-wide goals was its *Leadership Council*, a group of teachers, department chairs, and administrators who researched and implemented strategies to help meet the goals. This year, HS 4 engaged *Executive Functioning* strategies to address their "close-the-gap" goal, and *Project Based Learning* to "raise the bar." By design, all SPGs had to be tied to these goals in some way.

At Mark Sander's school, HS 3, their school team developed vertical, skills-based goals so that teachers could expect students coming out of 9th grade English, for example, to have certain writing and reading comprehension abilities. Each CT goal was committed to aligning to the idea of vertical goals specific to skills required in each content area, i.e. computation skills in math or research skills in Social Studies. And, because HS 3 also required CT/SPG alignment, every SPG was ultimately aligned with these vertical goals. Mark saw SPG policy as critical to making this happen. "You know, we wouldn't have done that as effectively and with as much weight, if it wasn't tied to the evaluation." In other words, HS 3 was more committed to this process because

116

teachers were held accountable through the evaluation process. Yet, they were using SPG policy to further objectives other than assessing teacher quality or holding teachers accountable for student achievement.

Several other participants mentioned the importance of school goals, although not as systematically as HS 4 or HS 3. Anthony Jenkins and Ruth Mercado, both at HS 18, perceived their school's commitment to tying SPGs to state EOC tests as running counter to their educational philosophies. Both considered it important to raise scores on these all-important tests, but desired higher standards based on critical thinking and problem solving. Michael Cox discussed the school improvement team's commitment to Project Based Learning (PBL) as "the key to getting things done instructionally." While, administrators did not require SPGs to link to it, he saw PBL as something they should aspire to do in the future. Therefore, depending on the level of school commitment, school team goals can indirectly or directly impact SPG policy. This, in combination with assistant principal authority to dictate their alignment with SPG, or not, created co-construction scenarios in schools.

School culture. School culture or climate, in the most general sense, has significant influence on everything that happens in a public high school. A detailed examination of school culture is beyond the scope of this study because I only interviewed assistant principals in the context of teacher evaluation policy. However, a few participants alluded to how their schools were so tightly aligned to a unified instructional mission, that it was clear the mission was impacting how they were co-constructing SPG policy. As a result of their commitment to instruction with all

stakeholders, these schools were less interested in individual teacher evaluation and more interested in a unified accountability for all student learning. The above examples of aligning SPGs with CT goals are a subset of this evidence. These examples are more global.

Deborah Ballard considered her school to be "a real tight school" that comes from a strong community atmosphere. School culture "is the biggest influence here at HS 4, it being a neighborhood school." Unlike most schools in District A, the majority of HS 4's boundaries are included a single town. The result inside the school has been a unified school culture more interested in carrying out school goals than focusing on individual SPGs. Therefore, their SPGs are very tightly aligned to school goals. HS 4's goals were cited above as raising the bar and closing the gap.

Heather Rogers spoke about how HS 7 was an instructionally focused school. This meant that SPGs, because they were aligned to instruction, took on more of a priority for teachers and administrators than say attendance or discipline. "This is a school culture, in my opinion, where kids stay after, they want to seek the help of their teacher." She was not sure where that culture came from, but was sure it was not the school's high SES. "Do I think socioeconomic status of this school's community has anything to do with me as an assistant principal discussing and working with the teacher? I don't think so . . . I just think attendance is high, the respect of the teacher is more positive here." If school administrators focus more on attendance than instruction, they might have to construct SPG policy differently, and without as much time.

Regardless of where school culture comes from, in some cases it clearly impacted policy implementation. School culture is too broad a subject to analyze in this study. The point here is that SPG policy was guided by school culture in these situations, not the other way around. Similar to other factors impacting policy co-construction, school culture demonstrates the weakness of SPG policy. Where some education policies have impact beyond their design to have positive or negative effects on teaching and learning, SPG is not strong enough. And, as demonstrated, SPG is often used to carry out other school objectives or affirm existing school culture.

Variables not impacting co-construction. Co-construction theory identifies many factors that might control policy implementation. In this study, participants clearly saw themselves as the most important and powerful actors in policy co-construction, with no oversight from their superiors in how they executed SPG policy. This and how tightly their schools were committed to other school goals, dominated their responses regarding policy implementation. Other variables identified by previous co-construction studies – personalities of actors, such as teachers; influence of other power actors such as principals, school board members, or district supervisors; environmental conditions, such as school SES or school resources; conflicting or concurrent policies; or "micropolitical interactions" (Datnow, 2000, p. 359) – were not identified by participants as important variables. Even when asked directly, "What outside forces or factors limit, enhance, or influence the way you implement SPG policy?" participants had trouble coming up with any response at all. Two variables this study anticipated might have profound influence on policy co-construction were years of administrator experience and school SES. As

119

explained above, however, these factors were not perceived by participants to have any influence.

When asked specifically about their principals' influence, no participants cited them as an important variable. Participants reported the principal's role was to set broad expectations for applying evaluation policy uniformly and fairly. Mark Sanders referred to this as his principal's "umbrella," and that the administrative team took cues from his broad expectations. Others reported their principals had no knowledge of how they dealt with teacher evaluations. Jose Foster stated, "I absolutely don't think he has any idea what's going on with the teacher evaluation system." Others see their principals as resources, like Anna Hill, "I've had multiple conversations, and there have certainly been times when I've said, 'What are your thoughts on this goal?' But, I have to say, at this point, I feel pretty comfortable in knowing what his expectations are."

My own personal experience as a principal in District A is similar to this characterization. I have little knowledge about what happens between assistant principals and teachers with regard to SPGs unless there is a problem. I trust my administrators to work in their administrative CT to be consistent and apply the evaluation process uniformly and fairly. But, at my school, instructional best practices and teacher quality are addressed in other ways – through the school improvement process, the Instructional Council, and regular, teacher-led professional development. District A has 24 different principals with 24 very different personalities. And yet, few seem to be involved directly in SPG policy implementation.

The potential does exist for schools to be influenced by state or district control. District A currently has no schools under state accreditation warnings. However, when Patricia Johnson's school, HS 10, was under a warning a few years ago, the region required SPGs, through the principal, to focus on raising math scores in order to meet state requirements. However, as soon as they were off warning, teachers went back to creating their own SPGs. Patricia doubted that those SPGs created under state influence had anything to do with their success in raising scores. She said, "Our success came from focusing on skills rather than content." Therefore, even when under the cloud of being labeled a failing school, outside influences, though they might seize control of SPG policy temporarily, this evidence suggests their impact is minimal.

Other environmental factors did not weigh heavily into the way participants applied SPG policy either. Participants, as explained above, dismissed school SES as an important factor. Time and workload constraints were noted as problems in implementing other aspects of the evaluation policy, like getting in all the observations done in time, but time was not a factor in developing and approving SPGs. Some participants cited difficulty in dealing with certain teacher personalities. Betsy White said some "curmudgeon" teachers needed more attention and cajoling. Anna Hill noted that this year in particular she saw a large increase in teachers who needed to revise their goals than in previous years. However, other factors checked these specific forces so that policy co-construction was not impacted. For example, Betsy White dealt with her "curmudgeons" with her interpersonal skill and utilizing team structures to bring them in

line. Anna Hill simply dedicated more time to working with her teachers, though she did not reduce expectations or change the policy.

It was evident from participant responses that the main drivers of District A's SPG policy co-construction were autonomous power and authority of assistant principals and the presence of school teams. Anticipated variables of years of administrator experience and school SES, the school principal, as well as a host of other factors suggested by co-construction theory, were non-factors. Considering the wealth of research on what impacts schools, especially with school SES, these may be surprising findings. However, when one considers the lack of research on assistant principals, and the extent to which scholars have overlooked these important power players, perhaps it should not be a surprise that they wield such control over policy co-construction. Additionally, the impact of school teams like PLC structures, has also been understudied. Having addressed how District A's SPG policy has been shaped by these factors, I now turn to an examination of the policy's impact on instructional leadership.

SPG Impact on Instructional Leadership

The third major finding of this study was the SPG policy's positive impact on some participants' instructional leadership capacity. SPG created opportunities for assistant principals to have conversations with teachers about individual students, to target subgroups of students for improved achievement, and to have meaningful dialogue around data. All participants did not consider SPG beneficial to increasing their instructional capacity. But, they were not distracted by SPG policy either. Therefore, the

general perception was that SPG was a positive influence on instructional leadership at best; and at worst, neutral.

That SPG policy had some positive impact on participants' instructional leadership capacity was surprising and one that appears unintended by policy makers. At the outset of this study, I had hypothesized that student achievement-based teacher evaluation policies were potentially problematic because they appeared to run counter to instructional leadership trends. I maintain, though, that the intended purpose of these accountability policies still runs counter to the purpose of instructional leadership because one is punitive and the other pro-active. Also, the results described elsewhere in this study demonstrate that the intentions of SPG are not being carried out. Rather, it seems the positive impact on instructional leadership is an unintended positive consequence of District A's SPG policy. So, despite policy maker intentions to hold teachers accountable for student achievement, the way it has been co-constructed by assistant principals provides them opportunities to increase their capacity for instructional leadership.

By name and by need. Participants in this study reported increased instructional leadership capacity because the SPG policy created opportunities to focus on individual students and their needs. District A's *Teacher Performance Evaluation Program Handbook* states, "The goal is set for only one defined group of students (elementary – one curricular area; middle and high school – one course section)" (p. 11). It is not clear why policy makers chose to craft SPGs this way, nor is it clear whether this is an option or mandatory. In any case, it has become customary for teachers in District A to select a

single class period, or a subset of students from within a class, i.e. special education students or ESOL students, as targets for their SPGs. This condition created opportunities for assistant principals to look at student specific data with their teachers and to have conversations about why these individual were succeeding or not.

During interviews, participants were able to share examples of teacher conservations about individual students. Deborah Walker called this knowing students "by name and by need." In going over assessments used for teachers' SPGs, Deborah talked to teachers about underperforming students in the data, "How does Matt look as compared to the rest of the class? How does Matt look as compared to the rest of the school?" Then, she would discuss instructional strategies with teachers on the best ways to reach that student. These conversations involved what Deborah referred to as "staying tight to the instructional cycle." The instructional cycle is a PLC model that describes data collection, data analysis, and responding to data by addressing where students are not showing achievement. According to Deborah, SPGs should not only select data to analyze, it should also provide a plan to address individual student deficiencies named in the data. Her conversations with teachers and within their CTs centered around these individual, struggling students.

Similarly, Amanda Nelson used SPG goal setting and resulting data collection to have individual student conversations with teachers. She shared a typical conversation she might have with a teacher at a mid-year SPG review.

When I'm looking at the data [with a teacher], I'll be like, "Okay, talk to me about Yolanda. I see she's getting worse. What's happening with this kid?" And,

really making the teachers speak to the individual students. Which I think helps

them remember that they have to pay attention to that.

Other participants spoke about other instructional leadership tools, such as professional development. However, both Deborah and Amanda defined their entire instructional leadership persona through data conversations and being able to drill down to the individual, student level.

Of the 16 participants, Deborah and Amanda were the only ones to discuss individual students so specifically. However, that the others did not should not diminish the importance of these conversations. SPG policy creates the opportunity, the potential for all assistant principals to have these conversations. This is notable because one of the criticisms several participants had about SPG policy is that the goals are not designed to measure all of a teachers' students all of the time. Yet, no matter what students SPGs target, the resulting conversations about these students with assistant principals, assuming the conversations lead to better instructional skill, certainly improves instruction for all students of that teacher.

Targeted subgroups. Several participants noted opportunities SPG policy created to target and support ESOL and special education teachers and their students. ESOL and special education teachers under evaluation often worry that SPGs might disadvantage them because these learners face more challenges than general education students. Deborah Ballard, who oversees special education exclusively, says, "I think special education teachers have it much tougher than everybody else these days." As a result of

these challenges, teachers of ESOL and special education populations tend to create

unrigorous goals. Some assistant principals did not accept this lowering of expectations.

Betsy White shared an example of a teacher, who tried to get by with an

unrigorous goal for her special education students,

> I had a teacher, she picked something that was very basic, and she thought it was
>
> okay for her Special Ed kids. And, she said that, you know, it's hard for them.
>
> And, I said, "Well, yes. But, don't you want to push them?"

During his interview, Ronald Coleman shared an Excel spreadsheet he shows his teachers

during goal setting conferences demonstrating they could still get effective ratings even if

they did not reach their SPG targets. So, his constant refrain to teachers was "push it."

On the other hand, setting uniform expectations for all students can be risky for

teachers. Timothy Hughes shared one conversation with a special education teacher, "I've

got some Sped English teachers that tell me, [Tim] no matter what I do this kid is just not

going to pass the SOL. So, then we start game planning what's it going look like for you."

Timothy's example demonstrates that setting expectations is not necessarily the cure-all.

There has to be follow through and support for teachers to help them meet the challenges

their students face.

As indicated above, teachers are wary of taking risks with rigorous SPGs because

it counts as 40% of their final rating. Mark Sanders, though, described a unique situation

where an Algebra 1 teacher targeted his SPG on his most challenging learners that

included ESOL students. "Honorable for him to do that when he could have picked his

AP [Calculus class]. He chose his hardest group of kids to have as his goal and worked

with them. I mean success was there." Even though the teacher did not meet his goal, "it motivated him to continue to work with those kids." This is an example of how product is much less important than process. Though students did not reach intended achievement targets, according to Mark, it was not due to a lack of attention or a lack of effort on the teacher's part, which is often the criticism when students fail. This example highlights the benefits and risks of setting high expectations for all students. Yet, when the risk is tied to 40% of a teacher's evaluation, one wonders if the policy limits risk-taking.

SPG also helped identify opportunities for general education teachers to interact with ESOL or special education teachers. Patricia Johnson shared an example of sharing strategies when a general education Science teacher was struggling to reach his SPG goal. In her estimation, reaching the goal required differentiating instruction for his struggling general education learners, a skill this teacher apparently lacked. Patricia's ESOL teachers were already doing that. "I was excited. Oh my gosh! My ESOL teachers are doing that every day with these kids." She took the opportunity to set up peer observations for the Science teacher in Level 1 ESOL classrooms. This led to a school-wide peer observation program.

Targeting subgroups for improvement can be a double-edged sword. If teachers take risks, or even maintain the same expectations for ESOL and special education students as general education students, they risk 40% of their teacher evaluation score. However, if expectations are lowered at the outset, or if teachers do not take risks, these populations are disadvantaged before they even start the school year. Therefore, the way

assistant principals work with SPG policy, not just the policy itself, is key to improving achievement with these targeted subgroups.

Data conversations. Data conversations helped assistant principals and teachers to think deeply about how to raise student achievement using various forms of student achievement data associated with SPGs. Despite the perception of most participants that SPG results were invalid and unreliable, they also reported that the process of designing, revising, and assessing SPGs was effective in improving instructional delivery or instructional strategies. Also, even if SPGs only measured small groups of students, or even if teachers failed to meet their SPGs, conversations were key components of instructional leadership because they promoted creative dialogue around improving instructional strategies. Improving instruction did not necessarily demonstrate a direct link to improved student achievement as the SPG policy intended. However, improving instruction was a goal of SPG as well, and a way for assistant principals to exhibit instructional leadership.

These conversations also created a sense that teachers were not alone in trying to meet their SPGs. It was a shared accountability. After initial goal setting issues regarding rigor were dealt with, and these were few and far between, conversations around SPG progress were generally supportive and encouraging. Through these conversations, participants signaled to teachers that the responsibility for meeting these targets was shared between teacher, student, and administrators. This is ironic because the stated purpose of SPG policy is to hold teachers directly accountable for student achievement.

Yet, the system created by assistant principals took at least some of the onus off the teacher.

Many participants reported that SPG data dialogues made conversations with teachers instructionally meaningful. Eric Alexander said that without SPG, "conversations would not have the meaning behind it. Because, the student progress goal is quantifiable." Then, he shared an example of trying to convince an English teacher to quantify writing goals using rubrics, something that team had not done before. No matter how that discussion turned out, the benefit to teachers was an opportunity to reflect on new and perhaps better instructional practices. Ruth Mercado had expressed her dislike for SPG throughout most of the interview. Yet, even she saw the value of SPG in creating dialogue, "At least [SPG] provides us the opportunity to focus on data and say, 'Well how are your kids doing in their end of course test, or the AP exam?' So, it does lend to the conversation, I think."

Other participants felt like SPG data dialogues gave them more capacity for instructional leadership. For example, Ronald Coleman stated, "I think just engaging teachers in those conversations about, 'Is there something you're willing to try differently?' I think makes an effective instructional leader." Even if participants were not certain about the policy's impact on their instructional leadership – some participants considered themselves excellent instructional leaders prior to SPG policy – some still mentioned data conversations. Eric Alexander stated, "I'm not sure it makes me a better instructional leader, or a more effective instructional leader, other than we could have more specific conversations." Similarly, Kathryn Campbell stated, "I wouldn't say the

129

[SPG] concept itself has inhibited anything. I think it's probably been good conversations for me to have with teachers." Therefore, even if participants felt SPG policy was not having an impact, conversations were important to giving them a sense of being part of the instructional process.

On the other hand, the presence of SPG data did not produce instructional leadership or meaningful discussion in isolation. Much of the value depended on how committed assistant principals were to creating positive relationships and engaging in dialogue. Amanda Nelson commented on this point,

I think, if you're just there, checking the box, if you're an administrator that's like, 'Oh, great. That's wonderful.' And, never talk to [teachers] about it again, until midyear and the final, I think [teachers] don't see the value in it.

Michael Cox's belief was that strong instructional leaders listen as much as they talk, if not more. "[SPG] and its impact on the other standards has helped fuel those conversations and the conversations are better if I listen actively."

These examples suggest that some quantity of instructional leadership capacity or interpersonal skill might already need to be in place in order for SPG dialogues to have positive impacts. Yet, even those participants who questioned their instructional leadership capacity were boosted by SPG. Heather Rogers doubted her ability to be a strong instructional leader, "I don't know how great of an instructional leader I may be." Still, when I asked her if she perceived herself more as an instructional leader under this new policy than the previous one without SPG, she replied, "I would say, yes. Honestly. Because, with this, now I feel all of us are working at this."

Additionally, the sense that teachers and administrators were working together stems from these data conversations. Many participants noted the stress teachers had over evaluations, particularly that the SPG is weighted at 40%. However, when participants engaged teachers in discussions about their SPGs and centered those discussion on data, it took personality and blame out of it because data are objective. Also, participants reported that teacher stress has been alleviated with time and an understanding that they are not alone in achieving SPGs. Amanda Nelson stated, "I think with the confidence of teachers, and the confidence of administrators, to really feel like, okay there's meaning here, we can actually achieve this." Her use of "we" is an indication of her shared accountability with the teacher. Similarly, Anthony Jenkins said, "Now all of a sudden, they have a little bit of guidance, 'Oh yeah, I can understand how I can monitor this over the course of the year.' I think it takes a lot of stress off of them."

The way participants worked with SPGs was not simply a stress reducing exercise. It also extended to a genuine philosophy of shared accountability. Michael Cox described how SPG policy held him accountable as much as it held teachers accountable:

> If I can't substantiate my decision on how a teacher did on his [SPG] then I'm not doing well. I should be able to talk to them about why we're on the same page. THE page. The [SPG] holds me accountable.

Michael referred to the SPG as creating a "reciprocal accountability" between him and his teachers. The tone here is indicative of most other participant philosophies as well. None of them were out to get teachers. In fact, Ruth Mercado, Ronald Coleman, and

Patricia Johnson all spoke the same phrase about this not being an 'I gotcha' process. Support and encouragement for teachers was clear from each participant.

Shared accountability, relieving stress, and data conversations with teachers have been in place in District A before SPG policy was implemented. Of course, the quality and frequency of conversations varies with how good the administrator is at his or her job. The difference SPG policy created was that, similar to the above examples of individual students, these conversations were data driven. The data provided an objectivity that took the onus off the teacher to a small degree. SPG policy still held them accountable for achieving SPG targets, but the sense participants provided is that they felt just as responsible for student achievement as the teacher. The policy did not prevent assistant principals from holding teachers directly and solely accountable without encouragement, support, or guidance. But, no policy is foolproof.

The irony here is that student achievement-based teacher evaluation policies were established nationwide as a way to hold teachers directly and solely accountable for student achievement. Even in District A, the SPG policy states clearly that holding teachers accountable is one of its goals. And yet, the policy in action, as it is co-constructed by assistant principals, actually serves to alleviate the burden on teachers.

Conclusion

This study examined assistant principals' perceptions of District A's SPG policy and how these perceptions impacted their instructional leadership capacity. After detailed analysis of 16, semi-structured interviews, I found that participants did not perceive District A's SPG essential to assessing teacher quality or to holding teachers accountable

for student achievement. I also found that SPG policy was co-constructed based on assistant principals' power and authority as well as the existence of goals created by school teams, such as CT or school improvement plan teams. Finally, I found that SPGs had positive impact on participants' capacity for instructional leadership.

These three major findings are significant because they contribute empirical evidence that student achievement-based teacher evaluation policies do not always work to hold teachers accountable and are not necessarily used to assess their quality accurately. The results also contribute valuable data as to what factors are important and unimportant in the co-construction of education policy. Further, this study presents surprising results about the unintended, positive consequences of how focusing student-specific data in the context of SPGs has increased benefit to instructional leadership capacity. These benefits directly relate to individual struggling students and students from ESOL and special education.

The 16 participants selected for interviews were from various demographic categories and were generally very engaged and passionate about this topic. The high interest among assistant principals, 35 out of 83 District A assistant principals replied to my invitation within a week, is indicative that teacher evaluation policy is of great interest to those who are primarily responsible for implementing it. District A will likely find these results beneficial and important in targeting professional development to school administrators and teachers as they work to create effective SPGs. Policy makers might also benefit from data that shows how SPG policy works in action, whether the outcomes are intended or unintended, as they work to craft future teacher evaluation

policies or revise current ones. Finally, the empirical evidence presented here contributes

to the scholarly debate on whether or not holding teachers directly accountable for

student achievement is wise or even possible. In the next and final chapter, I summarize

the study and its results, present topics for discussion, review the study's limitations, and

suggest avenues for future study.

Chapter Five: Conclusions

A study of high school assistant principals' perceptions of their teacher evaluation policy contributes to several discussions already present in the literature as well as in education and policy circles. First, identifying elements of teacher quality and inventing ways to hold teachers accountable for student achievement remains elusive. Second, factors impacting education policy co-construction depend on powerful actors and pre-existing school structures. Third, assistant principal roles continues to tend strongly toward instructional leadership. This study contributes empirical evidence to each of these ongoing discussions by analyzing high school assistant principal perceptions of a large, suburban public school system's student achievement-based teacher evaluation policy.

In this chapter I summarize my findings, discuss how results contribute to scholarly discourse, and demonstrate the significance of these findings. I also discuss limitations and suggest avenues for future study. In the end, I believe my conclusions provide stakeholders with ideas and direction on the way we perceive the purpose of teaching and our expectations of evaluation policies; ideas about assistant principals as instructional leaders; and ideas on how we attempt to hold teachers accountable for student achievement.

Summary of Findings

This study brought to light three important findings. Interviews of 16 high school assistant principals from a large, suburban public school district revealed that most participants are not using the policy to assess teacher quality or to hold teachers accountable for student achievement. It also found that policy implementation is influenced by assistant principal authority and school teams. Finally, the study found that there are positive impacts on assistant principals' instructional leadership capacity. These findings make significant contributions to existing scholarship and inform educational and policy practices about teaching and learning.

I began this research to examine the possible problem that student achievement-based teacher evaluation policies seemed to run counter to a trend toward assistant principals as instructional leaders. The study I designed to address this problem employed a qualitative, semi-structured interview guide to assess assistant principals' perceptions of an SPG teacher evaluation policy first implemented in 2012. The study was based on my goals to contribute empirical evidence to teacher evaluation, instructional leadership, and assistant principal literature; provide district leaders with evidence to inform evaluation and instructional leadership practices; and to provide an opportunity for assistant principals to reflect on practices and share ideas. My research questions were as follows:

1. What are high school assistant principals' perceptions of SPGs?
2. How do assistant principals' perceptions contribute to SPG policy co-construction?

3. How do these perceptions impact assistant principal capacity for instructional

   leadership?

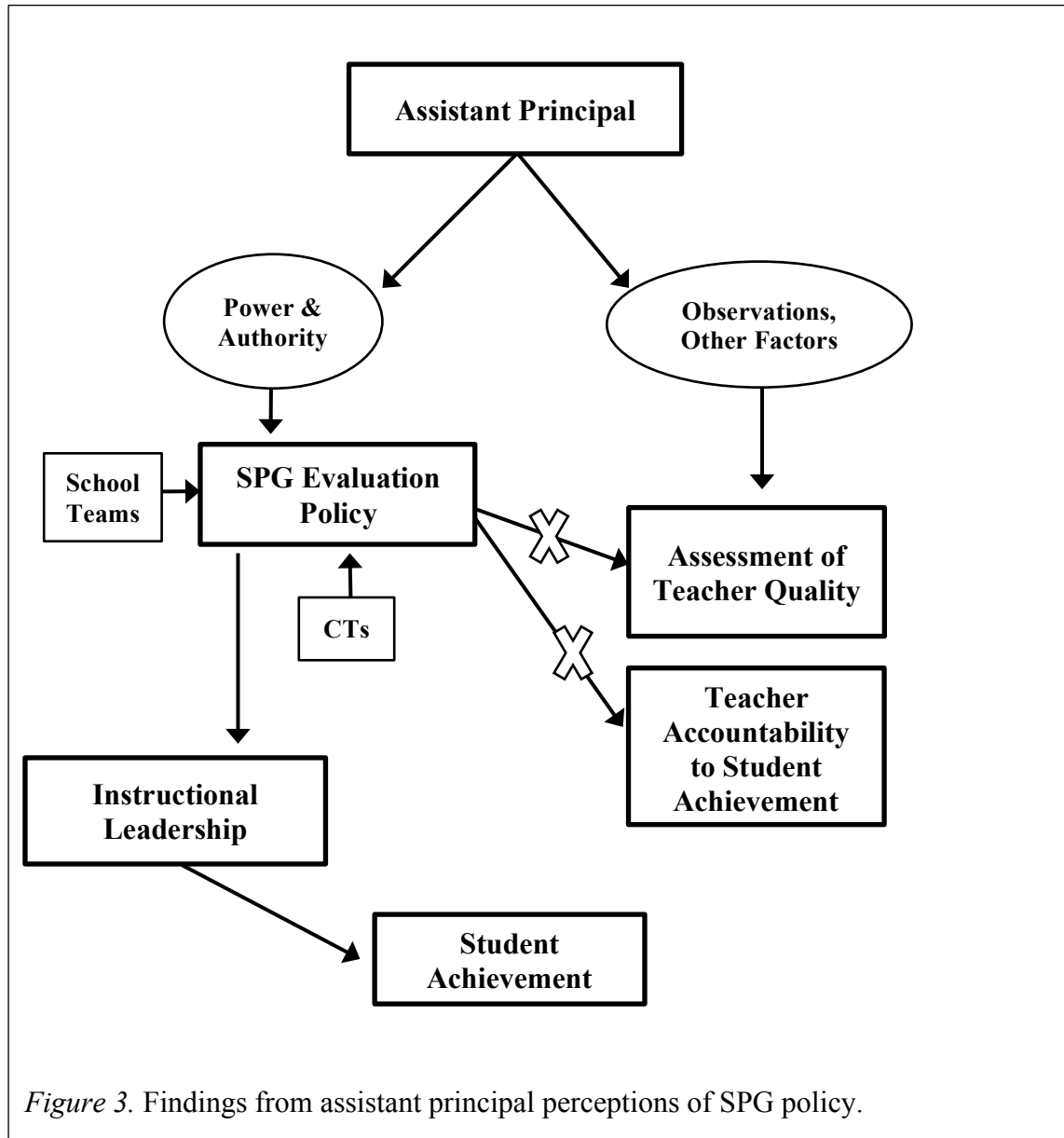*Figure 3.* Findings from assistant principal perceptions of SPG policy.

Figure 3 illustrates the findings of this study. Assistant principals exhibited a great deal of control over SPG policy. Yet, they did not use SPG as it was designed. Instead, they had impact on teacher quality and teacher accountability using other aspects of the teacher evaluation policy, observations, and their own previous notions of teacher quality. Participants did use their instructional leadership skills to impact student achievement. Sometimes their instructional leadership practices included SPG policy. Finally, Figure 3 shows that CTs and other school teams had an impact on how participants co-constructed SPG policy.

Previous research on teacher evaluation, instructional leadership, and assistant principal perceptions, looked at these phenomena independently. This study examined how policy and practice around these topics interacted. By way of a carefully designed deep examination of assistant principal perceptions, this study was able to produce rich evidence showing how and why evaluation policy was interpreted and practiced in a local setting.

This study does not account for all the reasons why assistant principals behaved this way toward District A's SPG policy. However, the overwhelming feeling from participants was that teachers need more guidance than the policy, by itself, allows. This comes from a teaching, as opposed to an evaluating, perspective. After all, participants were all teachers for at least 5 years. The average years of teaching experience for all participants was 11.8 (see Appendix C). Therefore, their goal seems to have been to improve teaching as they would have improved student learning, through guidance and support, as opposed to applying a single rating with no explanation or dialogue.

Discussion

The results of this study raise several questions about teaching and learning in the accountability era. In this section, I synthesize my findings with existing scholarship in order to promote discussion that will lead to further clarity on teacher evaluation, instructional leadership, assistant principals' roles, and co-construction theory. This discussion should provide a better understanding of what I believe is happening at the ground level of teacher evaluation policy implementation and with teaching and learning during the accountability era.

Teacher evaluation policy. Teacher evaluation scholarship has been bogged down by debate over the use of VAM and over a debate on the purpose of evaluations. As noted in Chapter Two: Literature Review, scholars have polarized the debate along fundamental lines. Economists are pitted against social scientists in what has occasionally devolved into personal attacks. This study offers a way out of that debate by contributing empirical evidence on what is actually happening with teacher evaluation policy in schools between administrators and teachers. In this study, participant practice with teacher evaluation policy did not match the stated purpose of the policy. Therefore, scholars would be well served by involving ground-level actors in the debate on policy purpose because, quite clearly from this study, policy practice is paramount to impacting teaching and learning.

Teacher evaluation policy purpose. My finding, that participants did not use SPG to assess teacher quality or to hold teachers accountable for student achievement, raises new questions about the stated purpose of student achievement-based teacher evaluation

policy and how these policies are actually being used. In other words, if the stated

purpose is for assistant principals to use SPG policy as a summative tool, but they use it

as a formative tool ignoring its summative features, then the debate on whether these

policies are summative or formative is moot.

This study suggests a change in the direction of scholarship on teacher evaluation

policy from a polarized debate, to one that might now focus on policy in practice and

how policy meshes with instructional leadership. The current literature pits those who

believe evaluations should be solely summative in purpose (Milanowski, 2006) against

those who believe evaluation should be formative (Darling-Hammond, 1986).

Milanowski (2006) goes a step further to suggest that formative practices to improve

teaching should be done with professional development, not teacher evaluation. In this

study, participants perceived the purpose of SPG policy as formative. They took

opportunities to examine data, even when they did not trust SPGs as valid, to have data

discussion with teachers and discuss teacher strategies that, they say, improved teaching.

Yet, the stated purposes of SPG policy are summative measures of teacher quality and

accountability to student achievement. Therefore, this study reveals an incongruence

between policy purpose and policy practice. It also reveals that assistant principals

believe improving existing teachers' quality is a priority for them.

Why this incongruence exists seems to originate from who assistant principals are

and what they believe is the best way to improve teaching and learning. Assistant

principals are teachers at their core. Many participants referenced their own teaching

experiences during interviews. Anthony Jenkins said his approach to evaluating teachers

was based on how he remembers being evaluated himself. Also, many participants noted how important it was to reduce teacher stress about the evaluation process. Assistant principals are therefore more concerned with teachers and teaching than evaluating because that role is congruent with who they are.

The incongruence between policy purpose and practice also appears to stem from what assistant principals think is an efficient or ethical use of their time. Clearly, participants would rather work with teachers on their SPGs than use it to evaluate them out. The reasons for this varied. Some said the system was not effective in evaluating teachers out. Others said it was the ethical drive within them that directed them to use the system to help teachers improve. So, regardless of assistant principals' perceptions of SPG policy, it was clear that policy purpose did not align with policy practice.

Teacher quality and accountability. If teachers and evaluators in District A, after five years of policy implementation, and having all manner of freedom to find ways to link teacher behavior to student achievement, could not find valid measures to make a link between teaching behaviors and student achievement, it seems unlikely that valid measures can ever exist. Despite the perception among assistant principals that teachers should be held accountable for student learning, they do not use SPG to do this.

Here again, there seems to be an incongruence between what assistant principals say they believe about accountability and how they practice accountability measures. This incongruence suggests that assistant principals may not actually believe teachers can or should be held accountable for student achievement. When probed more deeply on the accountability question, most participants qualified their beliefs stating that everyone –

parents, administrators, resource personnel, and the students themselves – is accountable for student achievement. So, how do we fairly assess teachers for shared accountability? That most participants objected to a 40% SPG weight seems to confirm this contradiction. Why would they object to 40% if they firmly believed teachers should be held directly accountable for student achievement?

Clearly, the focus of the participants in this study was to measure teacher effectiveness using the other six standards in the evaluation policy. These standards, in their collective opinion, assess teacher quality more accurately than SPGs even if they do not directly link teachers to student achievement. Timothy Hughes stated, "I believe that the data and all the numbers are just going to be a byproduct of those kids that are engaged." Therefore, one byproduct of good teaching should be increased student achievement.

Assistant principals are educators. Participants in this study have been evaluators of teachers, or evaluated as teachers, for most of their adult lives. Their answers indicated that they know what it is to be a good teacher, or at least they believed they knew. Therefore, they did not need a measure like SPG to define good teaching for them. They did not believe they needed to hold teachers accountable for student learning because good teaching would do that automatically. Add to that belief a deep distrust of SPG validity and it makes perfect sense that assistant principals did not use SPG to hold teachers accountable.

Timothy Hughes's above characterization of how teachers should be assessed, through student engagement, is just one example of the variety of ways participants

142

measured teacher quality. This independent analysis provides a way out of the current teacher evaluation debate because it suggests educators are skilled enough to hold teachers indirectly accountable for student achievement using a variety of observable means. This excuses teachers from being held accountable for variables beyond their control. For example, Berliner (2014) estimates that teachers are responsible for approximately 20% of the variables governing student success or failure in teachers' classes. Hanushek (2011), on the other hand, insists that teachers are the most influential factor on student achievement and measures can and should be devised to hold them directly accountable for it. However, in practice, if assistant principals can control teacher quality through the six standards, i.e. instructional delivery, instructional planning, etc., and they trust these factors to lead to student achievement, albeit indirectly, this should satisfy the concerns of both economists and social scientists.

Unexplored variables impacting assistant principal perceptions. On the other hand, this study also suggests that all variables impacting assistant principal perceptions have not yet been completely explored. Michael Cox's approach with SPG was an outlier that deserves some attention because the success with which he employed SPG to impact learning presents great potential for other administrators. Michael's philosophy was to use SPG as best he could to assess teacher quality and to hold teachers accountable. Why was Michael's approach so different from the other participants? It is not entirely clear. Further study, which I propose a section below, is needed to ascertain exactly why. However, there is enough evidence to offer some discussion points.

First, perhaps there is something about Michael's school, community, or his faculty that necessitates a more direct link between teachers and student achievement. If, for example, his school has the reputation that teachers are not serious or they do not care, assistant principals may be more likely to take a harder line on holding them accountable for clear learning results. His school does not have this reputation as far as I know. However, his school is in the low SES category and many schools in that category, across the nation, have poor academic reputations. If assistant principals work in environments where they are expected to clean up instruction or take hard lines on holding teachers accountable for showing up on time, attending scheduled meetings, taking attendance, as well as for learning – then perhaps their approach might be similar to Michaels.

The other possibility is that there is something about Michael's background that leads him to have a more pro-SPG perception. Perhaps Michael's high expectations for his teachers to get good academic results comes from his own personal success as a teacher or department chairperson. Or, perhaps it was his experience teaching in a different school district. If Michael found it easy as a teacher to get his students to learn, it would stand to reason that he might believe this was not much to ask of the teachers he supervises. Or, teaching in a system with a VAM policy, like Michael did for several years, may have impacted his perception of a student achievement-based evaluation policy that is far less strict than VAM. These ideas are speculative until we can learn more about assistant principals like Michael. What does seem to be strongly evident is

that participants' philosophies on teaching and learning had an impact on how they implemented SPG policy.

Findings of this study extends scholarship on teacher accountability by examining evaluator perceptions on how they balance their educational philosophies of teaching and learning, with holding teachers accountable. Derrington writes, "principals struggled to balance their belief that all teachers can grow and improve with a system of accountability that rates and ranks teachers" (p. 128). This struggle was evident in participant responses. Most ignored SPG as a tool to hold teachers accountable because they saw it as unreliable and invalid. Yet, they still held teachers accountable to their previously held notions of what good teaching looks like. Amanda Nelson exemplified most participants when she stated, "instructional delivery, instructional planning, those to me when I'm going in and observing, I get much more teacher quality." Again, these results suggest a way out of the debate. If assistant principals cannot rely on SPG as a measure of teacher quality, but they can use their professional judgment and expertise to improve teacher quality with observable standards, why would we need market-based measures to directly link teaching to student achievement, especially if those links are so elusive?

Ineffective teachers. Student achievement-based teacher evaluation policies owe their existence to scholars who criticized previous policies for their inability to get rid of ineffective teachers (Milanowski, 2006; Weisberg et al., 2009). My study, however, demonstrates that the SPG portion of District A's policy is not likely to perform any better. It is clear that District A's aim in creating this unique policy was not to

145

systemically purge its teacher workforce of poor teachers. Their approach was significantly more liberal than other districts who chose to strictly align state test results to teachers. However, it is unlikely that District A intended to make it more difficult for administrators to evaluate out ineffective teachers.

The majority of participants did not perceive SPG policy as way to evaluate out ineffective teachers. They indicated ineffective teachers were protected by SPG policy, either by "playing teacher," according to Amanda Nelson, or by being part of strong CTs. If a CT was successful in meeting their student achievement goal for the year, and that goal was connected to the SPGs of weak teachers on that team, then ineffective teachers were protected. Or, if teachers created unrigorous goals they received *highly effective* or *effective* ratings worth 40%, making it impossible to receive an *ineffective* overall rating. Scholars who suggested linking teachers to student achievement as a way to remove ineffective teachers did not predict assistant principals would reject attempts to these links, nor did they predict the power of assistant principals to turn a summative policy into a formative one. Therefore, student achievement-based teacher evaluation policies should be examined with policy practice and the perceptions of evaluators in mind.

This study also confirms and supports scholarship that predicts legal challenges against policies that cannot show clear, valid links between teacher ratings and student achievement (DeMitchell et al., 2012; Pullin, 2013). Mark Sanders suggested that teachers might be able to grieve their evaluations if evaluators use SPGs that are contingent on students who are absent, inattentive, or disruptive. He stated, "[if] at the end of the day their student progress goal isn't met, there's a lot more things that they

146

could, if they really want to, push to try to defend why it didn't get to where it needed to be." Why, then, would school districts risk using such systems? Perhaps the best we can do is hold teachers indirectly accountable for student achievement using observable methods that, in theory and in practice, serve to identify quality teaching.

Derrington (2014) and Harris et al. (2014) conclude from broad survey data that student achievement-based teacher evaluation policies seems to run counter to teacher quality traits principals consider important. This study contributes to this scholarship by providing rich, specific examples of how assistant principals are altering SPG policy to suit their values, or ignoring SPG when it runs counter to their values. Assessing teacher quality, holding teachers accountable for student learning, and ridding public school systems of ineffective teachers are not as simple as market-based strategies suggest. Perhaps this is because the marketplace does not have to deal with such hard to measure subjects such as teacher quality and student achievement. In other words, it is easier to evaluate the output of a product or a service than it is to measure teacher impact on student achievement.

Does this mean teaching is more art than science? Art is in the eye of the beholder. Its qualities are subjective, dependent on context, and filled with meaning that many outside the context may not understand. Certainly if teaching was more scientific – i.e. do this, under these conditions, and get this result all the time – it would be easier to measure. However, clearly from the literature, the variables are too numerous to predict; and the results of this study confirm that assistant principals consider teaching to be more art than science. Teachers must artfully adjust their craft to multiple conditions making

147

measuring their effectiveness truly more difficult than might be done in market-based environments.

Instructional leadership. The findings from this study provide evidence that high school assistant principals have taken another step toward full integration as instructional leaders as opposed to building managers. The measures Glanz (1994) and Sun (2012) developed to assess assistant principal roles demonstrated this evolution. Some participants in this study used SPG as an opportunity to increase their capacity for instructional leadership in ways unintended by the policy. Further, other participants used the weight of SPG to enforce other instructional goals such as meeting CT or school improvement goals.

SPG as an instructional tool. Many participants decided to use SPG as a tool to facilitate data discussions for the benefit of individual and groups of students, enforce CT expectations, or carryout school-wide missions such as cross-curricular and interdisciplinary instruction. Mark Sanders described this as a "backdoor" approach to facilitating his school's cross-curricular goals. Ronald Coleman used the weight of SPG as a stick to force a teacher to comply with his expectation that teachers contribute to their CTs. Using SPG as a tool for instructional leadership was evident. However, some measure or sense of instructional leadership needed to be present before participants decided to use SPG as an instructional tool. In other words, strong instructional leaders could potentially convert any policy into an instructional tool.

One of the hypotheses of this study was that SPG is problematic because it is philosophically opposed to ideas about instructional leadership. Evidence from this study

showed that participants used SPG as an instructional tool to have data discussions with teachers. Interviews also revealed that assistant principals took on strong instructional leadership roles apart from the evaluation policy and that they did not use SPG as it was intended. In other words, this evidence suggests that participants crafted SPG to their own personal beliefs about teaching and learning so that SPG policy was impacted by the existence of instructional leadership, not the other way around. Perhaps student achievement-based teacher evaluation policies do indeed run counter to instructional leadership. But, what seems also to be true based on this study's results is that instructional leaders convert the policy into a formative instructional tool, thereby negating the punitive and summative tone of the policy.

The way participants in this study used SPG to increase their instructional leadership capacity confirms and extends research in this field. All the suggestions from pre- and post-RttT instructional leadership scholarship are in full use in District A. Participants were clearly demonstrating transformational instructional leadership as guided by Leithwood et al. (2004); or shared instructional leadership, as suggested by Marks and Printy (2003); and changing teaching institutions into learning institutions, as suggested by Rigby (2015). Rigby stated that how principals used teacher evaluation as an instructional tool was "the key lever" (p. 387). This study provides evidence that no matter what the policy's stated purpose, strong instructional leaders indeed consider student achievement-based teacher evaluation policies as key levers to lead instruction.

Shared accountability. Assistant principals worked with teachers to share accountability for student achievement through SPGs. Instructional leadership

scholarship promotes this kind of collaboration and shared responsibility (Cuban, 2001; DuFour, 2004; Evans, 1996; Fullan, 2001; Hoyle, 2002; Leithwood et al., 2004; Marks and Printy, 2003; Smith & Ellett, 2002). As participants in this study demonstrated, instructional leadership is associated with excellent interpersonal and communication skills. Michael Cox's notion of "reciprocal accountability" showed how the characteristics of instructional leadership identified by the above scholarship can be combined with strong accountability measures. Even though Michael's approach was an outlier in this study (as addressed above), it nevertheless shows a new path for instructional leadership in the accountability era. If policies that may seem punitive and harsh on the surface are mitigated by expert instructional leaders, this allows policy makers to employ summative evaluation policies if they feel they must. But, the skill of the assistant principal also helps protect, inform, and guide teachers in ways policy can never hope to do.

To be sure, instructional leadership is still not the only role for assistant principals. Participants in this study reported discipline, attendance, and other management activity as distractions to their instructional leadership. However, that some used SPG as a vehicle for instructional leadership, and that all described other strategies such as professional development, demonstrates the power of instruction over managerial tasks. Eric Alexander's experience at two different low SES schools – he described one as having much more discipline than the other – is an example of how making instruction the focus of everything you do has the potential to reduce discipline and attendance problems. This example supports and extends the *Effective Schools Research* (Bracey,

1984; Edmonds, 1979; Hallinger, 1992) because the culture in District A is also a vision where all decisions are made with instruction in mind.

Like previous scholarship on instructional leadership, this study still does not connect good instructional practices to increased student achievement, demonstrate a reduction in the achievement gap, or show that American students raised scores on international tests. But, this is only because teacher quality and student achievement continue to be very hard things to measure. Nonetheless, participants in this study were using SPG to identify and strategize how to target struggling individual students and groups of students in ways not intended by the policy. Therefore, again indirectly, the assumption is that student achievement is indeed positively impacted through increased instructional leadership. Just because we are trying to measure two things with so many moving parts and that we may never be able to draw a straight line between teacher quality and student achievement, does not mean that we should discontinue efforts to improve instruction.

Assistant principals. Assistant principal scholarship is closely related to the above instructional leadership literature. This study provides rich data that extends research on what assistant principals actually do with both instructional leadership and teacher evaluation policy. It shows assistant principals as having strong instructional leadership capacity regardless of SPG policy. In this way, my study compliments research by Rintoul (2012) who describes assistant principals as important to resolving some of the most important issues facing school communities.

On the other hand, this study contradicts Militello et al. (2015) and Celikten (2001). Both claimed assistant principal power and authority was subject to the principal and other school functions. They did not consider assistant principals vital to instructional leadership or to teacher evaluation. This study tells a very different story because it pays close attention to assistant principal perceptions in ways that few other studies have before. Assistant principals exhibited their authority not just over SPG policy, but over many other instructional functions. For example, Jose Foster, Mark Sanders, and Patricia Johnson were all designated as teacher evaluation administrators in their buildings. They provided guidance to teachers and fellow administrators, not the principal. And, as noted in the evidence above, administrative CTs also guided policy without the principal. Most participants also noted providing both targeted, individual professional development to teachers as well as school-wide professional development programs. For example, Timothy Hughes instituted a peer observation program. These all were indicators of strong instructional leadership capacity.

With regard to SPG policy, the power and authority participants exercised in this study was unilateral. Principals had little to do with any decisions on SPG. In some cases principals laid out broad expectations, but they did not contribute to the evaluation policy beyond that. Colby et al. (2002), Cosner et al. (2015), and Peterson (2004), in their examinations of teacher evaluation policy, all found assistant principal authority to be minimal. Of course, nothing happens without principal approval or without their broadly defined vision, mission, and goals. Yet, it is clear from the interviews in this study that principals gave wide latitude to their assistant principals because they trusted their

capacity to lead instruction effectively. The results of this study were unambiguous on this point. Perhaps District A is unique. Maybe the culture of District A is more trusting of assistant principals as instructional leaders. This could stem from District A's commitment to a PLC culture, or its history of setting high expectations for teaching and learning. In any case, these findings suggest an alternative to existing scholarship that minimizes the role of what this study has determined are important and powerful instructional actors.

Policy implementation. Participants in this study exercised their authority and power over SPG policy in substantial ways. The evidence collected extends, but also contradicts co-construction scholarship in some ways. What is clear is that previously held notions of powerful actors are instrumental to policy implementation.

The presence of "micropolitics" (Datnow, 2000, p. 359) was not evident in this study. Therefore, this study is not a good example of pure co-construction. These factors may have been present, but participants did not mention personalities, relationships, or situations among their teachers that might have altered the success or failure of SPG policy implementation. It is not clear why micropolitical factors were ignored by participants. This could be where my role as a principal in the system I was studying impacted results. Participants may have been reluctant to bring up negative relationships or negative situations out of concern I might form a negative perception of them. They may have also avoided negativity about their principals because they knew I knew them well. Also, in my experience, most assistant principals in District A believe in the power of creating positive relationships. This mantra is incorporated in much of the district

153

rhetoric. Therefore, to discuss negative relationships would be counter intuitive. Co-construction suggests that success of policy implementation is difficult to predict because of so many micropolitical factors. However, without evidence of micropolitical factors, this study cannot support co-construction in this way.

On the other hand, this study shows that SPG policy implementation supports co-construction theory because implementation was impacted greatly by assistant principal power and authority as well as the presence of school teams. Assistant principal power and authority was described above. The presence of strong school teams was another important co-construction factor that changed SPG policy from one that holds teachers accountable for student achievement to one that holds teams accountable for student achievement. Therefore, that these internal forces existed to impact policy as it was implemented, validates and confirms co-construction theory.

Significance

The findings of this study on high school assistant principal perceptions on SPG policy are significant because they provide empirical evidence that student achievement-based teacher evaluation policies may not be working as intended. It also brings to light how teacher evaluation policy is being implemented at the ground level by those actors, assistant principals, most essential to policy success or failure. This study also provides readers with evidence that some accountability policies, regardless of their intention, has the potential to increase the instructional leadership capacity of assistant principals.

The significance of this research is threefold: First, this study contributes empirical evidence to the debate on student achievement-based teacher evaluation policy

that has been bogged down by controversy over VAM. Movement from this debate has been difficult because scholars have been caught up on how to link teacher behavior with student achievement, instead of asking whether teacher behavior *should* be linked to student achievement. Results of this study suggest education policy makers and practitioners would be hard pressed to invent a straight line between teaching and learning that can be accurately and validly measured with quantifiable data. Instead, policy makers and practitioners would be better served by policies that improve instruction because this is what is controllable. Student learning, as this study suggests, is beyond the control of policy makers, and even beyond the control of teachers.

The second significant contribution this study makes is to policy makers who might be considering implementing, changing, or eliminating student achievement-based teacher evaluation policies in the future. This study offers no direct guidance to district policy makers on evaluation policies because District A is a unique setting with a unique policy. However, the findings are instructive to policy makers in that this study provides firm evidence that policy purpose does not necessarily equate to policy practice. Policy makers would be well served to take into account assistant principal perceptions before engaging in policy construction and implementation. Engaging those who directly evaluate teachers will guide any professional development that would need to take place prior to implementation. Also, knowing what assistant principals think and how they approach instructional leadership might also inform the policy before it is written.

Lastly, this study is a significant source of data for District A. Because of the way District A crafted this policy, it might not come as a surprise to its leaders that SPG is not

being used to assess teacher quality or to hold them accountable for student achievement. However, understanding how assistant principals are using SPG to increase their instructional leadership capacity might very well be surprising and useful information. This study offers several avenues for further study to assist District A in developing its leaders, improve teaching, and create a better understanding of teacher/administrator relationships and how these are impacting teaching and learning in their schools.

Future Study

Studies of student achievement-based teacher evaluation policies need to be conducted with more frequency across the nation. This study is just a small part of what needs to happen if we are to fully understand all the nuances of such intrusive policies in so many different contexts. Replication of this study might serve scholarship and school districts well. These studies would be able to account for the many different contextual factors that exist in individual districts.

In addition to replicating this study, District A would benefit from a correlational study to assess how assistant principals are rating SPGs as compared to the other six standards. With its on-line teacher evaluation data collection system, I imagine this could be done easily. The district could pull all Standard 7 scores, then pull Standard 1-6 scores combined, and devise statistical queries to assess whether or not these measures correlate. Such a study would inform the district more precisely on the use of SPGs, highlight potential misunderstandings about rating SPGs, and target professional development to administrators to improve the system.

Also, District A would benefit by replicating this study for teachers and principals. Teachers have been studied more frequently in the literature with regard to teacher evaluations, but District A has yet to survey or interview its own teachers to obtain data on how they perceive the policy, their level of stress, if they value SPGs, and their perceived impact on their own quality and on their students' achievement. Results of this study suggest that at least some teachers find value in the data discussions that ensure from the SPG process. Principals would also be an interesting population to study with regard teacher evaluation policy and instructional leadership. District A principals have been largely left out of the SPG policy creation and implementation process and their opinions of how this policy fits into the broader vision and mission of their schools would be an interesting study. Contextual studies are also recommended to assess contributions of CTs and other school teams mentioned in this study. Clearly, school teams have a large impact on SPGs, instructional leadership, and on teacher quality.

More study is also needed to examine outlier assistant principals who attempt to use SPG as it was intended – to assess teacher quality and hold them accountable for student achievement. Such an investigation may uncover contextual factors that might be forcing some assistant principals to behave differently than others. Or, this research may explain in more detail some assistant principals' philosophies about teaching and learning. Perhaps this would lead to better understanding of how to link teaching and learning. In any case, to ignore the minority view would be detrimental to scholarship in this field.

Finally, confounding variables apparent in this study need to be sorted out and studied independently. Of particular importance would be a study of how classroom observations has impacted assistant principal instructional leadership capacity. When, in 2012, District A's new teacher evaluation policy came out, the two major changes were SPGs and the large increase in required classroom observations. Past evaluations required three teacher observations, the new one requires six. Participants in this study all noted the importance of classroom observations in supporting and guiding teachers. This impact on instructional leadership should be studied more closely and independently of SPG.

Limitations

The findings of this study may be transferable to other public school systems that employ SPGs or SLOs as one method of linking teacher quality to student achievement. However, caution should be used as District A's teacher evaluation policy is quite unique. The findings of this study are not generalizable or transferable to other school systems that have student achievement-based teacher evaluation policies using VAM or other strictly defined methods to link teacher behaviors to student achievement. A key condition of this setting was the flexibility District A had in its SPG policy and subsequent decisions made by participants and teachers as they tried to work with the policy. Not all public school districts have or desire this kind of flexibility.

Further, this study was limited by a relatively small sample size of 16 participants. Also, by agreement with District A, I was not able to include racial makeup or background of participants. Though I made no assumptions about how these demographics might have related to perceptions of SPGs or instructional leadership, race

has typically been a key demographic to study because it often impacts perceptions. Also, personal background statistics such as where participants taught, who they taught, or where they went to school, may have had significant impact on how they formed their philosophies of education or attitudes toward teacher accountability. These perceptions were key to the findings of this study and should be explored in more detail.

This study is also limited by the time of year in which it was conducted. There were benefits to interviewing assistant principals at the close of the first marking period – SPG goal setting conferences were just finishing up and so these questions were fresh in their minds; administrators may have been more optimistic and less tired in the Fall as opposed to the Spring; there were fewer distractions at schools during the teacher workdays. Interviewing participants in the Spring may have resulted in different answers, perhaps more informed answers as the evaluation cycle would have been more complete. Or, toward the end of the year responses may have been more negative due to assistant principals heavy work load. Future replication studies should take time of year into consideration.

Conclusion

Discussion and debate about teaching and learning has changed as the result of federal involvement in public education. NCLB and RttT have led to many accountability expectations in states and districts all over the nation. One of the most impactful expectations has been the call to create student achievement-based teacher evaluation policies that attempt to hold teachers directly accountable for student learning. This study

brings up two issues indicating that policy makers and educational practitioners ought to think carefully before adopting future policies of this kind.

The first concern this study sheds light on is whether or not it is possible to ever hold teachers accountable for student achievement. Despite the liberal use of SPGs in District A and the many different ways teachers invented to try and draw straight lines between their performance and their students' achievement – assistant principals find these measures invalid and unreliable. This evidence is an indication of two possible realities: perhaps nobody can ever know what a student has learned; or, perhaps the purpose of teaching is not to raise student achievement but to provide students with opportunities to raise their own achievement.

Measuring student achievement has been an elusive task since the beginning of NCLB, and even before. There has been much said and written about the validity of NCLB mandated end-of-course tests, common core assessments, SATs, PISA, etc. There is no perfect measure of student learning. Plus, most of the variables that exist to prevent or boost students' chances for succeeding or failing exist outside the classroom. Therefore, is it really fair to assess teacher quality based on situations they cannot control and by using imperfect measures to assess what cannot be controlled? VAM claims to have found a way to account for these outside variables, but most social scientists and educators are unwilling to trust VAM creators' word that VAM is a valid measure of teacher effectiveness without further information on how this statistical model works (Sanders, 1998).

If, on the other hand, teachers were not held directly accountable for student achievement and instead held accountable for facilitating their students' own achievement, then it may be easier to evaluate them fairly and accurately. Certainly assistant principals would not have to manipulate data, or ignore SPG policy if they could measure how teachers provide opportunities for students. Assessing teachers in this way puts the onus of student achievement on the student who, arguably, has more control over his or her achievement than the teacher. Further, this approach does not absolve teacher responsibility to ensure learning with all the tools, resources, and opportunities available to them.

The second concern this study brings to light are assistant principals preexisting notions of teacher quality. Like student achievement, teacher quality has been hard to assess. Participants in this study clearly manipulated SPG ratings in order to fit their preconceived ideas of whether or not teachers were effective or not. How do they know good teaching? Assistant principals in this study used the other six standards and their extensive observations of teachers to come to conclusions about the quality of their teachers. However, these are very subjective measures. How are they validated or normed between assistant principals and between schools?

Again, we come to the question of purpose. As described above, scholarship on the purpose of teacher evaluations is polarized into formative and summative camps. This study raises serious concerns over any evaluation policy that is purely summative in its purpose because of the power assistant principals wield over policy co-construction. Because of how they apply subjective measures in implementing SPG, when assistant

principals use SPG as a summative tool, there is no guarantee that good teachers are rated effectively and no guarantee that bad teachers are rated ineffectively. On the other hand, assistant principals seemed to have little problem using teacher evaluation policy, even the SPG component, to inform and improve teaching practices.

Demand for teachers is at an all-time high. Teachers are overworked and underpaid. Teacher evaluation systems that support teacher improvement, assess their needs for and target professional development, encourage them to try new things, and otherwise support and guide them, are productive. Assistant principals are key players in implementing and using teacher evaluation policy to improve teacher quality. Whether or not student achievement-based teacher evaluation policies contribute to assistant principals' instructional leadership capacity depends on how they perceive the purpose of the policy. If they perceive the policy as important to their roles as instructional leaders they make good use of it in that manner. If they perceive the policy as working against their instructional leadership, they either ignore it or re-construct the policy.

Appendix A

IRB Approval Materials

GEORGE MASON UNIVERSITY

Office of Research Integrity and Assurance

Research Hall, 4400 University Drive, MS 6D5, Fairfax, Virginia 22030
Phone: 703-993-5445; Fax: 703-993-9590

DATE: June 2, 2016

TO: Diana D'Amico, PhD
FROM: George Mason University IRB

Project Title: [884955-1] Assistant Principal's Perceptions of Student Progress Goals and
Their Impact on Instructional Leadership

SUBMISSION TYPE: New Project

ACTION: DETERMINATION OF EXEMPT STATUS
DECISION DATE: June 2, 2016

REVIEW CATEGORY: Exemption category #2

Thank you for your submission of New Project materials for this project. The Office of Research Integrity
& Assurance (ORIA) has determined this project is EXEMPT FROM IRB REVIEW according to federal
regulations.

Please remember that all research must be conducted as described in the submitted materials.

Please note that any revision to previously approved materials must be submitted to the ORIA prior to
initiation. Please use the appropriate revision forms for this procedure.

If you have any questions, please contact Katherine Brooks at (703) 993-4121 or kbrook14@gmu.edu.
Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within George Mason
University IRB's records.

Interview Guide for
***Assistant Principals' Perceptions of Student Progress Goals and Their Impact on
Instructional Leadership***

Interview Guide *rev. 10-14-16*

Re: IRBNet 884955-1

1. In 2012, FCPS implemented it's new teacher evaluation policy that includes
   student progress goals as 40% of teachers' final rating. Generally, what are your
   thoughts of this policy?

   FOLLOW UP – Have your perceptions of this policy changed over time?

2. What personal values or philosophies do you think led you to these perceptions?

3. What do you think of the student progress goal standard as it relates to teacher
   quality? Does it help you assess quality, provide support to improve teacher
   quality?

4. To what extent do you think teachers should be held accountable for student
   learning?

5. Have you ever evaluated teachers using different evaluation policies?

   a. If so, does the student progress goal change the way you developed
      teachers or their instruction under those other systems? How?

   b. If not, can you envision any other ways to develop teachers or their
      instruction other than student progress goals? In other words, how else
      would you/do you support or help to improve teacher quality through an
      evaluation system?

6. To what extent do you have power or authority to implement the student progress
   standard?

7. What outside forces/factors limit, enhance, or influence the way you implement the student progress goal standard?

   PROMPT – Personalities; school culture; principal's authority, instructional guidance; influence from region/district/parents/students; collegial influence.

8. Does who you are evaluating matter when implementing the student progress goal? In other words, do you implement this standard the same way with everyone? Same from year to year? Does it help you improve or eliminate ineffective teachers?

9. Can you give me some examples of teacher created student progress goals – good, bad, typical?

   FOLLOW UP – What defines good or bad progress goals? Have you ever rejected a teacher's student progress goal or asked them to revise it? Examples? What happens when a teacher fails to meet the student progress goal, but you feel, from the other standards that they are an effective teacher?

10. Do you think there is a strong correlation between Standard 7 ratings and teacher quality?

11. In what ways, if any, are you an instructional leader? Can you give me examples?

    FOLLOW UP: Are any of these answers related to the teacher evaluation process? Student progress goals?

12. What, in your opinion, are the best ways to improve instruction or teacher quality?

13. What aspects of the current evaluation system allow you to be an instructional leader?

14. Does the student progress standard play a role in how you perceive yourself as an instructional leader?

    FOLLOW UP – Since the new system has been implemented, are you more, less, or the same in terms of instructional leadership? In other words has the student progress goal distracted you from, or enhanced your capacity to be an instructional leader?

Appendix C

Descriptive Statistics of Participants

Years of Administrator Experience

Figure B1. Participants' years of administrator experience.

Figure B2. Departments Supervised. In order to further conceal participant identity, no pseudonyms were used. The x-axis is the number of departments represented by participants. Most participants supervised more than one department.
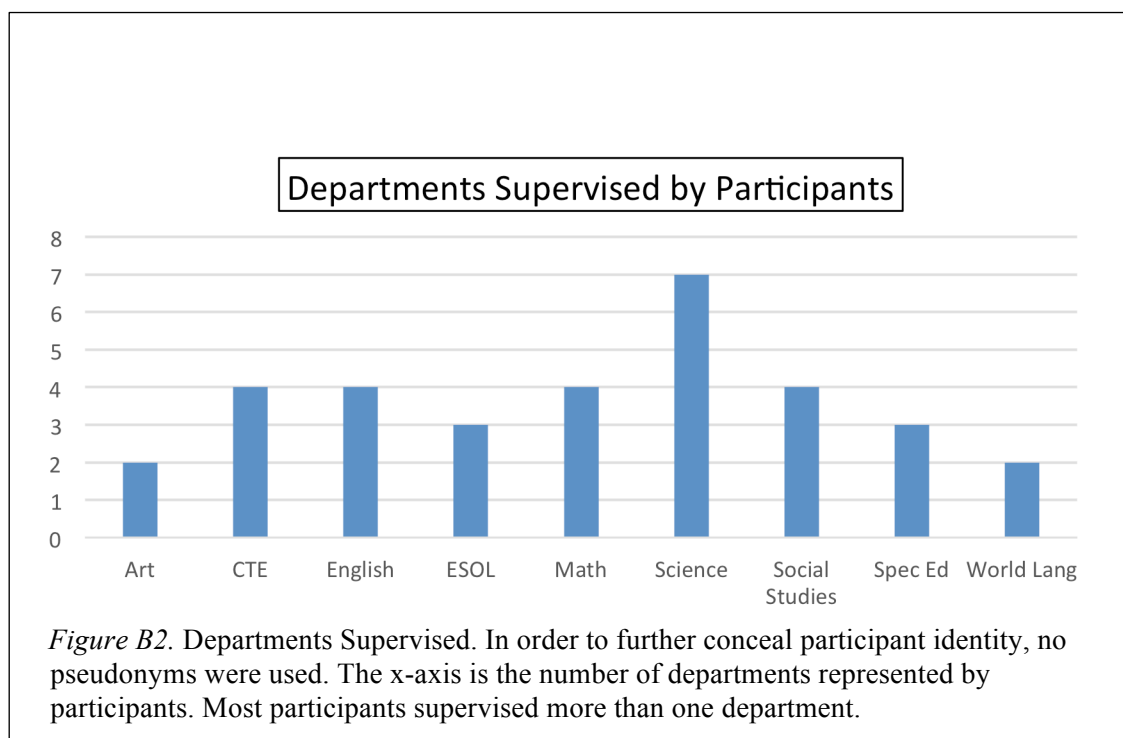
Table B1

Participants' gender and SES schools represented.

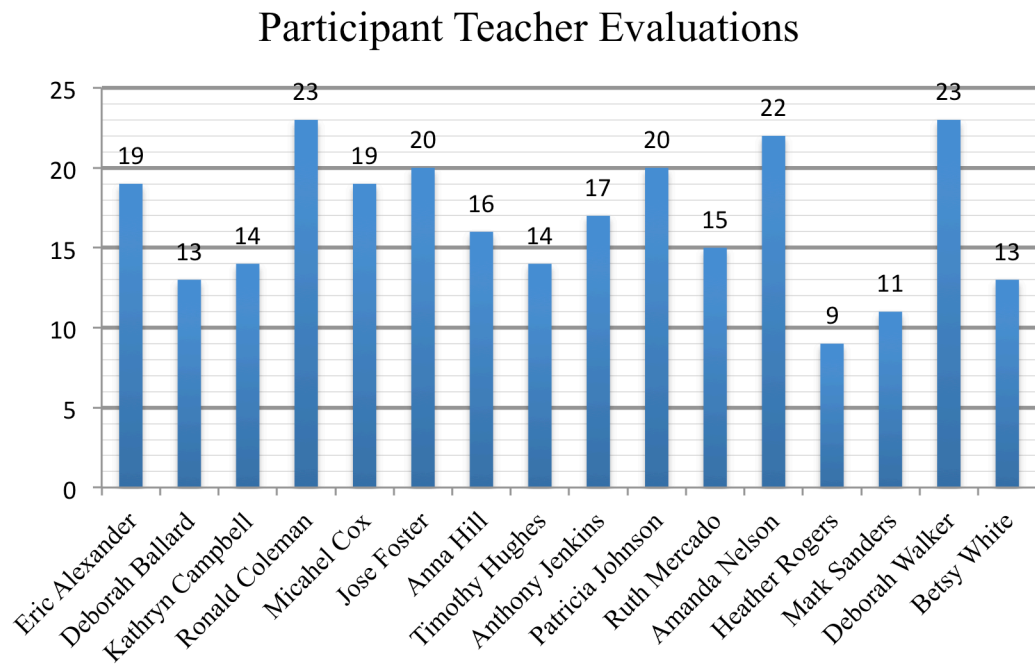| Category	Number of Participants
Male Participants	7
Female Participants	9
From Low SES Schools	8
From High SES Schools	8

Figure B3. Participants' number of teacher evaluations. Mean = 16.75. Mode = 19.

Table B2

Participant teaching and other experience.

Participant	Yrs Teaching	Subjects	Current HS	Prev. Exp
Assistant Principal 1	15	Math	HS 23	HS 8
Assistant Principal 2	23	Sped	HS 4	
Assistant Principal 3	12	Band	HS 6	
Assistant Principal 4	10	Sped Math	HS 1	HS 5
Assistant Principal 5	9	Social Studies	HS 21	
Assistant Principal 6	16	Phys Ed	HS 12	
Assistant Principal 7	14	Social Studies	HS 22	
Assistant Principal 8	7	Business	HS 7	
Assistant Principal 9	8	Math	HS 18	
Assistant Principal 10	10	Math	HS 10	HS 2
Assistant Principal 11	9	Spanish	HS 18	
Assistant Principal 12	5	Social Studies	HS 2	
Assistant Principal 13	8	English	HS 7	HS 24
Assistant Principal 14	10	Social Studies	HS 3	
Assistant Principal 15	17	Science	HS 1	Central
Assistant Principal 16	<u>16</u>	Science	HS 15	Central
Average Yrs Teaching	11.8			

Note. In order to further protect confidentiality, pseudonyms were not used. Schools are listed by confidential school codes. Previous experience refers to other non-teaching experience only. "Central" is central office administrator. If participants were assistant principals at other District A high schools, those schools are listed by confidential school codes. If blank, participants either did not have or did not share previous non-teaching experience.

Appendix D

Data Analaysis Themes and Codes

Name	Notes	Sources	References
Accountability		0	0
Admin on Student Achievement		2	2
Admin to Policy Implementation		2	3
District Accountability		4	4
Outside on Student Ach		3	3
SPG Implementation Accountability		2	4
Students on Student Ach		4	6
Teacher on Student Ach		5	12
Teacher to Job		1	2
Admin Role		4	9
Admin Time & Workload		4	15
Admin Values		2	5
All Students Can Learn		2	2
Congruence with Policy		2	2
Engagement, Est. Relationships		2	12

Incongruent with Policy		1	1
Values Data		2	2
Administrator Teacher Relationship		1	1
Conversations		4	15
Expectations for Teachers		4	9
Power & Authority		4	9
Trust		4	20
Impact of Teacher Eval Policy		1	1
General Policy Impact		2	3
Ineffective Teachers		4	10
Instructional Leadership, Impact On		4	12
Observations, Impact On		3	7
SPG Only, Impact		2	4
Student Achievement		5	14
Teacher Admin Relations		1	2
Teacher Quality, Impact On		4	12
Teacher Stress		3	4
Teacher Time & Workload		0	0
Instructional Leadership		2	2
Building Teacher Leadership Capacity		2	3
CTs, PLC		3	8

Data Conversations		4	13
Effecting Teacher Quality		3	7
Feedback on Teaching Strategies		4	16
Identifies Resources for Teachers		3	5
Professional Development		3	6
Teacher Goals		1	4
Overall Perceptions & Feelings		0	0
Anxious, Not Prepared		1	4
Comfortable in Role, Practice		2	2
Comparison to Other Policies		2	3
Dislikes Evaluation Policy		1	1
Eval Process & Procedures		3	7
Evaluation Examples		2	2
Likes Evaluation Policy		3	9
Meaningfulness		4	9
Other Aspects of Policy		2	3
Suggestions for Policy Changes		2	3
Validity of Final Eval Scores		4	13
Policy Implementation		1	1
Change Over Time - Implementation		5	5
Consistency - Admin Team, Schools		5	23

Differentiated Implementation		5	10
Forces, Factors Influencing Implementation		4	14
Implementation Fidelity		3	7
SPG Process & Steps		2	4
School Level Functions		0	0
School Culture & Values		3	7
School Improvement Plans		1	4
School Stakeholders		1	2
SPG Admin Perceptions		0	0
Confidence in SPG		2	6
Dislike or Frustration		3	3
Perfunctory		3	3
Supports SPG in Theory		4	9
Technical Criticisms		1	1
Unsure of SPG		2	5
SPG Characteristics		1	2
Other Measures of Progress		2	2
SPG Defined		4	7
SPG Rigor		3	11
SPG Validity		6	34
SPG Weight		4	7

Use of Data		3	6
SPG vs. Other School Goals		1	1
Congruence with CT or SIP		3	7
School, Admin Discourages Links to CT		0	0
School, Admin Doesn't Require CT Link		2	3
School, Admin Encourages Links to CT		3	7
Student Teacher Relationship		2	7
Understanding of SPG		5	14

Appendix E

Recruitment Materials

Invitation Email to District A Assistant Principals

Dear <Assistant Principal>,
        I am writing to ask for your participation in my dissertation study of high school assistant principals' perceptions of teacher evaluation policy. I am a doctoral student at George Mason University (GMU), and Principal at [a District A High School]. I have received permission from [District A] and GMU to conduct this study. The purpose of my study is to discover how high school assistant principals perceive the student progress goal standard of the FCPS teacher evaluation policy, and how these perceptions might impact your instructional leadership. I am seeking between 15 and 19 participants for a 35 minute interview that I will arrange at your complete comfort and convenience.
        The results of this study should make a significant contribution to existing literature on teacher evaluation and instructional leadership – two areas of research where assistant principal voice has been largely absent. I also believe this study will be useful to our district as we seek to better understand how we implement evaluation policy and instructional leadership practices. Further, I believe that you might benefit from the time spent discussing and reflecting on evaluation and instructional leadership ideas.
        Should you agree to help me in this study, the identities of you, your school, our district, and anyone you might refer to during the interview are strictly confidential. Codes linking any identifiable information are only known by me throughout the entire research project, and will be destroyed at the conclusion of the project.
        Your decision to participate is completely voluntary and you are under absolutely no obligation to agree or to even respond. I am sending this email to all [District A] high school assistant principals who evaluate teachers. I will not track non or negative replies.
        If you have any questions before deciding, please do not hesitate to contact me directly via my cell phone at [phone #], or by personal email at [my personal email address]. (Please do not use my [District A] email, as it would violate the agreement I have made with [District A]).
        Thank you in advance for considering this invitation.
        Sincerely,
        Matt Ragone

Informed Consent Form

Informed Consent Form for Research for
Assistant Principals' Perceptions of Student Progress Goals and Their Impact on Instructional Leadership

Diana D'Amico, Principal Investigator
Matt Ragone, Student Researcher
George Mason University

RESEARCH PROCEDURES
This research is being conducted to investigate how high school assistant principals perceive teacher evaluation policy, particularly the student progress goals component, and how these perceptions impact their capacity to lead instruction. If you agree to participate, you will be asked several questions in a semi-structured interview format. The interview will take approximately 30 to 45 minutes and be conducted in a time and location most convenient for you. If necessary, a brief follow-up phone conversation to verify or confirm responses may be requested after the interview. All interviews are audio recorded to ensure accuracy of answers. (If a follow-up phone call is necessary, it will not be recorded).

RISKS
There are no foreseeable risks for participating in this research.

BENEFITS
Benefits to your participation in this study are minimal. However, you may benefit by the reflective dialogue as it pertains to your daily work. Studies have shown that time to reflect on how professionals go about their work are rare. But, when provided the opportunity to reflect professionals benefit in a variety of ways from the experience (Blumberg, 2015).

CONFIDENTIALITY
The data in this study will be confidential. You have been assigned a randomly generated pseudonym. The study report will include no personal identifiers of you, your school, the teachers you supervise, or the district where you work. The district is referred to only as a large, suburban public school district. And, your school has been assigned a random number and is only referred to by that number.

Descriptive statistics about you will be collected for this study. However, the final report will not link this data to any personal identifying. Descriptive information includes – years of administrator experience, years of teaching experience, age, ethnicity, gender, departments you supervise, and number of teachers you evaluate.

All personal information is linked to your personal identifying information only during the data collection phase for the sole purpose of contacting you to collect or verify information. Once all information is collected and verified, all personal identifiers will be destroyed. I will have the only access to the code key.

CONSENT TO BE AUDIO RECORDED

Your interview will be audio recorded. Your name is linked to the beginning of the audio recording, again, for collection and verification purposes only. Your name will be deleted from the audio recording upon completion of the data analysis process. Audio recording will be conducted using my personal iPhone. The audio file will be transferred to my personal laptop, after which, I will delete the iPhone recording. Audio files will be transcribed into word files for data analysis.

All personal identifying information is deleted immediately after data analysis. However, audio recordings and their transcriptions must be preserved for five years at George Mason University. These records are secured. Only the Dean of the College of Education and Human Development, the chair of my dissertation committee, and I will have access to the files. After five years, these files will be deleted.

_____ I agree to audio taping.

_____ I do not agree to audio taping.

PARTICIPATION

Your participation is voluntary, and you may withdraw from the study at any time and for any reason. If you decide not to participate or if you withdraw from the study, there is no penalty or loss of benefits to which you are otherwise entitled. There are no costs to you or any other party.

CONTACT

This research is being conducted by Matthew J. Ragone, Doctoral Student, at George Mason University. I can be reached at 703-435-5002, or via email at mjragone@outlook.com, for questions or to report a research-related problem. My faculty advisor is Dr. Diana D'Amico. She can be reached at 703-993-5596, or via email at ddamico2@gmu.edu.You may contact the George Mason University Office of Research Integrity & Assurance at 703-993-4121 if you have questions or comments regarding your rights as a participant in the research.

This research has been reviewed according to George Mason University procedures governing your participation in this research.

CONSENT

I have read this form, all of my questions have been answered by the research staff, and I agree to participate in this study.

_____   _____
Signature   Printed Name

_____
Date of Signature

REFERENCE

Blumberg, P. (2015). How Critical Reflection Benefits Faculty as They Implement Learner-Centered Teaching. *New Directions for Teaching & Learning*, *2015*(144), 87–97. http://doi.org/10.1002/tl.20165

References

2015 best counties to live in. (2015). *Niche*. Retrieved February 15, 2016, from https://local.niche.com/rankings/counties/best-counties/

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95–135. http://doi.org/10.1086/508733

Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher*, *37*(2), 65–75. http://doi.org/10.3102/0013189X08316420

Andrews, R. L., & Soder, R. (1987). Principal leadership and student achievement. *Educational Leadership*, *44*(6), 9–11.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, *29*(1), 37–65.

Bates, R. C., & Shank, J. G. (1983). The associate principalship: A better, more effective way to manage schools. *NASSP Bulletin*, *67*(462), 111–114. http://doi.org/10.1177/019263658306746217

Berliner, D. (2014). Exogenous variables and value-added assessments: A fatal flaw. *Teachers College Record*, *116*(1), 1–31.

Bernanke, B. (2007, September). *FRB: Bernanke, education and economic competitiveness*. Washington, D.C. Retrieved from http://www.federalreserve.gov/newsevents/speech/bernanke20070924a.htm

Bossert, S. T., Dwyer, D. C., Rowan, B., & Lee, G. V. (1982). The instructional management role of the principal. *Educational Administration Quarterly*, *18*(3), 34–64. http://doi.org/10.1177/0013161X82018003004

Bracey, G. W. (1984). Of time and the classroom. *The Phi Delta Kappan*, *66*(1), 70.

Buckley, K., & Marion, S. (2011). *A summary of approaches used to evaluate educators in non-tested grades and subjects* (The Colorado Education Initiative). National

Center for the Improvement of Educational Assessment. Retrieved from
http://www.coloradoedinitiative.org/wp-content/uploads/2014/04/Summary-of-
Approaches-for-non-tested-grades_7-26-11.pdf

Burns, J. M. (1978). *Leadership*. New York, NY: Harper Collins.

Busch, S., & MacNeil, A. (2012). Critical advice from practicing assistant principals for
assistant principal preparation programs. In A. Shoho, B. Barnett, A. Tooms
(Eds.), *Examining the assistant principalship: New puzzles and perennial
challenges for the 21st Century*, (pp. 35–57). Charlotte, NC: Information Age
Publishing Inc.

Calabrese, R. L. (1991). Effective assistant principals: What do they do? *NASSP Bulletin*,
75(533), 51–57. http://doi.org/10.1177/019263659107553311

Callister Everson, K., Feinauer, E., & Sudweeks, R. R. (2013). Rethinking teacher
evaluation: A conversation about statistical inferences and value-added models.
Harvard Educational Review, *83*(2), 349–370.

Carraway, J. H., & Young, T. (2015). Implementation of a districtwide policy to improve
principals' instructional leadership principals' sensemaking of the skillful
observation and coaching laboratory. *Educational Policy*, *29*(1), 230–256.
http://doi.org/10.1177/0895904814564216

Celikten, M. (2001). The instructional leadership tasks of high school assistant principals.
Journal of Educational Administration, *39*(1), 67–76.
http://doi.org/10.1108/09578230110380742

Center on Education Policy. (2015, November 20). Retrieved February 14, 2016, from
http://www.cep-dc.org/index.cfm?DocumentSubTopicID=48

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D.
(2010). *How does your kindergarten classroom affect your earnings? Evidence
from Project STAR*. National Bureau of Economic Research. Retrieved from
http://www.nber.org/papers/w16381

[City Public Schools] Teacher Performance Evaluation System. (2012, December).
Retrieved from
http://web.[city].k12.[state].us/Portals/0/assets/ProfessionalDevelopment/pdfs/Tea
cherPerformanceEvaluationrevised73012a-12-12-12.pdf

Cizek, G. J. (2000). [Review of *Review of contemporary assessment for educators*, by C.
I. Chase]. *Journal of Educational Measurement*, *37*(2), 179–185.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school. *Journal of Human Resources*, *45*(3), 655–681.

Coburn, C. E. (2005). Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Educational Policy*, *19*(3), 476–509. https://doi.org/10.1177/0895904805276143

Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002). Perceptions of teacher evaluation systems and their impact on school improvement, professional development and student learning. Retrieved from http://eric.ed.gov.mutex.gmu.edu/?id=ED464916

Collins, C., & Amrein-Beardsley, A. (2013). Putting growth and value-added models on the map: A national overview. *Teachers College Record*, *116*(1), 1-32.

Cosner, S., Kimball, S. M., Barkowski, E., Carl, B., & Jones, C. (2015). Principal roles, work demands, and supports needed to implement new teacher evaluation. *Mid-Western Educational Researcher*, *27*(1), 76–95.

Cruickshank, D. R. (1990). *Research that informs teachers and teacher educators*. Phi Delta Kappa. Retrieved from http://eric.ed.gov.mutex.gmu.edu/?id=ED325476

Cuban, L. (1993). *How teachers taught: Constancy and change in American classrooms 1880-1990* (2nd ed.). New York, NY: Teachers College Press.

Cuban, L. (2001). *How can I fix it? Finding solutions and managing dilemmas: An educator's road map*. New York, NY: Teachers College Press.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching, 2nd Edition* (2nd ed.). Alexandria, VA: ASCD. Retrieved from http://www.ascd.org/publications/books/106034.aspx

Darling-Hammond, L. (1986). A proposal for evaluation in the teaching profession. *The Elementary School Journal*, *86*(4), 531–551.

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

Darling-Hammond, L., Newton, X., & Wei, R. C. (2010). Evaluating teacher education outcomes: A study of the Stanford Teacher Education Programme. *Journal of Education for Teaching*, *36*(4), 369–388. http://doi.org/10.1080/02607476.2010.513844

Datnow, A. (2000). Power and politics in the adoption of school reform models. *Educational Evaluation and Policy Analysis*, *22*(4), 357–374. http://doi.org/10.3102/01623737022004357

Datnow, A., Hubbard, L., & Mehan, H. (1998). Educational reform implementation: A co-constructed process. Research report 5. Retrieved from http://eric.ed.gov/?id=ED428430

Davis, D. R., Ellett, C. D., & Annunziata, J. (2002). Teacher evaluation, leadership and learning organizations. *Journal of Personnel Evaluation in Education*, *16*(4), 287–301.

DeMitchell, T. A., DeMitchell, T. A., & Gagnon, D. (2012). Teacher effectiveness and value-added modeling: Building a pathway to educational malpractice? *Brigham Young University Education & Law Journal*, June 2012(2), 257–301.

Department of Education Awards $200 Million to Seven States to Advance K-12 Reform. (2011, December 23). [Press Releases]. Retrieved April 16, 2014, from http://www.ed.gov/news/press-releases/department-education-awards-200-million-seven-states-advance-k-12-reform

Derrington, M. L. (2014). Teacher evaluation initial policy implementation superintendent and principal perceptions. *Planning & Changing*, *45*(1-2), 120–137.

Digest of education statistics, 2012. (2012, June). Retrieved March 23, 2014, from http://nces.ed.gov/programs/digest/d12/tables/dt12_036.asp

[District A] ESEA flexibility map page. (2015, March 31). [U.S. Department of Education]. Retrieved February 14, 2016, from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/map/[state].html

[District A] home page (2015, October 14). Retrieved November 29, 2015, from http://www.[districta].edu/about/index.shtml

Dodson, R. L. (2015). Kentucky principal perceptions of the state's new teacher evaluation system: A survey analysis. *Educational Research Quarterly*, *39*(2), 53–74.

Donaldson, M., Cobb, C., LeChasseur, K., Gabriel, R., Gonzales, R., Woulfin, S., & Makuch, A. (2014). *An evaluation of the pilot implementation of Connecticut's system for educator evaluation and development*. UConn Center for Education Policy Analysis, Neag School of Education.

DuFour, R. (2004). What is a "professional learning community"? *Educational Leadership*, *61*(8), 6–11.

Duke, D. L. (2005). *Education empire: The evolution of an excellent suburban school system*. Albany, NY: State University of New York Press.

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, *37*(1), 15-24.

ESEA Flexibility. (2012, June 7). U.S. Department of Education. Retrieved from https://www.ed.gov/sites/default/files/esea-flexibility-request-w3.doc

Evans, R. (1996). *The human side of school change: Reform, resistance, and the real-life problems of innovation*. San Francisco, CA: Jossey-Bass Inc.

Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher*, *43*(2), 100–107. http://doi.org/10.3102/0013189X14521864

Frase, L. E., & Streshly, W. (1994). Lack of accuracy, feedback, and commitment in teacher evaluation. *Journal of Personnel Evaluation in Education*, *8*(1), 47–57.

Fuhrman, S., Clune, W., & Elmore, R. (1991). Research on education reform: Lessons on the implementation of policy. In *Education Policy Implementation* (pp. 197–218). Albany, NY: State University of New York Press.

Fullan, M. (2001). *The new meaning of educational change* (Third). New York, NY: Teacher College Press.

Glanz, J. (1994). Redefining the roles and responsibilities of assistant principals. *The Clearing House*, *67*(5), 283–287.

Glesne, C. (2011). *Becoming qualitative researchers*. Boston, MA: Pearson Education, Inc.

Goldhaber, D. D., & Brewer, D. J. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources*, *32*, 505–523. http://doi.org/10.2307/146181

Goldhaber, D. D., & Brewer, D. J. (1999). A three-way error components analysis of educational productivity. *Education Economics*, *7*(3), 199–208.

Goldhaber, D., Walch, J., & Gabele, B. (2014). Does the model matter? Exploring the relationship between different student achievement-based teacher assessments. *Statistics and Public Policy*, *1*(1), 28–39.

Goldrick, L. (2002). *Improving teacher evaluation to improve teaching quality* (pp. 1–8). National Governors Association. Retrieved from http://www.nga.org/files/live/sites/NGA/files/pdf/1202IMPROVINGTEACHEVAL.pdf

Good, T. L. (2014). What do we know about how teachers influence student performance on standardized tests: And why do we know so little about other student outcomes. *Teachers College Record*, *116*(1), 1–41.

Greenfield, W. D. (1985). Developing an instructional role for the assistant principal. *Education and Urban Society*, *18*(1), 85–92. http://doi.org/10.1177/0013124585018001005

Guidelines for Uniform Performance Standards and Evaluation Criteria for Teachers. (2011). [State] Department of Education. Retrieved from http://[state]-sig-training.wmwikis.net/file/view/W-[state]DOE+Guidelines.docx

Hallinger, P. (1992). The evolving role of American principals: From managerial to instructional to transformational leaders. *Journal of Educational Administration*, *30*, 35–48.

Hallinger, P., & Murphy, J. (1985). Assessing the instructional management behavior of principals. *The Elementary School Journal*, *86*(2), 217–247.

Halverson, R., Grigg, J., Prichett, R., & Thomas, C. (2015). The new instructional leadership. *Journal of School Leadership*, *25*(3), 447–481.

Hansen, M., Lemke, M., & Sorensen, N. (2013). Combining multiple performance measures: Do common approaches undermine districts' personnel evaluation systems. *National Center for Analysis of Longitudinal Data in Education Research (CALDER)*, (Working Paper #118). Retrieved from http://files.eric.ed.gov/fulltext/ED553415.pdf

Hanushek, E. (1971). Teacher characteristics and gains in student achievement: Estimation using micro data. *The American Economic Review*, *61*(2), 280–288. http://doi.org/10.2307/1817003

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, *30*(3), 466–479. http://doi.org/10.1016/j.econedurev.2010.12.006

Hanushek, E. A., & Rivkin, S. G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of Policy Analysis & Management*, *28*(3), 366–393. http://doi.org/10.1002/pam.20437

Hanushek, E. A., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *The American Economic Review*, *100*(2), 267–271. http://doi.org/10.2307/27805002

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. http://doi.org/10.3102/0002831213517130

Hazi, H. M., & Rucinski, D. A. (2009). Teacher evaluation as a policy target for improved student learning: A fifty-state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives*, *17*(5), 1–19.

Heck, R. H. (1992). Principals' instructional leadership and school performance: Implications for policy development. *Educational Evaluation and Policy Analysis*, *14*(1), 21–34. http://doi.org/10.3102/01623737014001021

Heneman III, H. G., & Milanowski, A. T. (2003). Continuing assessment of teacher reactions to a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, *17*(2), 173–195. http://doi.org/10.1023/B:PEEV.0000032427.99952.02

Herlihy, C., Karger, E., Pollard, C., Hill, H., Kraft, M., Williams, M., & Howard, S. (2014). State and local efforts to investigate the validity and reliability of scores from teacher evaluation systems. *Teachers College Record*, *116*(1), 1–28.

Hoff, D. J. (2008, August 27). Democrats air dueling ideas on education. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2008/08/27/01dems.h28.html?qs=NCLB+stalled

Honig, M. (2006). *New directions in education policy implementation: Confronting complexity*. Albany, NY: State University of New York Press.

Hoyle, J. R. (2002). *Leadership and the force of love: Six keys to motivating with love*. Thousand Oaks, CA: Corwin.

Hubbard, L., Stein, M. K., & Mehan, H. (2006). *Reform as learning: School reform, organizational culture, and community*. New York, NY: RoutledgeFalmer Press.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2006). *What does certification tell us about teacher effectiveness? Evidence from New York City* (Working Paper No. 12155). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w12155

Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Evaluating teacher effectiveness. *Education Next*, *11*(3), 54–60.

Killian, B. R. (2010). *Administrators' and teachers' perceptions of the efficacy of the Missouri Performance-Based Teacher Evaluation model* (Doctoral dissertation). St. Louis University, St. Louis, MO. Retrieved from http://search.proquest.com.mutex.gmu.edu/docview/753941094/abstract

Kimball, S. M., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, *45*(1), 34–70. http://doi.org/10.1177/0013161X08327549

Klein, A. (2012, July 18). More than half of states now have NCLB waivers. *Education Week*, *31*(36), 22–23.

Klein, K. J., & Sorra, J. S. (1996). The challenge of innovation implementation. *The Academy of Management Review*, *21*(4), 1055–1080. http://doi.org/10.2307/259164

Koru, J. M. (1993). The assistant principal: Crisis manager, custodian, or visionary? *NASSP Bulletin*, *77*(556), 67–71. http://doi.org/10.1177/019263659307755609

Kowalski, T. J., & Dolph, D. A. (2015). Principal dispositions regarding the Ohio Teacher Evaluation System. *AASA Journal of Scholarship and Practice*, *11*(4). Retrieved from http://ecommons.udayton.edu/eda_fac_pub/16/

Kupermintz, H. (2002). Teacher effects as a measure of teacher effectiveness: Construct validity considerations in TVAAS (Tennessee Value-Added Assessment System) (Technical Report No. 563). University of California. Retrieved from http://www.cse.ucla.edu/products/Reports/TR563.pdf.

Lacireno-Paquet, N., Morgan, C., & Mello, D. (2014). How states use student learning objectives in teacher evaluation systems: A review of state websites. *Washington, DC: US Department of Education, Institute of Education Sciences*. Retrieved from http://vamboozled.com/wp-content/uploads/2015/02/here.pdf

Lavigne, A. (2014). Exploring the intended and unintended consequences of high-stakes teacher evaluation on schools, teachers, and students. *Teachers College Record*, *116*(1). Retrieved from http://www.tcrecord.org.mutex.gmu.edu/library/pdf.asp?ContentId=17294

Lavigne, A. L., & Good, T. L. (2013). *Teacher and student evaluation: Moving beyond the failure of school reform*. Routledge.

Lazear, E. P. (2003). Teacher incentives. *Swedish Economic Policy Review*, *10*(2), 179–214.

Le Fevre, D. M., & Robinson, V. M. J. (2015). The interpersonal challenges of instructional leadership: Principals' effectiveness in conversations about performance issues. *Educational Administration Quarterly*, *51*(1), 58–95. http://doi.org/10.1177/0013161X13518218

Lefgren, L., & Sims, D. (2012). Using subject test scores efficiently to predict teacher value-added. *Educational Evaluation and Policy Analysis*, *34*(1), 109–121. http://doi.org/10.3102/0162373711422377

Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004). Review of research: How leadership influences student learning. Retrieved from http://conservancy.umn.edu/handle/11299/2035

Levin, H. (2012). More than just test scores. *Prospects (00331538)*, *42*(3), 269–284. http://doi.org/10.1007/s11125-012-9240-z

Lochmiller, C. R. (2016). Examining administrators' instructional feedback to high school math and science teachers. *Educational Administration Quarterly*, *52*(1), 75–109. http://doi.org/10.1177/0013161X15616660

Louis, K. S., Febey, K., & Schroeder, R. (2005). State-mandated accountability in high schools: Teachers' interpretations of a new era. *Educational Evaluation & Policy Analysis*, *27*(2), 177–204.

Maharaj, S. (2014). Administrators' views on teacher evaluation: Examining Ontario's teacher performance appraisal. *Canadian Journal of Educational Administration & Policy*, *152*, 1–58.

Mangiante, E. M. S. (2011). Teachers matter: Measures of teacher effectiveness in low-income minority schools. *Educational Assessment Evaluation and Accountability*, *23*(1), 41–63. http://doi.org/10.1007/s11092-010-9107-x

Marion, S., DePascale, C., Domaleski, C., Gong, B., & Diaz-Bilello, E. (2012). Considerations for analyzing educators' contributions to student learning in non-tested subjects and grades with a focus on student learning objectives. *Center for Assessment*. Retrieved from http://www.nciea.org/publication_PDFs/Measurement%20Considerations%20for%20NTSG_052212.pdf

Marks, H., & Printy, S. (2003). Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*, *39*(3), 370–397.

Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, *86*(10), 727–735.

Maslow, V. J., & Kelley, C. J. (2012). Does evaluation advance teaching practice? The effects of performance evaluation on teaching quality and system change in large diverse high schools. *Journal of School Leadership*, *22*(3), 600–632.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). SAGE Publications, Inc.

Maxwell, J. A., & Miller, B. A. (2008). Categorizing and connecting strategies in qualitative data analysis. In P. Leavy, & S. Hesse-Biber (Eds.), *Handbook of emergent methods* (pp. 461–477). New York, NY: Guilford Press.

McGuinn, P. (2012). Stimulating reform Race to the Top, competitive grants and the Obama education agenda. *Educational Policy*, *26*(1), 136–159. http://doi.org/10.1177/0895904811425911

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education*, *79*(4), 33–53.

Milanowski, A. T. (2006). Split roles in performance evaluation: Field study involving new teachers. *Journal of Personnel Evaluation in Education*, *18*(3), 153–169. http://doi.org/10.1007/s11092-006-9017-0

Militello, M., Fusarelli, B. C., Mattingly, A., & Warren, T. (2015). We do what we're told: How current assistant principals practice leadership and how they wish they could. *Journal of School Leadership*, *25*(2), 194-222. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=10526846&AN=102221899&h=03%2Fjn8VfFYoUjRx65%2Fle8oLRMBW6Thg7CaFeqSQOM9OendZOHtfVsMtvfiGhocPU1DohjhT8vn2tQhQc8uKVgQ%3D%3D&crl=c

Mulligan, C. B. (1999). Galton versus the human capital approach to inheritance. *Journal of Political Economy*, *107*(6), S184–224. http://doi.org/http://www.jstor.org/action/showPublication?journalCode=jpoliecon

Murnane, R. J. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings. *Journal of Policy Analysis and Management*, *19*(4), 547–568. http://doi.org/10.1002/%28ISSN%291520-6688/issues

Murphy, J. (1988). Methodological, measurement, and conceptual problems in the study of instructional leadership. *Educational Evaluation and Policy Analysis*, *10*(2), 117–139. http://doi.org/10.3102/01623737010002117

Murphy, J., Hallinger, P., & Heck, R. H. (2013). Leading via teacher evaluation the case of the missing clothes? *Educational Researcher*, *42*(6), 349–354. http://doi.org/10.3102/0013189X13499625

National Commission of Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, D.C.: U.S. Government Printing Office. Retrieved from http://datacenter.spps.org/uploads/SOTW_A_Nation_at_Risk_1983.pdf

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. http://doi.org/10.1177/0013161X12456700

Newton, X. A., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Education Policy Analysis Archives*, *18*(23), 1–23.

Oleszewski, A., Shoho, A., & Barnett, B. (2012). The development of assistant principals: A literature review. *Journal of Educational Administration*, *50*(3), 264–286. http://doi.org/10.1108/09578231211223301

Ovando, M. N., & Ramirez Jr., A. (2007). Principals' instructional leadership within a teacher performance appraisal system: Enhancing students' academic success. *Journal of Personnel Evaluation in Education*, *20*(1-2), 85–110. http://doi.org/10.1007/s11092-007-9048-1

Peterson, K. (2004). Research on school teacher evaluation. *NASSP Bulletin*, *88*(639), 60–79. http://doi.org/10.1177/019263650408863906

Petrides, L., Jimes, C., & Karaglani, A. (2014). Assistant principal leadership development: A narrative capture study. *Journal of Educational Administration*, *52*(2), 173–192. http://doi.org/10.1108/JEA-01-2012-0017

Praetorius, A.-K., Lenske, G., & Helmke, A. (2012). Observer ratings of instructional quality: Do they fulfill what they promise? *Learning and Instruction*, *22*(6), 387–400. http://doi.org/10.1016/j.learninstruc.2012.03.002

Pullin, D. (2013). Legal issues in the use of student test scores and value-added models (VAM) to determine educational quality. *Education Policy Analysis Archives, 21*(6/7), 1–27.

Ramirez, A., Clouse, W., & Davis, K. W. (2014). Teacher evaluation in Colorado: How policy frustrates practice. *Management in Education*, *28*(2), 44–51. http://doi.org/10.1177/0892020613511264

Range, B., Scherz, S., Holt, C., & Young, S. (2011). Supervision and evaluation: The Wyoming perspective. *Educational Assessment, Evaluation & Accountability*, *23*(3), 243–265. http://doi.org/10.1007/s11092-011-9123-5

Ravitch, D. (2014). Foreword. *Teachers College Record*, *116*(1), 1–4.

Rigby, J. (2015). Principals' sensemaking and enactment of teacher evaluation. *Journal of Educational Administration*, *53*(3), 374–392. http://doi.org/10.1108/JEA-04-2014-0051

Rintoul, H. (2012). The new urgency of vice-principalship: Ethical decision making. In A. Shoho, B. Barnett, A. Tooms (Eds.), *Examining the assistant principalship: New puzzles and perennial challenges for the 21st Century* (pp. 35–57). Charlotte, NC: Information Age Publishing Inc.

Riordan, J., Lacireno-Paquet, N., Shakman, K., Bocala, C., & Chang, Q. (2015). Redesigning teacher evaluation: Lessons from a pilot implementation. Retrieved from http://files.eric.ed.gov/fulltext/ED552484.pdf

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417–458. http://doi.org/10.2307/3598793

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*(2), 247–252.

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

Rousmaniere, K. (2013). *The principal's office: A social history of the American school principal*. Albany, NY: State University of New York Press.

Ruffini, S. J., Makkonen, R., Tejwani, J., & Diaz, M. (2014). Principal and teacher perceptions of implementation of multiple-measure teacher evaluation systems in Arizona. *Regional Educational Laboratory West*. Retrieved from http://eric.ed.gov/?id=ED548539

Sanders, W. L. (1998). Value-added assessment. *School Administrator*, *55*(11), 24-28.

Sanders, W. L. (2004). Compounding errors. *Phi Delta Kappan*, 86(2), 174-175.

Shoho, A., Barnett, B., & Tooms, A. (2012). *Examining the assistant principalship: New puzzles and perennial challenges for the 21st Century*. Charlotte, NC: Information Age Publishing Inc.

Smith, W., & Ellett, C. D. (2002). *A mixed methods study of small jazz combo as a metaphor for leadership density in schools*. Presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, *72*(3), 387–431. http://doi.org/10.3102/00346543072003387

Spring, J. H. (2014). *Political agendas for education: From Race to the Top to saving the planet* (5th ed.). New York, NY: Routledge. Retrieved from http://samples.sainsburysebooks.co.uk/9781134626366_sample_553267.pdf

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, *62*(4), 339–355. http://doi.org/10.1177/0022487111404241

Sullivan, J. (2012). A collaborative effort: Peer review and the history of teacher evaluations in Montgomery County, Maryland. *Harvard Educational Review*, *82*(1), 142–152.

Sun, A. (2012). Sharing instructional leadership as assistant principals in an accountability-oriented environment. In A. Shoho, B. Barnett, A. Tooms (Eds.), *Examining the assistant principalship: New puzzles and perennial challenges for the 21st Century* (pp. 153–178). Charlotte, NC: Information Age Publishing Inc.

Supovitz, J. (2012). *The linking study - first year results: A report of the first year effects of an experimental study of the impact of feedback to teachers on teaching and learning*. University of Pennsylvania. Retrieved from www.cpre.org

Teacher Performance Evaluation Program Handbook. (2013, August). [District A Public Schools]. Retrieved from http://www.[districta].edu/hr/epd/evaluations/teacher/PDF/TeacherHB.pdf

Terosky, A. L. (2014). From a managerial imperative to a learning imperative experiences of urban, public school principals. *Educational Administration Quarterly*, *50*(1), 3–33. http://doi.org/10.1177/0013161X13488597

These are the 10 best counties in America (2014, August 1). *Movoto*. Retrieved February 15, 2016, from http://www.movoto.com/blog/top-ten/best-counties-in-america/

Torff, B. (2005). Getting it wrong on threats to teacher quality. *Phi Delta Kappan*, *87*(4), 302–305.

Toscano, S. (2014). Ranked: These are officially the 10 best counties in America (2014, August 28). *Huffington Post*. Retrieved February 15, 2016, from http://www.huffingtonpost.com/2014/08/28/best-counties-america_n_5722930.html

Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.

U.S. Department of Education. (2009). Race to the Top Fund. *Federal Register*, *74*(144), 37804–37814.

Viadero, D. (2008, April 23). Researchers gain insight into education's impact on nations' productivity. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2008/04/23/34riskresearch_ep.h27.html?qs=competition+with+finland

Weber, G. (1971). *Inner-city children can be taught to read: Four successful schools* (No. 18). Washington, D.C.: Council for Basic Education. Retrieved from http://files.eric.ed.gov/fulltext/ED057125.pdf

Weick, K. E. (1995). *Sensemaking in organizations*. Newbury Park, CA: SAGE.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect. *Education Digest*, *75*(2), 31–35.

Weiss, R. (1994). *Learning from strangers: The art and method of qualitative interview studies*. New York, NY: The Free Press.

Whitehurst, G. J. R., Chingos, M. M., & Lindquist, K. M. (2014). Evaluating teachers with classroom observations. *Brown Center on Education Policy: Brookings Institute*. Retrieved from http://www.brookings.edu/~/media/research/files/reports/2014/05/13-teacher-evaluation/evaluating-teachers-with-classroom-observations.pdf

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1985). Teacher evaluation: A study of effective practices. *Elementary School Journal*, *86*(1), 60–121.

Wood, C. J., & Pohland, P. A. (1983). Teacher evaluation and the "Hand of History." *Journal of Educational Administration*, *21*(2), 169–181.

Young, T., & Lewis, W. D. (2015). Educational policy implementation revisited. *Educational Policy*, *29*(1), 3–17. https://doi.org/10.1177/0895904815568936

Young, S., Range, B., Hvidston, D., & Mette, I. (2015). Teacher evaluation reform: Principals' beliefs about newly adopted teacher evaluation systems. *Planning & Changing*, *46*(1/2), 158–174.

Youngs, P., & King, M. B. (2002). Principal leadership for professional development to build school capacity. *Educational Administration Quarterly*, *38*(5), 643–670. http://doi.org/10.1177/0013161X02239642

Biography

Matthew J. Ragone graduated from Mary Washington College, Fredericksburg, Virginia in 1990, having already begun work on his teaching license at George Mason University in that same year. He received is first Masters in Curriculum and Instruction from George Mason in 1993. Matthew worked as a Social Studies teacher in the district of study for 13 years. Matthew finished work on his Education Specialist degree in Administration and Supervision in 2003. The following year, he became an assistant principal in the same district. In 2008, Matthew was appointed to his first principalship, and moved to his second school in the Fall of 2014, where he is currently serving.