

Implementing a Standards-Based Teacher Evaluation System:  
Learning Experiences for Administrators in an Urban School District

by

Chris Canelake

A Dissertation Presented in Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Education

Approved March 2012 by the  
Graduate Supervisory Committee:

Debby Zambo, Chair  
Cory Hansen  
Tacy Ashby

ARIZONA STATE UNIVERSITY

May 2012

## ABSTRACT

Policymakers at the national level have recently initiated K-12 education reforms focused on teacher quality and teacher evaluation. Far-reaching legislation was subsequently enacted in the state of Arizona requiring schools to adopt standards-based teacher evaluation systems and link them to student outcomes. The end product is to result in annual summative measures of teacher effectiveness. Because of this, Arizona school administrators have become concerned about rapidly becoming experts in high-stakes teacher evaluation. Principals rarely have time to come together to talk about teacher evaluation, and consider the reliability of their evaluations and how to use teacher evaluation to help teachers improve their practice.

This action research study focused on a group of nine administrators in a small urban district grappling with a more complex and high-stakes teacher evaluation system. An existing community of practice was engaged to help administrators become more effective, fair, and consistent evaluators. Activities were designed to engage the group in dynamic, contextualized learning. Participants interacted in small groups to interpret the meaning of newly adopted evaluation instruments and professional teaching standards, share practical knowledge, and compare teacher evaluation experiences in classrooms.

Data were gathered with mixed methods. Prior to, and immediately after engaging in this 20-week innovation, principals and district administrators were given two surveys and interviewed about teacher evaluation. Additionally, a detailed record of this project was kept in the form of meeting records and a

research journal. Quantitative and qualitative data were triangulated to validate findings.

Results identified concerns and understandings of administrators as they attempted to come to a shared consensus regarding teacher evaluation, increase inter-rater reliability, and use teacher evaluation to improve professional practice. As a result of working and learning together administrators lowered their concerns about inter-rater reliability. Other concerns, however, remained and grew. Administrators found the process of gaining a common understanding of teacher evaluation to be complex and far more time-consuming than anticipated. Intense concerns about alignment of the evaluation system with other reforms led these administrators to consider modifications in their evaluation system. Implications from this study can be used to help other administrators grappling with common concerns.

## DEDICATION

My wife and partner in life, Jayne Hartley: Your support, patience and encouragement throughout this project kept me marching along this path. You shared my vision to achieve this goal and were always generous in allowing me the time and space to get it done. We are ready to have fun again!

My mom, Xanthi Canelake: You always gave of yourself to help others, encouraged and believed in me from the beginning, and stayed strong through the best and worst of times. Your memory will always inspire me.

My dad, John G. Canelake: You are my role model for creativity, leadership, and respect. I've especially appreciated your encouragement in our daily phone calls over these last few years. You will always be my favorite hero, and you make me feel like one every day.

My three children, Sarina, Nick and Sam Canelake: You are all so different, yet all filled with a unique brand of promise and spirit. You make me proud and happy, and I look forward to seeing you take command of this world in ways I can't begin to imagine.

## ACKNOWLEDGMENTS

I am grateful to many individuals at Arizona State University, my school district, and school districts across the state for their time, talents and inspiration that helped inspire my work.

The greatest appreciation is reserved for my dissertation chair, Dr. Debby Zambo. Her incredible dedication to students has been unmatched in my many years as a university student. She kept me on track with detailed feedback and provocative questions that changed my perspective at every turn. Dr. Zambo helped me come to understand the power of establishing theoretical frames as the basis for my work decisions, and I have become a better practitioner by learning from her. Through her expert guidance, I learned to look a little closer, think a little more creatively, and never use the words “final draft” on any enterprise.

Special thanks also go to my committee members. Cory Hansen and Tacy Ashby. Their encouragement and feedback gave me new insights as a practitioner and a scholar. Dr. Hansen’s respectful and concise assistance helped me stay the course and find meaning in this project, and Dr. Ashby provided the unique perspective of an accomplished practitioner and a highly respected colleague in education leadership.

Finally, I have great appreciation and respect for my school district colleagues: superintendent, directors, and principals. They are bright, hard working leaders with a remarkable passion for their work. I especially appreciate their engagement and honesty during the course of this action research project. They rarely left me to wonder what they thought!

## TABLE OF CONTENTS

	Page
LIST OF TABLES.....	ix
LIST OF FIGURES .....	x
CHAPTER	
1 INTRODUCTION.....	1
Teacher Evaluation Reform in My District .....	6
Research Questions.....	13
2 LITERATURE REVIEW .....	16
Theoretical Frames .....	19
Change Theory: Diffusion of Innovations .....	20
Change Theory: Concerns-Based Adoption Model.....	22
Change Theory: Change Leadership.....	24
Learning Theory: Social Learning Theory .....	26
Vygotsky Space Learning Theory.....	28
Learning Theory: Communities of Practice.....	31
3 METHODS AND INTERVENTION.....	33
Methodology .....	33
Setting .....	35
Participants.....	36
Timetable and Action Plan .....	38
Measures: Data Collection .....	43
Measure 1: Surveys .....	43

CHAPTER	Page
Measure 2: Interviews .....	47
Measure 3: Meeting Records .....	47
Measure 4: Research Journal .....	48
4 ANALYSIS AND RESULTS.....	49
Results of the Quantitative AES Construct Survey .....	50
Reliability of AES Survey.....	51
Descriptive Statistics of Participants' Pre- and Post-Survey Responses.....	52
Comparison of Pre- and Post- Survey Response Means.....	53
Results from Analyses of Participant Characteristics and Survey Outcomes.....	56
Correlations Among the Four Constructs.....	58
Results of the Stages of Concern Questionnaire.....	59
Peak and Secondary Stage Score Interpretation.....	60
Profile Interpretation .....	62
Results of Qualitative Data Analysis .....	65
Participant Concerns.....	66
Participant Learning .....	76
Analysis of Leadership.....	83
5 FINDINGS.....	90
Data Triangulation .....	90
Validity of the Data .....	91

CHAPTER	Page
Changes in Administrators' Concerns: Research Question #1 .....	92
Concerns Frequent and Only Slightly Changed .....	92
Concerns About the Evaluation System .....	93
Concerns About Inter-Rater Reliability .....	94
Concerns About Level of Training .....	95
Teacher Evaluation Concerns About Alignment and Integration .....	96
Administrator Learning: Research Question #2 .....	98
How I Led This Process of Change: Research Question #3.....	104
6 CONCLUSION .....	108
Achieving the Purpose of This Action Research.....	109
Answering the Action Research Questions.....	111
New Learning About Implementing Teacher Evaluation .....	112
Leadership Insights .....	114
Limitations of This Action Research .....	116
Implications for Further Research.....	117
Closing Thoughts.....	119
REFERENCES .....	120
 APPENDIX	
A PROFESSIONAL TEACHING STANDARDS ADOPTED IN 2010 .....	126



APPENDIX	Page
B INSTITUTIONAL REVIEW BOARD APPROVAL .....	128
C PARTICIPANT CONSENT LETTER .....	130
D SURVEY INSTRUMENTS .....	132
E KEY INFORMANT INTERVIEW .....	142
F SoCQ QUESTIONNAIRE PEAK AND SECONDARY SCORES .....	145
G SoCQ INDIVIDUAL PROFILE LINE GRAPHS .....	147

LIST OF TABLES

Table	Page
1. Administrator Demographic Data .....	47
2. Coefficient-Alpha Estimates of Internal-Consistency Reliability for Teacher Evaluation Survey Instrument .....	61
3. Participant Survey Response Difference by Construct from Pre- Intervention Survey to Post-Intervention Survey .....	65
4. Participant Survey Response Difference by Particular Sub-Group ...	67
5. Teacher Evaluation Survey: Correlations Between Four Constructs on the Post-Survey .....	68

## LIST OF FIGURES

Figure	Page
1. Stages of concern about innovation: Definitions .....	32
2. Kotter's eight steps to leading change in context of our teacher evaluation system .....	34
3. Vygotsky Space theoretical lense applied to innovation .....	38
4. Outline of intervention activities .....	51
5. AES survey construct correlation to research questions .....	56
6. Frequencies for construct one: Teacher evaluation survey .....	62
7. Administrators' perception of teacher evaluation .....	63
8. SoCQ questionnaire peak and secondary scores on pre-and post-survey .....	71
9. Hypothesized Development of Stages of Concern .....	73
10. Stages of Concern profile for participant 1 .....	74
11. Overall frequency of participant concerns .....	76
12. Frequency of participant concerns pre- and post-intervention .....	76
13. Frequency of occurrence of Vygotsky Space learning demonstrations .....	86
14. Frequency of occurrence of leadership elements in qualitative data .	93
15. Research questions and data sources .....	100

## Chapter 1 Introduction

Since the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983), the school reform movement has offered many contrasting ideas and initiatives to make our nation's schools more effective and globally competitive. Reformers promoted the value of smaller schools and smaller class sizes, outcome-based education, top-down and bottom-up literacy instruction, back to basics, and an entire range of other programs and practices. The federal government entered into the reform dialog with the passage of No Child Left Behind (NCLB, 2001). NCLB steered the reform movement toward an intensified focus on accountability for student achievement at the school and district level. Districts were required to report disaggregated data on standardized achievement tests from each school. These annual reports held substantial consequences and became part of the public record for every public school and district. Educators responded by developing a keener focus on academic standards, standards-based instruction and data-driven instruction techniques. These efforts concentrated on having each disaggregated subgroup of students meet steadily increasing achievement goals.

As the pressure for accountability grew, reformers began to identify the importance of the teacher quality as a critical predictor of student achievement. Early on, the National Board for Professional Teaching Standards (NBPTS, <http://www.nbpts.org>) fueled the movement to heighten the importance of the individual teacher when it created a set of standards that outlined a description of

accomplished teaching practice, along with a certification system to assess and qualify teachers who reached this vision.

Charlotte Danielson (1996, 2007) led the development of standards-based teacher evaluation systems with her book *Enhancing Professional Practice: A Framework for Teaching* (1996). Danielson's work represented a milestone in the development of a standards-based teacher evaluation system, and over the next decade, many others joined in by developing similar tools for evaluating the effectiveness of teachers. As these advances were made, others began to study the usefulness and validity of these tools. Odden (2004) investigated the reliability and validity of the standards-based evaluation instruments developed by Danielson and others and sought to determine whether ratings from these instruments could reliably differentiate the effectiveness of teachers in relation to their ability to affect positive changes in student achievement. Odden (2004) found a positive correlation between teacher ratings on the new standards-based teacher evaluation instruments and increased student achievement gains. Odden's findings on the correlation between standards-based evaluation ratings and student outcomes in content areas were replicated (Borman & Kimball, 2005; Kimball, White, Milanowski, & Borman, 2004; Milanowski, 2004; Odden, Borman, & Fermanich, 2004) and generated greater interest in using teacher evaluation instruments as key facets of accountability systems.

Nye, Konstantopoulos, and Hedges (2004) released a study that provided additional evidence regarding the levels of influence individual teachers have on student achievement.

These findings would suggest that the difference in achievement gains between having a 25<sup>th</sup> percentile teacher (a not so effective teacher) and a 75<sup>th</sup> percentile teacher (an effective teacher) is over one-third of a standard deviation in reading and almost half a deviation (.48) in mathematics. Similarly, the difference in achievement gains between having and 50<sup>th</sup> percentile teacher (an average teacher) and a 90<sup>th</sup> percentile teacher (a very effective teacher) is about one-third of standard deviation (.33) in reading and somewhat smaller than half a standard deviation (.46) in mathematics... These effects are certainly large enough effect to have policy significance. (p. 253)

The Arizona State Legislature responded to the teacher quality movement when it passed Chapter 12 of Arizona House Bill 2011 (Arizona State Legislature, 2009). Although this legislation did not dramatically alter existing statutes requiring teacher evaluation, it had the effect of stimulating reform in the area of judging and acting on teacher quality by mandating a significant increase in administrative authority to make decisions based on teacher effectiveness. The legislation contained significant changes to statutes affecting personnel procedures by relaxing due process laws affecting teacher terminations and prohibiting the use of seniority or tenure when a reduction in force (RIF) was necessary. The legislation was perceived to significantly weaken the concept of tenure for experienced teachers and was met with considerable resistance and formal appeals from the Arizona Education Association. Challenges were voiced, but all aspects of the legislation were upheld.

The Bill prohibited districts from using seniority to select teachers for a RIF, and had the intended effect of driving school districts to seek more substantive methods of sorting teachers to determine eligibility for a RIF. Without seniority as the default measure, districts turned to their teacher evaluation measures. The typical evaluation instruments being used in most districts, however, predated the standards-based instruments pioneered by Danielson (1996), or were watered-down versions of the newer systems. Typical evaluation systems gave an overwhelming majority of teachers' ratings of adequate or proficient with no apparent regard for success in raising student achievement. Since virtually all teachers rated at a proficient level, it was clear that these evaluation systems did not differentiate adequately enough to rank teachers the way seniority numbers had in the past. The inability of teacher evaluations to discriminate effective from ineffective teachers is probably why teachers found evaluation instruments to be ritualistic and a waste of time (McLaughlin, 1990).

Chapter 12 of Arizona House Bill 2011 created an immediate need for school districts to adopt a more fine-grained means to measure the quality of teachers in their schools. In effect the State Legislature created a new requirement to refine district evaluation systems so they could stand up to the rigorous – and often legally scrutinized – task of standing up to challenges and sorting teachers to determine who retains a continuing employment contract. The call had been made for Arizona school systems to create far more sophisticated sets of instruments and procedures than most school districts had in place.

Feeling the urgency for reform in our state, Tom Pickrell, *General Counsel of Mesa Public Schools* addressed an audience of over 100 administrators at a January 2010 meeting of the Arizona School Personnel Administrator's Association. He asked whether anyone in the room felt confident that their current evaluation systems were strong enough to use for employment decisions. His question was met with general laughter and just one or two hands were raised. Clearly, the administrators in the room that preside over teacher evaluation in their districts did not think their evaluation systems could be used to differentiate among their current teachers for high-stakes employment decisions.

The message was reinforced and the stakes were raised even higher when on February 17, 2009, President Obama signed into law the American Recovery and Reinvestment Act of 2009 (ARRA). The ARRA provided \$4.35 billion for the Race to the Top (RTTP) competitive grants to the states, and cash-strapped states like Arizona took note. RTTP added a new dimension to the dialog on reform of teacher evaluation by requiring districts to have sophisticated standards-based evaluation systems, and further requiring that student achievement outcomes be included as an evaluation component for each teacher. This addition of student achievement data represented a politically sensitive step and an unprecedented shift toward increased accountability for individual teachers. Not only would this requirement have the potential to be controversial among teacher groups, it would stretch the resources and technical ability of many school districts to fairly and effectively measure student performance and tie the results to individual teachers (Braun, 2005).



Arizona made the commitment to participate in the RTTP but was not selected for the award in the first two rounds. Even so, districts remained committed to compliance with the U.S. Department of Education specifications including the newly required standards for teacher evaluation. Districts viewed these RTTT grant requirements as a reflection of the spirit of the new leadership of the U.S. Department of Education and a leading indicator of future expectations for teacher accountability. Even if districts did not choose to participate in RTTP, the Arizona Legislature moved in concert with the Federal government by enacting Senate Bill 1040 in the second session of 2010. This Bill stated that by the school year 2013 quantitative achievement data on student achievement gains must account for 35-50% of evaluation outcomes for both principals and teachers.

Together, these three measures, Arizona HB 2011, U.S. Department of Education RTTT and Arizona SB 1040 sent a clear directive to school districts in Arizona to take teacher evaluation to new levels of sophistication and accountability. Districts had little choice but to create evaluation systems that assured individual teacher performance was measured using standards-based methods, and integrate student achievement outcomes into those newly developed standards.

### **Teacher Evaluation Reform in My District**

I became a district administrator for my small urban district in 2009, just months before HB 2010 was passed. My district consists of five campuses set in a six square mile neighborhood filled with a vigorous retail, office, and

manufacturing infrastructure. Although there are several high-end residential condominium communities, a majority of families that attend our schools live in poverty. Over 86% participate in the National Free and Reduced Lunch program and many are refugees from Mexico or Somalia. Over 31% of the students are second-language learners. There is also a citywide center for homeless families situated in the district and many of these children are enrolled in district schools. The 172 teachers employed for the district have an average experience of ten years. Over 33% of teachers have five or less years of experience and 15% have 20 or more years of experience. A majority (85%) of teachers are Caucasian and 67% are Caucasian female (Internal Equal Employment Opportunity data; School District Employment Report, 2011).

My administrative responsibilities include supervision of teacher evaluation and responsibility to ensure that teachers are evaluated effectively and in a manner that supports the teaching and learning mission of the school district. Upon my arrival, principals and other administrators made me aware that the district's evaluation systems were not highly regarded. The evaluation tools were overly simplistic and not adequately descriptive to inform plans for professional growth and development.

In my administrative role in the district, and as a doctoral student, I conducted a first cycle of action research in an effort to remedy the limitations of our evaluation systems. This study began with a set of surveys and the formation of a district Teacher Evaluation System (TES) committee. Group membership included all principals, most district office administrators and a representative

group of teachers from each school. Initial surveys were sent to all 172 teachers with responses returned by 98. All nine of the targeted administrators participated in the survey. The TES committee studied the survey results and learned that teachers were more ambivalent in their opinions about the instrument than administrators. They were initially much more likely to favor keeping the current system, while administrators were nearly unanimous in calling for a total overhaul. Even though teachers did not see an immediate need for change, they expressed a general dissatisfaction with the teacher evaluation, stating that it had little correlation to the realities of teachers' jobs in high-needs schools. Teachers reported that the current system was superficial and for the most part too easy on teachers. One said, "I think most teachers are happy that it's over and don't take much from the process." Many commented that teachers don't often reflect on the evaluations after their completion and noted that evaluation should entail more than a single classroom observation.

The TES committee began by reviewing summaries of the surveys I generated. The group met formally on nine occasions and engaged in dozens of online conversations between meetings. Initial movement toward a complete revamping our evaluation systems proceeded as administrators and teachers became exposed to more sophisticated standards-based teacher evaluation systems. The group looked at examples of teacher standards and evaluation systems from various sources published nationally, and several others developed by Arizona school districts. This exposure to new teaching standards and standards-based evaluations prompted the group to agree that our current

evaluation tools lacked sufficient scope and sensitivity to guide the development of a modern urban teacher.

Both teachers and administrators found the indicators too simplistic and lacking in adequate scope to measure the expectations for teaching in our high-needs urban schools. Committee participants also noted that overall the new teaching standards were far more congruent with recent system-wide professional development initiatives focused on improving instructional design and pedagogy. They noted that the standards-based evaluation systems were far more reflective of our current practices. Additionally, committee members looked back at our existing evaluation system and expressed a new level of concern that there were no descriptive statements associated with the evaluation indicators. After exposure to standards-based instruments, they could see that each indicator of our old system was merely a simple statement that could be open to wide latitude of interpretations. The indicators lacked a detailed set of rubrics with specific descriptions of behaviors, knowledge and job performance necessary to meet stated expectations.

Principals and teachers were also dissatisfied with the binary levels of rating for each indicator. They could choose “meets expectations (ME)” or “below expectations (BE)” with no ranking in between. In practice, the BE rating was rarely used, and saved almost exclusively for instances where a teacher was being coached out of the profession or threatened with dismissal for poor teaching performance. As a result principals expressed that they had little choice but to rate almost every teacher with all ME’s, or be put in a position to move toward

dismissal. There was little middle ground and the only place to detail greater insight was with the evaluator's summary comment box at the end of the evaluation. In essence, administrators agreed that the evaluation system currently in place was a very poor tool to guide or encourage improvement in teacher quality.

Using these insights, our group reviewed the current evaluation systems against seven standards-based options that were in use either locally or nationally. Our committee decided to completely abandon our existing evaluation system and embrace a completely new option. We narrowed our choices to several standards-based systems and eventually came to favor a complete package of teacher evaluation tools developed by a nationally recognized educational service agency. The package of evaluation tools that was selected appealed to teachers and administrators because it had a positive tone that encouraged teachers and principals to reflect on attainable increments of teacher competencies. It had undergone extensive field-testing statewide in the southeastern U.S. They also felt the outcome of the evaluation cycles could be used to create meaningful and productive professional development plans for groups and individuals. The committee brought a recommendation to the Governing Board and gained approval to use the new teacher evaluation system in April 2010 for the 2010-2011 school year.

Preparations were made to put our new system in effect district-wide. A copy of the professional teaching standards, as adapted for my school district, can be seen in Appendix A. Forty district teachers and administrators participated in

training for the new system, and as they learned more, they began to anticipate the significant impact this level of change would have on both evaluators and evaluatees. Teachers outside of our initial adoption committee learned this new system required far more engagement from individual teachers. Rather than simply being observed and evaluated once or twice annually, teachers were now expected to complete a detailed self-assessment, analyze their professional growth needs, and create a professional development plan. Most importantly, teachers became aware that the evaluation instruments would be used to rate each aspect of their professional practice. They learned that although they could become engaged and in control of some aspects of the evaluation process, not all aspects of their evaluation were in their hands. In the end our evaluators would have the power to create a summative annual evaluation that could have a significant effect on their employment with the school district.

Administrators also faced adoption challenges with the new evaluation system. They moved from a simplistic checklist rating system that could be completed quickly and easily, to a comprehensive standards-based assessment that took much time and thought. Our evaluations could no longer be done quickly or in isolation. They required extensive communication between administrator and teachers on issues of self-assessment and professional development. Rather than a simple classroom observation and write-up, principals were now required to review and evaluate multiple measures of teacher performance, compare these against the teacher's self-perceived traits in each area, and then assist each teacher in developing a coordinated professional

development plan designed to improve performance in specific areas tied to this new evaluation system.

As a district leader I became concerned about whether our administrators—district administrators and principals – would be able to sustain the necessary energy and focus needed to ensure a successful implementation. The system was launched with a high level of interest and enthusiasm for the new innovation, but to be adopted with full integrity would require sustained involvement from both administrators and teachers. No matter how committed our principals and teachers were to this new process, they faced many competing interests for their time and focus. Complicating matters, recent budget cuts caused principals to begin the 2010-11 school year with fewer individuals to assist with teacher coaching and instructional leadership. Additionally, four of the five principals no longer had assistant principals and had to absorb additional management duties on a day-to-day basis. Principals could no longer delegate teacher evaluation tasks to assistants or coaches as much as they had in the past. Additionally, the teacher evaluation adoption was not the only major initiative taking the attention of teachers and administrators. The district was awarded two school improvement grants, one for an individual campus and another for the entire district. Each of these grants represented opportunities for far-reaching reforms affecting curriculum and instruction. These initiatives required the energy and attention of already busy principals and teachers be diverted, at least to some extent, from our new evaluation system. Initiation of these reforms in the midst of the adoption of

a new teacher evaluation system also raised the potential of overlapping or even conflicting reform agendas.

With a high level of activity in the district, an ample set of initiatives on the table, and diminished resources at hand, this adoption became a major challenge that led to many concerns among the administrative team. My focus for the 2010-11 school year was to ensure teachers and principals had sufficient and high-quality training to become comfortable and proficient using and managing our new evaluation system. Although our adoption was monitored and managed carefully, principals did not always have positive experiences using the system. They had the opportunity to develop a sense of the reflection, pacing, and effort needed to get through a yearlong cycle, but experienced challenges adapting to this more demanding system of evaluation. Adoption concerns among our administrative team remained high and there was great potential for growth coming into the second year using our new evaluation and professional growth tools.

### **Research Questions**

I knew it was essential to take a leadership role to plan and implement strategies to ensure our new evaluation tools were used fairly, consistently, and effectively. My primary sphere of influence was with the six building-level and three district-level administrators, all of whom were members of an existing community of practice (CoP) and key voices to understand the challenges and potential success of using standards-based teacher evaluation. My primary



objective was to sustain and enhance the CoP with the administrative team to help them become more effective, fair, and consistent evaluators.

As I pursued this end I knew that it would challenge my leadership skills to effectively implement this learning project. I had to work closely with my superintendent and my district administrator peers to balance all of the competing interests and cooperatively develop a plan to fit this professional development in with the rest of what we were trying to accomplish within our district. It was difficult to prioritize this effort to ensure this initiative got the necessary attention to help our administrators learn and use this new evaluation system in a way that helps teachers improve their practice.

As I pursued this innovation I sought to understand more about how our administrators learn and use our new teacher evaluation tools, all the while seeking answers to the following research questions:

1. What concerns do administrators within my district have about effective implementation of our new teacher evaluation system, and how and to what extent will those concerns change as our CoP learns, practices, and engages in discourse?
2. What will administrators say, do, and feel as our CoP attempts to:
  - 1) develop a common understanding of the professional teaching standards;
  - 2) increase inter-rater reliability on teacher evaluation instruments;

- 3) understand the purpose of our teacher evaluation system to improve professional practice;

**3.** How do I lead this process of change?

## Chapter 2 Literature Review

Teacher quality has been cited as the most important single factor in determining student success (Odden, 2004; Nye, Konstantopoulos, & Hedges, 2004; Borman & Kimball, 2005; Kimball et al., 2004; Milanowski, 2004; and Odden et al., 2004). Further, a growing body of scholarly literature indicates that quality instruction is clearly connected to improvement in student learning (Darling-Hammond, 2000; Gamoran, Porter, Smithson, & White, 1997; Sanders & Horn, 1998; Westbury, 1993). With evidence that teacher quality is a primary determinant of educational outcomes, teacher evaluation has gained prominence among strategies to reform education (Danielson 1996).

Until recently, however, teacher evaluation has not been reviewed favorably in the literature. In 1987, Medley and Coker found that traditional evaluations by principals were not effective enough to truly differentiate between levels of proficiency among teachers. Not surprisingly, they found these same results of little use as a basis to help teachers improve instruction. In 1990, McLaughlin found that many teachers describe evaluations as ritualistic and largely a waste of time. Administrators seeking to use teacher evaluation as a means to improve student achievement will note the finding by Linda Darling-Hammond (1990) that many evaluations had very little influence on personnel, staff development, or instructional methods. Danielson (1996) found that traditional evaluations of teachers by principals were inadequate in determining teacher proficiency or as a tool for helping guide improvements in instructional skills. More recently, Peterson and Peterson (2006) found that most

evaluation protocols usually lacked connection to identified teaching standards and scoring rubrics. By design they tended to have limited means to collect useful data, and are hampered by idiosyncratic interpretation of performance areas by those doing the evaluations.

This dismal look at teacher evaluation is strongly contrasted by literature focused on newer models of standards-based teacher evaluation. Charlotte Danielson's, *Enhancing Professional Practice: A Framework for Teaching* (1996) represents a beginning to create a positive connection between what teachers know and can do and student achievement outcomes. As a foundation to understanding the potential benefits of this new generation of teacher evaluation systems, Milanowski and Kimball (2003) posit that standards-based teacher evaluation systems promote a common conception of good teaching and act as a performance competency measure useful to identify how to improve instruction, affect teacher selection and retention, and guide teachers to improve their skills. Odden (2004) studied examples where evaluations systems with criterion-validity results strong enough to indicate that higher teacher evaluation scores positively correlate with increased learning gains by students. Odden suggests that these systems are “good enough to use for consequential decisions such as pay increases” and goes on to caution that “results used to trigger pay increases should be fully understood by teachers, produce reliable results across multiple assessors, and be valid—that is, have positive linkages between evaluation scores and value-added student learning” (p. 130)

Gallagher's study (2004) was conducted at a school with a contemporary standards-based evaluation system designed to improve instructional practice and enhance student learning. Gallagher found significant variation in achievement attributable as classroom level effects, and that the "teacher evaluation system had a statistically significant relationship to classroom effects, that is, value-added learning growth" (p. 100). Gallagher's study demonstrated an instance where an enhanced standards-based evaluation system produced results that correlated a teacher's ability with student achievement.

Overall, this emerging body of literature focused on standards-based evaluation systems is positive and gaining momentum. The effectiveness of evaluations to identify and reliably measure teacher qualities and behaviors that positively affect student outcomes has been shown. Successful implementation, however, is not automatic with the adoption of a teacher evaluation system and depends on effective leaders with knowledge and skills about teaching and evaluation (Danielson 1996, 2006, 2010). The research of Davis, Ellett and Annunziata (2002) demonstrated that school leadership is essential to utilize teacher evaluation as

A vehicle to improve teaching and learning... In the extreme case, leadership makes the difference between perfunctory summative teacher evaluation and meaningful assessment of the teaching and learning process that has the potential to enhance the quality of teaching and student learning. (p. 288)

Danielson (2010) states,

A credible system of teacher evaluation requires higher levels of proficiency of evaluators than the old checklist, “drive-by” observation model. Evaluators need to be able to assess accurately, provide meaningful feedback, and engage teachers in productive conversations about practice.

(p. 39)

Danielson goes on to outline four steps for training evaluators:

- 1) Familiarization with the instruments and associated teaching standards,
- 2) Recognition of sources of evidence for each standard, 3) Learning how to interpret evidence against the instruments’ rubrics, and 4) Learning how to calibrate their judgments against those of their colleagues.

### **Theoretical Frames**

The theoretical frames that guided my work focused on change and learning. My first frame addressed the change process, organizational change, and change leadership. Change theories guided me to understand how administrators might accept and embrace new thinking in the area of teacher evaluation. The second set of theoretical frames focused on explorations of sociocultural learning. I looked at social learning theory to better understand the learning processes that would take place among the administrative team as they worked together to create common understandings and common practices for using the evaluation system. I examined principles of sociocultural learning and looked specifically at the framework of Vygotsky Space to better plan for and understand how the administrators would transfer their learning from a social setting to their work context. I also reviewed sociocultural learning theories applied to communities of

practice (CoPs) to better understand how the dynamics of this group could be optimized for learning.

### **Change Theory: Diffusion of Innovations**

Adoption of a standards-based evaluation system requires an ongoing change in practice for our administrative team. Davis, Ellett and Annunziata (2002) recognize that this task requires a direct focus on the teaching and learning process and a major shift from typical checklist evaluations. Everett M. Rogers' book, *Diffusion of Innovations* (1995), outlines several theoretical perspectives about the concept of the diffusion of change. Diffusion is defined as the process of adoption of an innovation within an organization or community. Rogers outlines four influencing factors for the adoption of an innovation. The first factor is the innovation itself, the second factor is the communication channels used to disseminate information about the innovation, the third is time, and the fourth is the nature of the society in which the innovation is introduced (Rogers, 1995). Rogers (1995) outlines four theories related to the diffusion of innovations. These are 1) innovation-decision process theory, 2) the individual innovativeness theory, 3) the rate of adoption theory, and 4) the theory of perceived attributes.

The innovation-decision process theory describes five stages. The first stage is knowledge. Rogers' theory states that future adopters must first learn about the innovation. Next, they would be persuaded about the qualities of the innovation. Third, they must make the decision to adopt the innovation. Fourth, once adopted, they must actually act to implement the innovation. Fifth and

finally, the implementers must confirm they made the right decision. Upon achievement of these stages, diffusion results (Rogers, 1995).

The individual innovativeness theory addresses who adopts the innovation and at what time they adopt. This effect is usually plotted on a bell-shaped curve showing the categories of adopters of an innovation. The first category is the risk-taker innovators (2.5%) who are seen as the pioneers leading the way. The second category is the early adopters (13.5%) who spread the word to others. The third and fourth categories are the early majority (34%) and late majority (34%) of those adopting the innovation. The role of the innovators and early adopters is to communicate positively to the early majority. The late majority, as the name implies, waits to be sure there is little risk in adopting. The final group, the laggards, represents the last 16%. The laggards are highly skeptical and resist the change to the extent that many never adopt the innovation (Rogers, 1995).

The theory of rate of adoption illustrates adoption of innovations with an s-curve on a graph. The theory states that the adoption will grow slowly in the beginning, followed by an accelerated period of growth that tapers, becomes stable, and eventually declines (Rogers, 1995).

The theory of perceived attributes postulates that individuals become more likely to adopt an innovation when they perceive the following five ordered attributes:

1. The innovation demonstrates an advantage over the status quo or a competing innovation.



2. The innovation is perceived to be compatible with existing values and practices.
3. The innovation is not overly complex.
4. The innovation has trialability (meaning that it can be tested for a specified time before full adoption.)
5. The innovation must present concrete, observable results (Rogers, 1995).

### **Change Theory: Concerns-Based Adoption Model**

The Concerns-Based Adoption Model (CBAM), developed over 30 years ago by Hall, Wallace, and Dossett (1973), is a theory to help researchers understand concerns about how individuals accept change when implementing new innovations. The theory develops a means to gather information on participants' levels of adoption to continuously facilitate the change process (George, Hall & Stiegelbauer, 2008). Frances Fuller (1969) worked with student-teachers and recorded an apparent progression of implementation concerns that moved developmentally from completely unrelated concerns, to more simple concerns about self, to concerns about completing tasks, and finally to concerns about what impact they may have. The CBAM grew out of Fuller's work and described the developmental Stages of Concern (SoC) in greater detail. The development of the CBAM identified and confirmed a series of seven developmental SoC about any innovation (Hall, 1979; Hall & Hord, 1987, 2006; James, 1981). Fuller's developmental stages can be seen compared to the current seven SoC in Figure 1 below:

Fuller's Stages of Concern	Seven Stages of Concern	Definitions of Stages of Concern by Focus of Participants
Impact	Refocusing	Focus on exploring more universal benefits from innovation, including possible major changes or replacement with more powerful alternative. Individual has definite ideas about alternatives to the proposed or existing innovation.
	Collaboration	Focus on coordination and cooperation with others regarding use of innovation.
	Consequence	Focus on impact of the innovation on "clients" in the immediate sphere of influence.
Task	Management	Focus on the processes and tasks of using the innovation and best use of the information and resources. Issues related to efficiency, organizing, managing, scheduling, and time demands are utmost.
Self	Personal	Individual uncertain about demands of innovation, inadequacy to meet demands, and role with the innovation. This includes analysis of role in relation to the reward structure of the organization, decision-making, and consideration of potential conflicts with existing structures or personal commitment. Financial or status implications of the program self and colleagues may also be reflected.
	Informational	General awareness of the innovation and interest in learning more detail about it is indicated. The person seems to be unworried about self in relation to innovation. She/he is interested in substantive aspects of innovation in selfless manner, such as general characteristics, effects, and requirements for use.
Unrelated	Unconcerned	Little concern about or involvement with the innovation indicated. Concern about other things is more intense.

*Figure 1.* Stages of Concern about Innovation: definitions. Adapted from George, Hall, & Stiegelbauer, 2006.

The CBAM, described by Hall (1979) and others (Hall & Hord, 1987, 2006; James, 1981), has become well-known as a tool to measure the extent to which a program has been implemented (James, 1981), and has become a trustworthy model for examining change in organizations (Kelly & Staver, 2005). Hall and Hord (2006) view the CBAM in terms of a means for leaders to

strategically guide the change process and direct leadership of professional development when adopting new practices.

### **Change Theory: Change Leadership**

John Kotter's, *Leading Change* (1996) provides an additional theoretical perspective about the process of adopting change in an organization and what effect leaders can have on promoting successful change initiatives. His work outlines eight steps to leading change that are relevant to my leadership role in facilitating this change for our district. The steps that he outlined were used in planning the learning activities with the administrators. Figure 2, below, illustrates the main concepts of Kotter's eight steps and is referenced with demonstrations of this theory into the current adoption and goals for continued application of his theory.

Step	Fundamental Idea	Applied Strategies
<b>Develop a Sense of Urgency</b>	Capture the attention and interest of key stakeholders.	This strong sense of urgency occurred naturally with the government interventions (House Bill 2010, Senate Bill 1040, and Race to the Top). I continued to publicize at superintendent and district administrative levels by making this a recurring agenda item. I also sent several all-user blast emails to update progress to all teachers and support staff.
<b>Develop a Guiding Coalition</b>	Gather a group of key stakeholders to lead change.	A teacher evaluation committee was convened in first round of action research and consisted of teachers, principals and district administrators. This committee was called again to help advise in the transition to compliance with SB 1040. An administrative committee was also brought together beginning in August to learn more about teacher evaluation and state requirements.
<b>Create a Vision for Change</b>	Determine and develop a short summary of values central to the change to focus the future. Create a strategy to execute that vision.	A district leadership team developed vision statement and key talking points for adoption of new evaluation instruments in first round of action research. This team revisited the topic and refined the message based on the context of new directions from SB 1040.
<b>Communicate the Vision</b>	Create opportunities to talk about change vision. Address peoples' concerns and anxieties. Apply vision to all aspects of operations.	A vision statement and talking points were integrated into Board presentations in March, August and November. These were also included in summer mailings, superintendent presentations and all-user blast emails.
<b>Remove Obstacles</b>	Continually check for barriers to change and remove obstacles to empower people to execute vision.	Online implementation of the evaluation tools was problematic and required constant attention. Principals complained about the workload and some minor adjustments were made to ease their concerns in this area. There were also unintended consequences of defensiveness and mistrust. Additionally the administrative teacher evaluation meetings were framed as a support group for the team to share concerns and collectively mitigate problem areas.
<b>Create Short-term Wins</b>	Create achievable short-term targets with little room for failure.	Positive feedback from teachers was highlighted at administrative meetings. Usefulness of evaluation tools in helping RIF process was emphasized for administrators. Administrators were scheduled to attend regional administrator conferences and presented on our innovations. Comprehensive professional development plans created by administrators and teachers were highlighted at administrative meetings.
<b>Build on the Change</b>	Use each success as opportunity to build on what went right, to expand and improve.	Begin to integrate new teacher evaluation tools with district pedagogical training initiatives. Continued to find ways to integrate student achievement outcomes into evaluation as possible way to strengthen integrity of system and meet SB1040 requirements. Continue to integrate teaching standards into teacher development training
<b>Anchor the Changes in Culture</b>	Make continuous efforts to ensure change is seen in every aspect of organization for solid place in your organization's culture.	The district teaching standards (upon which the TES is based) have been published. TES was used as measure for performance pay and reduction in force rubrics. Posters of the district teaching standards were published. Plans made to integrate standards into cultural norms established in new staff handbooks.

Figure 2. Kotter's eight steps to leading change in context of our teacher evaluation system

Much of Kotter's strategy for success in adopting organizational change relies on communication. Schlechty (2009) reinforces the importance of communication when he recommends that leaders "must be adept at painting vivid word pictures..." and goes on to say that "This requires that they learn to think metaphorically as well as systemically... and tell compelling stories"

(p. 210). Senge (1990) emphasizes the need to build a shared vision, and more importantly, the capacity in the organization to share a vision of the future. This strategy seeks to develop a vision that all members embrace and is not anchored in one charismatic leader. The vision and strategy must sustain itself even when current leaders leave and be designed to ensure that universal commitment is the end goal in sight.

### **Learning Theory: Social Learning Theory**

Rogoff, Matusov, and White (1996) state that “coherent patterns of instructional practices are based on instructional models, and instructional models are based on theoretical perspectives on learning” (p.389). Theories about teaching and learning should inform planning and decision making, yet according to Wilhelm, Baker and Dube (2001), “these theories are typically under-articulated, unrecognized, underspecified, and quite often inconsistent.” (p. 1). Application of principles of social learning theory to this study allowed a more focused perspective on what was to be accomplished with the learning planned in this social setting.

Social Learning Theory is based the work of Russian psychologist Lev Vygotsky (1896-1934). Vygotsky’s (1978) work is one of the foundations of constructivist theory of learning. Constructivism is defined by Learning Theories Knowledgebase (<http://www.learning-theories.com/> 2011) as:

A reaction to didactic approaches such as behaviorism and programmed instruction, constructivism states that learning is an active, contextualized process of constructing knowledge rather than acquiring it. Knowledge is

constructed based on personal experiences and hypotheses of the environment. Learners continuously test these hypotheses through social negotiation. Each person has a different interpretation and construction of knowledge process. The learner is not a blank slate (*tabula rasa*) but brings past experiences and cultural factors to a situation.

In contrast to Jean Piaget's belief that cognitive development necessarily precedes learning, Vygotsky's theory asserts that social interaction plays a fundamental role in the process of cognitive development (Vygotsky, 1978; Daniels, 2001). Vygotsky states, "Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (interpsychological) and then inside the child (intrapsychological)" (1978, p. 57). Vygotsky's social learning theory reveals the importance of interactions between people and the sociocultural context in which they share their experiences (Crawford, 1996). Key to experience in this social context is dialog among individuals. He saw this learning as beginning with external discourse and then becoming internal.

Vygotsky's social learning theory reveals the importance of interactions between people and the sociocultural context in which they share their experiences (Crawford, 1996). The key to experience in this social context is dialog among individuals occurring as external discourse and then becoming internal. Vygotsky's theory promotes dynamic learning contexts in which students play an active role in learning. Vygotsky notes the importance of active learning in contexts that allow students to take an active role in their own learning and the

learning of others. In these contexts, the roles of the teacher and student are shifted. The teacher collaborates with learners to help individuals construct meaning from the experience in a reciprocal experience for both students and teacher (Wertsch & Sohmer, 1995).

### **Vygotsky Space Learning Theory**

Vygotsky Space is based on this sociocultural view of learning initially conceptualized by Vygotsky (1978) and then further developed by Harre (1984) and Gavelek and Raphael (1996). Kong and Pearson (2003) state a fundamental idea for this theory by saying, “psychological processes originate in purposive social interactions among human beings within an environment in which cultural tools and artifacts are present”(p. 2). Kong & Pearson (2003) go on to state that learning occurs when individuals interact with more and less knowledgeable individuals in a variety of social contexts.

The Vygotsky Space metaphor developed by Harre (1984) uses these concepts to frame learning experiences along the continuums of *public-to-private displays of learning* and *collective-to-individual learning activities*. In the public to private continuum, learning performances range from being observable to unobservable and the collective to individual continuum reflects the extent to which learners “either use the knowledge learned from others or make what was learned their own” (Gavelek & Raphael, 1996, p. 187).

Figure 3 is a representation of Vygotsky Space, and is based on the work of Harre(1984) and McVee, Gavelek, and Dunsmore (2005). The illustration conceptualizes four quadrants through which learners are theorized to recursively

cycle through. In the context of my study, Quadrant I (Public Setting/Collective Learning) is where more or less formally scheduled presentations occurred with the administrative team. Quadrant II (Private Setting/Individual Learning) is the space where administrators conceptualized the presentations and began to create meaning. Quadrant III (Private Setting/Individual Learning) is where administrators worked in their setting, more or less independently, on going through real or mock problems associated with evaluation. Finally, Quadrant IV (Public Setting/Individual Learning) is the space where administrators demonstrated learning in their actual practice.

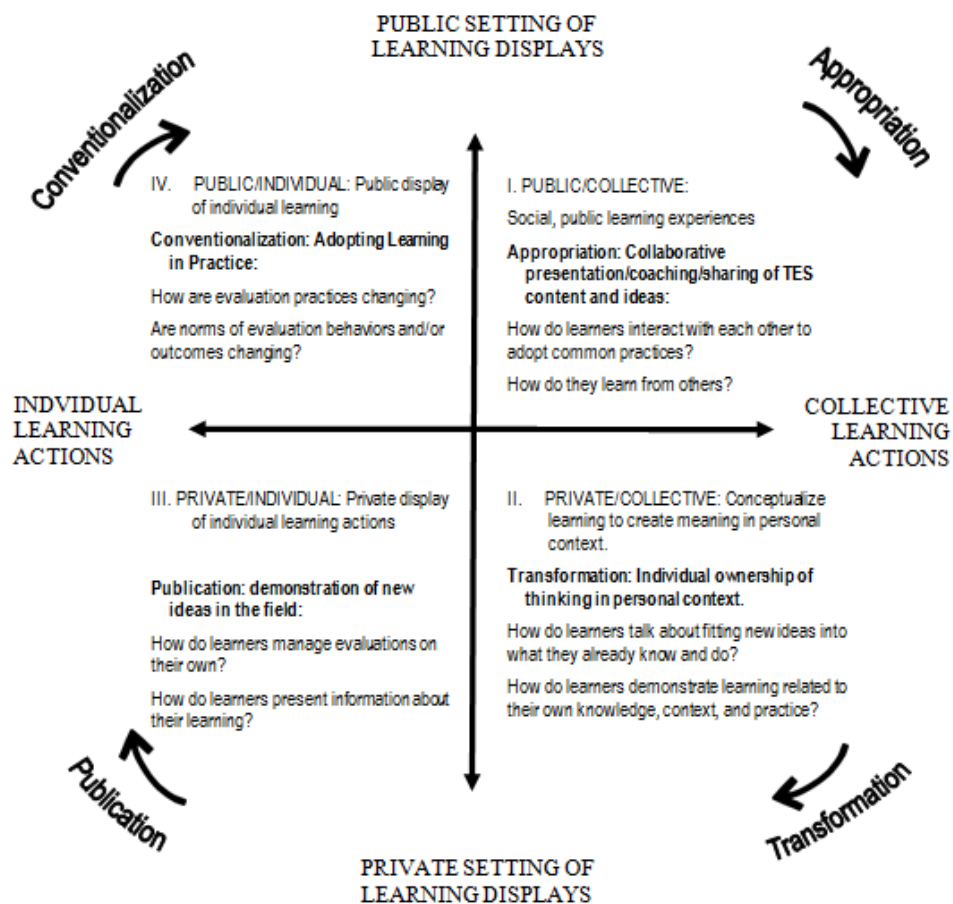


Figure 3. Vygotsky Space theoretical lens applied to innovation



Gallucci, Van Lare, Yoon and Boatright (2010) comment on the Vygotsky Space, “The process is cyclical and evolutionary, in the sense that learning and change operate in a cumulative and reciprocal way at both individual and collective levels.” (p. 925). She notes that the learner may at any time be functioning in any of the quadrants and that the essential focus is on the reiterative transition between quadrants. The theory describes recursive cycles of appropriation, transformation, publication, and conventionalization as learners transition through four quadrants of learning. The following bulleted description is adapted from Gallucci (2008, p. 549).

- *Appropriation*: Individual appropriation of particular ways of thinking through interaction and sharing of knowledge in a social setting represented in transition from Quadrant I to II.
- *Transformation*: Individual process of transformation and ownership of that thinking in the tangible context of one’s own work, represented in transition from Quadrant II to III.
- *Publication*: Publication of tangible new learning through talk or by practicing new ideas in their work, represented in transition from Quadrant III to IV.
- *Conventionalization*: The process whereby those public acts become conventionalized in the practice of that individual, adopted in the work of others, or both.

Vygotsky’s Space was used as a model for planning collaborative learning opportunities for the administrative team, and tracking their progress as they

grappled with this change and worked to become more proficient in their use of the teacher evaluation tools.

### **Learning Theory: Communities of Practice**

The concept of communities of practice (CoP) is another social learning theory that was useful in planning for this intervention with district administrators. Wenger, McDermott and Snyder (2002) define a community of practice (CoP) in its essence as “groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis” (p.4). In these CoPs members explore, expand, create, and share knowledge to pursue their joint enterprise.

Wenger et al. (2002) describe three elements of CoPs: domain, community, and shared practice. The element domain describes shared interests within the group and is the essential purpose of the community's interactions. The element of community describes the CoP's culture that develops positive and trusting relationships among CoP members. The third element, shared practice, describes the CoP's shared information, ideas, beliefs and tools that act to advance the knowledge base of the members.

The concept of situated learning developed by Lave and Wenger (1991) is also relevant to organizational change. The authors describe learning as "legitimate peripheral participation in communities of practice" (p. 31) and participating in a social world where one is almost continuously learning from experiences, actions, and connections to sources of knowledge. The authors state,

Moving toward full participation in practice involves not just a greater commitment of time, intensified effort, more and broader responsibilities within the community, and more difficult and risky tasks, but, more significantly, and increasing sense of identity as a master practitioner.

(p. 111)

Wenger et al. (2002) state “organizations need to cultivate communities of practice actively and systematically, for their benefit as well as the benefit of the members and communities themselves” (p. 12). They go on to say that organizations can’t force a true CoP, but that they can create conditions that help them to grow. Their seven design principles “are not recipes, but rather embody our understanding of how elements of design work together” (p. 51). These principles are the foundations for the enhancement of the CoP with the five principals and will be relevant in different ways.

### **Chapter 3 Methods and Intervention**

In the previous chapters I introduced the topic of change in the field of teacher evaluation and the growing sense of urgency to change how teacher evaluation is conducted in my school district. I outlined the specific challenges I planned to address and framed a set of research questions for this study. I reviewed current literature about teacher evaluation, as well as theoretical frames that would be relevant to the intervention that I planned. In this chapter I will describe the methodology of the research, as well as the intervention that was planned and executed in an effort to improve practice in my school district.

#### **Methodology**

This mixed methods study was conducted in an action research framework (Kochendorfer, 1997; Hinchey, 2008; Stringer, 2007; Mills, 2007). Action research, as defined by Mills (2007) is, “any systematic inquiry conducted by teacher researchers, principals, school counselors, or other stake holders in the teaching/learning environment to gather information about how well their students learn” (p.5). Kochendorfer (1997) identified several reasons action research is performed. These include: a) changing practice, b) creating new understandings, c) developing new relationships, and d) seeking answers to problems. Insiders conduct action research to make things better and close the theory/practice divide (Hinchey, 2008; Stringer, 2007). In my position as a district administrator, I played the role of action researcher, operating on the inside of the workings of my school district. This study developed as a systematic inquiry into how our organization grappled with change to make sense of our practice with the

new evaluation system. More specifically, I addressed the effect of a series of professional development activities created for this district's administrative team and focused on our newly adopted teacher evaluation system.

I used a triangulation mixed methods design in which different but complementary data were collected on the same focus (Creswell & Plano Clark, 2006; Johnson & Onwuegbuzie, 2004). In my study, research methods were used to investigate how principals traversed through Vygotsky's Space as they learned and implemented our new evaluation system. Vygotsky's theory predicted that as principals learned about, and practiced using the teacher evaluation tools they would move along a continuum of learning that would take them through public-to-private and individual-to-social space. Theoretically, this movement would positively affect their learning and make it conventionalized, or their own (Gavelek & Raphael, 1996).

To better understand administrators' perceptions and skills with our evaluation tools, I arranged multiple occasions to join individual administrators in the process of conducting brief teacher observations. Independently of one another, we rated teachers on scales from our teacher evaluation instruments, and shortly afterward came together to informally compare the degree of similarity and variation between our ratings. We repeated this process as we traveled through five or six classrooms per visit and repeated this process on several occasions. Qualitative data from these comparative observations were captured in the form of notes in my research journal and were used to help inform understanding of administrator learning and perception of inter-rater reliability.

Additionally, since the adoption of the new evaluation tools appeared to have created multiple concerns among administrators, I sought to understand what current concerns the administrators had and if they changed as a result of the work of our CoP. To help answer this question I used the Stages of Concern Questionnaire (SoCQ) developed by the Research and Development Center for Teacher Education at the University of Texas at Austin to track administrators' concern profiles over the time of this study (Hall et al., 1973). I also used a custom-designed Administrator Evaluation Survey (AES) that I developed and field-tested in February 2011. The AES served as a means to get more specific understanding of administrators' concerns about their common understanding of the professional teaching standards and their ability to use the teacher evaluation to improve professional practice.

I used a mixed methods design because of the complexity of the questions I sought to answer (Stringer, 2007; Fraenkel & Wallen, 2006; Wooley, 2009). Quantitative and qualitative data were brought together to take advantage of the strengths of both and to compare and potentially validate results.

### **Setting**

The district where this study took place consists of five campuses set in a six square mile neighborhood. A majority of families live in poverty as measured by 86% participation in the subsidized lunch program. The student population is 69% Hispanic, 8% Caucasian, 15% Black, 2% Asian, and 5% Native American. There is also a citywide center for homeless families situated in the district and many of these children are enrolled in district schools. There are 172 teachers

employed for the district with an average total teaching experience of ten years. One-third of the teachers have five or less years of experience, and 15% have 20 or more years of experience. A majority (85%) of teachers is Caucasian and 67% are Caucasian female.

### **Participants**

The participants for this study were nine administrators, each of whom were involved in teacher evaluation for the district. The group consisted of five principals, one assistant principal and three district office administrators (a director of curriculum, student services and instructional technology). Study participants were selected as a purposive sample (Miles & Huberman, 1994). Their selection was made because these individuals, in their collective role as the administrator team, were the entire group of administrators learning about and adapting to our new standards-based evaluation system for teachers. The professional development activities were designed specifically to assist these individuals as they grappled with the challenges of adopting this new system. Each was invited to participate in this study and each confirmed his or her willingness by signing a consent form approved by my university's Institutional Review Board (Appendix C).

Five principals and one assistant principal were key informants on the topic of teacher evaluation and were the core participants. Even though the directors evaluated only a handful of teachers, they played key roles in planning and directing teacher evaluation activities and were included in the activities and this study because of their important role in guiding the process of adopting the

teacher evaluation system. Of the six principals, three are Caucasian males, one is a Hispanic male and two are Caucasian females. Two of the six principals were once teachers in the district and the three directors were brought in from outside the school district three years ago. They include two female Caucasians and a Hispanic female.

Director participants average nine years of classroom experience and six years experience in leadership roles where they have evaluated teachers. Three of the six principals have been in the district from four to 13 years, and the other three were just brought on board this year. All of the directors have four years experience in the district, although one left for several years and came back. Participant profiles can be seen in Table 1 below.



Table 1

*Administrator Demographic Data*

Role	Ethnicity	Gender	Years Experience			
			Teaching	Evaluator	Educator	District
Principal	White	M	9	4	13	10
Principal	White	M	8	4	8	1
Principal	Hispanic	M	3	9	19	4
Principal	White	M	7	7	17	1
Principal	White	F	22	15	23	1
Principal	White	F	8	4	12	13
Director	White	F	8	2	15	10
Director	White	F	7	6	13	3
Director	Hispanic	F	9	3	12	3

*Note.* M= male, F=female

**Timetable and Action Plan**

This cycle of action research took place over the first four months of the 2011-2012 school year. The professional development activities were conducted from August through December, at meetings that I initiated and in which I attended as a participant observer. The purpose of this intervention was to help the administrative team a) mitigate concerns about the adoption of the teacher evaluation adoption, b) develop a common understanding of the professional teaching standards, c) gain confidence about their inter-rater reliability on teacher

evaluations, and d) understand the purpose of teacher evaluation to improve professional practice.

The professional development activities were structured to provide opportunities for participants to cycle through experiences of direct learning, collegial dialog, guided practice in field experiences, and back to direct learning and collegial dialog. This process was derived from Gavelek and Raphael's (1996) and Gallucci's (2007) adaptation of Harre's (1984) Vygotsky Space. The process is conceptually illustrated in Figure 3.

Plans for these interventions were conceptualized relative to the four phases of Vygotsky Space where learning takes place in a sociocultural setting and is *appropriated* by individuals, *transformed* based on individual context of needs and uses, *publicized* in ways that may influence others, and eventually *conventionalized* as part of the accepted practice of our school district. I approached learning in these activities as cyclical and evolutionary on both the individual and collective levels.

Although activities were designed to help participants learn through cyclical iterations through Vygotsky's Space, I approached these activities knowing that participants were functioning at any given time in any of the quadrants and not necessarily in sync with my intentions. I also approached participants individually, with the understanding that each has a different profile on the seven developmental SoCQ relative to this innovation (Hall, 1979; Hall & Hord, 1987, 2006; James, 1981). Findings from the SoCQ surveys (George et al., 2008) helped me understand individual differences and to know how to facilitate

implementation strategies for each. These findings also provided relevant information about the group as a whole.

I began with collaborative learning activities centered on describing each of the five professional teaching standards from our newly adopted teacher evaluation system (Appendix A). In accord with the Vygotsky Space concept of appropriation (Gavelek & Raphael, 1996; Harre, 1984), participants were engaged in classroom activities where rubrics for each element from the five standards were reviewed and collectively interpreted by the group. Participants were encouraged to offer clarification about how they interpreted each rubric in their evaluative settings. Participants posted ideas by editing shared electronic documents. These ideas were collectively developed and refined through a mediated discussion of the group, both electronically and in person. In so doing, knowledge of the facilitators, myself or another director or principal, was combined with the knowledge of other administrator participants to refine the rubrics and create common understandings of their meaning in evaluation practice.

These refined rubrics were recorded and shared again with participants electronically through Google Docs, a website designed to allow members to share and collaboratively edit documents outside of the group meetings. Participants were encouraged to test these refined rubrics at their sites as they continued the evaluation process with teachers. They were further led to reflect on their learning in the group meetings and in paired observations with me as a participant observer. These reflective experiences were used as a means to

encourage the transformation of information, evidenced as they took ownership of the learning in the context of their practice of teacher evaluation and development.

Participants returned to the classroom setting periodically to share evaluation experiences with others, review the rubrics, and test and demonstrate their learning in the social context of the group. These shared experiences provided opportunities for additional appropriation as information was shared and individuals learned from others' experiences. Additionally, participants had opportunities for transformation as they continued to gain ownership for learning in the context of their own understandings, and publication as they presented their evaluation and professional growth plan findings within the group and for district administration. The outline of this intervention is summarized in Figure 4, with references to dates, activities and participants.

<b>Dates</b>	<b>Participants/ Initiated Activity</b>
August 18, 25	<u>Administrator CoP:</u> <ul style="list-style-type: none"> <li>• Introduction and purpose setting.</li> <li>• Present review overall concept of teacher evaluation.</li> </ul>
August 29 September 20 October 12, 20, 29 November 21, 29 December 1	<u>Administrator CoP:</u> <ul style="list-style-type: none"> <li>• Collaborative presentations on five teacher evaluation standards</li> <li>• Collective negotiation of locally standardized rubrics for each element</li> <li>• Collective negotiation of locally developed scheme for integrating student achievement outcomes</li> </ul>
November 21, 29 December 1	<u>Administrator CoP:</u> <ul style="list-style-type: none"> <li>• Case studies brought by administrators for collective discussion of teacher evaluation/growth plan.</li> </ul>
Open Schedule: October-December	<u>Individual Administrators and Researcher:</u> <ul style="list-style-type: none"> <li>• Researcher and administrator shadow administrators practicing evaluation/professional growth at sites</li> </ul>

*Figure 4.* Outline of intervention activities

Securing confidentiality and providing ethical protection for each participant and the site location was paramount to this study. A request to conduct the study was submitted to the Arizona State University Office of Research Integrity and Assurance for institutional review board approval. Their acceptance of the request was granted and can be seen in Appendix B. Each participant signed and retained a copy of an informed consent form describing the parameters of the study, participant involvement, measures of protections, including the right to withdraw at any time, and the intended use of the data (Appendix C). The researcher did not identify participants or the specific location of school or district

sites. In no case was any staff or student identified by the researcher or in the research.

### **Measures: Data Collection**

Measures for this study were adopted to answer the following research questions:

- 1.) What concerns do administrators within my district have about effective implementation of our new evaluation system and how and to what extent will those concerns change as our CoP learns, practices and engages in discourse?
- 2.) What will administrators say, do, and feel as our CoP attempts to develop a common understanding of the professional teaching standards, increase inter-rater reliability, and understand the purpose of teacher evaluation to improve professional practice.
- 3.) How do I lead this process of change?

The following quantitative and qualitative measures were triangulated (Johnson & Onwuegbuzie, 2004) to validate findings relevant to the research questions listed above.

**Measure 1: Surveys.** Two questionnaires were administered for this study at pre- and post-intervention intervals. Quantitative data were gathered from a pre- and post-survey of the Stages of Concern Questionnaire (SoCQ). The Administrator Evaluation Survey (AES) was also given pre- and post survey in August and December to each administrator participating in the professional development experiences. These two surveys, the 66-item SoCQ, and the 31-item

AES, were administered together in two clearly separated sections (See Appendix D). I piloted both sections of the survey instrument six months earlier and revised each of the instruments to improve reliability and readability.

The SoCQ was developed by the Research and Development Center for Teacher Education at the University of Texas at Austin (Hall et al., 1973). The SoCQ is closely associated with the Concerns-Based Adoption Model (CBAM) and has been used to measure Seven Stages of Concern (SoC) about any innovation. These seven SoC were identified and confirmed to exist about any innovation (Hall, 1979; Hall & Hord, 1987, 2006; James, 1981). The CBAM, as described by Hall (1979) and other researchers (Hall & Hord, 1987, 2006; James, 1981), has been used to measure the developmental degree to which innovations have been implemented in school settings (James, 1981). According to Hall and Hord (2006), the SoCQ has, “strong reliability estimates (test/retest reliabilities range from .65 to .86) and internal consistency (alpha-coefficients range from .66 to .83)” (page 80). The SoCQ instrument is a 35-item questionnaire with a 7-point Likert-type scale ranging from this statement is “very true of me now” to “not true of me now.” This questionnaire has been found to be a reliable and valid survey to identify and characterize the stages of concern of individuals involved with innovations in school settings (Hall, George, & Rutherford, 1998). Hall and colleagues investigated the validity by studying scores on the seven stages to see how they relate to one another and other variables from concerns theory. They used these intercorrelation matrixes, interview data and confirmation of expected group differences and changes over time to prove the validity of the SoCQ scores.

They found that, “all [tests] . . . provided increased confidence that the SoCQ measures the hypothesized Stages of Concern” (p. 20).

The SoCQ survey I adapted for this study was a 35-item questionnaire, with an 8-point Likert-type scale and one open-ended question. My SoCQ was administered according to specific guidelines established by George et al. (2008). Survey responses on the SoCQ were gathered using an eight-point Likert scale ranging from 7 to 0. The range begins with 7 to indicate *this item is very true of me now* to 1 to indicate *this item is not at all true of me at this time*. The response of 0 indicates *this item seems irrelevant to me now*. The survey was produced using SurveyMonkey software and transmitted electronically to each participant with assurance of complete anonymity. As a means to allow pre- and post-survey comparisons and maintain confidentiality, participants were asked to generate a unique four-digit number for identification purposes.

The AES survey was based on four constructs related to the intended outcomes of the intervention and this study’s research questions (Iarossi, 2006). Table 4 illustrates a description of each construct. Survey responses were gathered using a five-item Likert scale ranging from 5 *strongly agree*, 4 *agree*, 3 *not sure*, 2 *disagree*, and 1 *strongly disagree*. There was an open-response section following each construct.

The first construct, “perception of inter-rater reliability” was intended to measure perceptions of how likely the evaluation system was to produce similar rating outcomes for the same teacher among different administrators. The second construct, “potential for evaluation system to help teachers” was designed to



measure participants' perceptions about their ability to use the evaluation tools to engage teachers to improve their professional practice. I used results from these items to learn more about to what extent administrators feel able to use the evaluations to help teachers improve their professional practice and how engaged teachers become in this process. The third construct, "administrator perception of teacher evaluation tools" was designed to find the level of confidence administrators have in the instrument as a tool to improve professional practice. The fourth construct, "perception of professional growth plans for teachers" was designed to reveal administrators' perceptions of the usefulness of professional growth plans as a means for improving teachers' professional practice.

Each construct was designed to provide additional data for one or more research questions. For example, the second, third and fourth constructs, taken together, were designed to provide another measure of administrators' understandings of the purpose of teacher evaluation to improve professional practice. Alignment of survey questions to research questions is shown in Figure 5.

<b>Survey Construct</b>	<b>RQ:</b> Development of a common understanding of the professional teaching standards	<b>RQ:</b> Understanding of the purpose of teacher evaluation to improve professional practice	<b>RQ:</b> Confidence in inter-rater reliability of teacher evaluation	<b>RQ:</b> Concerns administrators have about implementation
perception of inter-rater reliability	X		X	X
potential for evaluation system to help teachers	X	X		X
administrator perception of teacher evaluation tools	X	X	X	X
perception of professional growth plans for teachers	X	X		X

*Figure 5.* AES survey construct correlation to research questions

**Measure 2: Interviews.** Pre-innovation and post-innovation interviews were used to address issues relevant to the primary research questions (Kvale, 1996; Suzuki, Ahluwalia, Arora, Mattis, 2007). The interview protocol (see Appendix E) shows each question and the corresponding research questions addressed by interviewees. Questions were open-ended to help me answer the research questions (Anderson, Herr, Sigrid Nihlen, 2007). A digital audio recording device was used to record the interviews. Audio files were kept in a secure location and respondents were assured of complete confidentiality.

**Measure 3: Meeting records.** Written records were kept at each of the meetings of the administrators' CoP. These minutes were recorded as a means of gathering additional detailed information about the discourse and learning

activities of the participants as they talked with each other in the full group. The minutes were taken in such a way as to ensure complete anonymity of participants. I dictated a more detailed accounting of each of the meetings using a digital recording device. These recordings were later transcribed for analysis.

**Measure 4: Research journal.** Throughout my innovation, and beginning with the difficult process of aligning leadership priorities to lay the groundwork for this innovation, I kept a research journal. This journal was kept primarily to provide a personal perspective to answer the following research question: How do I lead this process of change? The journal became a personal forum to document my leadership process, but also expanded to include many entries describing the details of participant activities in the process of implementing teacher evaluation. These entries were made at least weekly and whenever there was a major development in the innovation. The research journal was later transcribed for more detailed analysis. The content also included copious references to administrators' concerns and learning processes, and thereby served also to inform my first and second research questions.

## **Chapter 4 Analysis and Results**

Chapter 3 addressed the design of this study and how data was collected.

This chapter focuses on the results from the study and is presented in three sections. The first section presents results from the quantitative data gathered from the AES Construct pre- and post-surveys given to measure participants' perceptions about teacher evaluation. The second section presents the results of the pre- and post-surveys for the SoCQ instrument measuring participants' concerns profiles. The third section presents results for the qualitative data gathered in surveys, interviews, meeting notes and my research journal.

Review and analysis of these sources of data provided insight from multiple sources to answer the research questions posed by this study, 1) What concerns do administrators within my district have about effective implementation of our new teacher evaluation system, and how, and to what extent, will those concerns change as our CoP learns, practices and engages in discourse about the teacher evaluation?; 2) What will administrators say, do and feel as our CoP attempts to develop a common understanding of the professional teaching standards, increase inter-rater reliability on our teacher evaluation instruments, understand the purpose of teacher evaluation to improve professional practice?; and 3) How do I lead this process of change? For research questions one and two, I also explored whether certain demographic characteristics of the participants make a difference.

## **Results of the Quantitative AES Construct Survey**

The quantifiable responses to the questions on the pre- and post-surveys from the AES survey were exported from Survey Monkey and formatted for import into SPSS 20, a statistical analysis software package. As a means to establish the internal reliability, a Cronbach's Alpha coefficient was calculated for each construct of the AES survey, as well as to establish the internal reliability of the overall survey (Cronbach, 1951). Descriptive statistics were generated to include frequencies, means, effect sizes and standard deviations (Miles & Huberman, 1994). Statistical analysis also included an independent samples *t*-test to determine whether various participant demographic attributes have any substantial relationship to individual or clusters of responses. All responses were analyzed using both a frequency count function and a *t*-test. The *t*-test allowed a comparison between the means of the pre- and post-surveys to determine whether or not the pre- and post-scores among the participants were statistically significant or simply a chance finding (Gay, Mills & Airasian, 2009). Additionally, correlations among the four constructs were analyzed to find significant correlations between demographic subgroups of participants and correlations among items and constructs (Smith & Glass, 1987).

Nine participants were invited to participate in the AES pre- and post-surveys. Eight responded to the pre-survey and seven responded to the post-survey. Results from the AES Construct Survey are reported in four sections, addressing:

1. Reliability of the survey
2. Comparison of pre- and post-survey means
3. Descriptive statistics related to participants' perceptions of their experiences with teacher evaluation
4. Correlations of participant demographics to survey outcomes
5. Correlations among survey items.

**Reliability of AES survey.** The pre- and post-intervention survey used to evaluate participants' perceptions about teacher evaluation consisted of four constructs designed to assess perceptions of: (a) inter-rater reliability, (b) potential for evaluation system to help teachers, (c) the teacher evaluation tools, and (d) professional growth plans for teachers. Reliability for this survey was positively established during planning and development of the AES, and again after administration of the pre- and post-survey. Positive reliability was demonstrated when alphas were greater than .70. Internal consistency was shown for four of the subsets representing participants' perceptions about teacher evaluation. The coefficient-alpha results are provided in Table 2.

Table 2

*Coefficient-Alpha Estimates of Internal-Consistency Reliability for Teacher Evaluation Survey Instrument*

Administrator Opinions	Within Factor Items	Coefficient Alpha Estimate of Reliability Pre-Survey
Total survey	Items 1-21	0.84
Perception of inter-rater reliability	Items 1,2,4,5	0.73
TES potential to help teachers	Items 6,7,10-14,16	0.78
Teacher engagement with the TES	Items 15,17,18	0.79
Extent of positive perception of TES tools	Items 8,19-21	0.80

**Descriptive statistics of participants’ pre- and post-survey responses.**

When administrators were asked opinions about their perception of inter-rater reliability on the teacher evaluation instrument, participants responded to five questions in a pre- and post-survey. Figure 6 shows additional detail regarding administrators’ increased confidence in inter-rater reliability among the participants. Items #1 and #2 showed the greatest growth and item #4 dipped slightly between pre- and post-surveys, partly because one participant strongly disagreed that “...peers are not overly generous or overly severe in rating teachers on the TES.” In the pre-survey, only three of eight administrators agreed or strongly agreed on #5 that, “Teachers can be confident of the consistency of

evaluations of teachers on our campuses,” and two disagreed. On the same question in the post-survey four of seven agreed or strongly agreed with this statement and the rest were unsure. No one disagreed. Figure 6 shows the entirety of results from this group of survey items.

<b>Administrators’ Opinions</b>	<b>Item</b>	<b>M SD</b>	<b>Strongly Agree (5)</b>	<b>Agree (4)</b>	<b>Not Sure (3)</b>	<b>Disagree (2)</b>	<b>Strongly Disagree (1)</b>
#1. Our TES is likely to result in consistent ratings between administrators.	Pre n=8	2.50 .926	n=0 0.0%	n=1 12.5%	n=3 37.5%	n=3 37.5%	n=1 12.5%
	Post n=7	3.43 .787	n=0 0.0%	n=4 57.1%	n=2 28.6%	n=1 14.3%	n=0 0.0%
#2. I am confident that most administrators’ ratings of the same teacher would be similar to my own.	Pre n=8	2.50 .535	n=0 0.0%	n=0 0.0%	n=4 50.0%	n=4 50.0%	n=0 0.0%
	Post n=7	3.29 .756	n=0 0.0%	n=3 42.9%	n=3 42.9%	n=1 14.3%	n=0 0.0%
#4. My administrative peers are not overly generous or overly severe in rating teachers on the TES.	Pre n=8	2.88 .641	n=0 0.0%	n=1 12.5%	n=5 62.5%	n=2 25.0%	n=0 0.0%
	Post n=7	2.71 1.13	n=0 0.0%	n=2 28.6%	n=2 28.6%	n=2 28.6%	n=1 14.3%
#5. Teachers can be confident of the consistency of evaluations of teachers on our campuses.	Pre n=8	3.00 .765	n=0 0.0%	n=2 25.0%	n=4 50.0%	n=2 50.0%	n=0 0.0%
	Post n=7	3.71 .756	n=1 14.3%	n=3 42.9%	n=3 42.9%	n=0 0.0%	n=0 0.0%

Figure 6. Frequencies for construct one: Teacher evaluation survey

**Comparison of pre- and post-survey response means.** For purposes of reporting results, I interpreted average survey scores by labeling the range between 1.00 - 1.80 to mean *strongly disagree*, 1.81 - 2.60 to mean *disagree*,



2.61 - 3.40 to mean *not sure*, 3.41 - 4.20 to mean *agree*, and 4.21 - 5.0 to mean *strongly agree*. Figure 7 shows administrators' level of agreement with the four constructs of the teacher evaluation survey with a comparison of pre- and post-survey results. The strongest mean level of agreement is shown on the pre-survey opinions about the potential of the evaluation system to help teachers. This level dipped slightly in the post-survey. The largest gain from pre- to post-survey level of agreement is seen on the construct measuring confidence in inter-rater reliability. Administrators averaged close to *not sure* (3.0) on all measures.

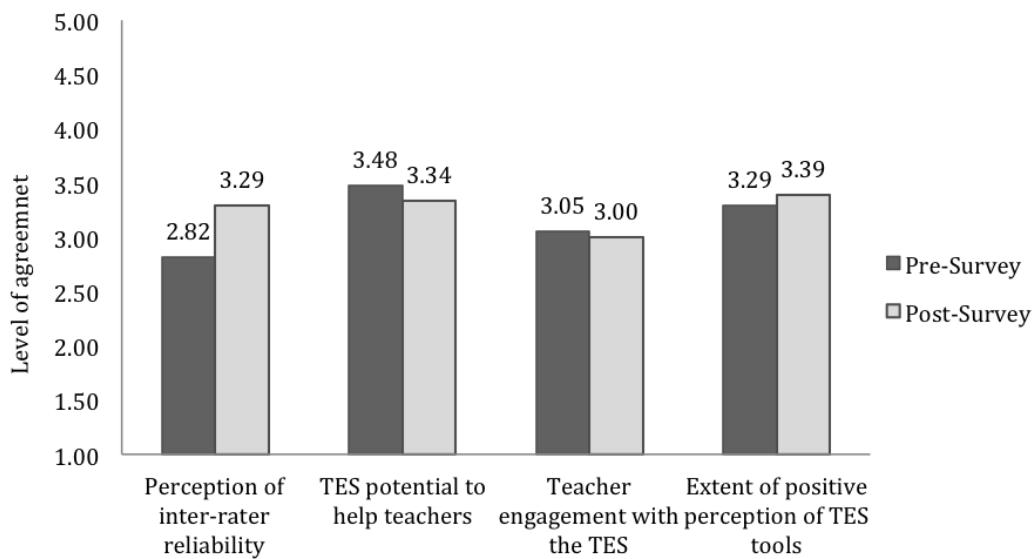


Figure 7. Administrators' perception of teacher evaluation

A paired-samples dependent *t*-test was performed to determine if any significant differences in responses on the pre- and post-survey could be attributed to participation in this innovation. These differences were measured by

groups of survey items based on four constructs. Significant differences were found only in the construct rating administrator confidence in inter-rater reliability. The level of agreement on this group of survey questions increased from a mean of 2.82 ( $SD = 0.49$ ) on the pre-survey to 3.29 ( $SD = 0.47$ ) on the post-survey. The difference between the two means (0.92) was statistically significant at the  $p < 0.05$  level, indicating that the administrators felt more confident in their inter-rater reliability after participation in the innovation. Cohen's (1988) definition of effect size was applied and indicated that an effect size ( $d = 0.97$ ) means that the administrators posted 0.97 standard deviations in growth, which might be classified as a large effect. The pre-survey average for this construct, 2.82, falls in the range of *not sure*. The post-survey average score of 3.29 is a significantly higher level of agreement, however this still falls in the *not sure* range on the survey.

Differences between the means of pre- and post-survey of the other three constructs changed only slightly and these differences were not statistically significant. Changes in perception of teacher evaluation potential to help teachers went down just slightly, beginning with a stronger 3.48, in the *agree* range, and moving slightly lower to 3.34 to fall in the *not sure* range. Survey measures of change in perception of teacher engagement with evaluation also dipped slightly. Pre-survey and post-survey average agreement both fell in the *not sure* range, with the pre-survey at 3.05, and post-survey at 3.00. The level of agreement about positive perceptions of the evaluation tools rose slightly from pre- to post-survey, but again both fell in the *not sure* range. The pre-survey level of agreement was

3.29, and the post-survey was 3.39. Table 3 shows these comparisons for all four of the survey constructs.

Table 3

*Participant Survey Response Difference by Construct From Pre-Intervention Survey to Post-Intervention Survey*

Construct		Pre-Survey	Post-Survey	$m_2 - m_1$	p	d
Perception of inter-rater reliability	<i>M</i>	2.82	3.29	.46	*.02	0.97
	<i>SD</i>	.49	.47			
TES potential to help teachers	<i>M</i>	3.48	3.34	-.16	.22	0.38
	<i>SD</i>	.41	.41			
Teacher engagement with the TES	<i>M</i>	3.05	3.00	-.05	.89	0.07
	<i>SD</i>	.76	.61			
Extent of positive perception of TES tools	<i>M</i>	3.29	3.39	.11	.71	0.22
	<i>SD</i>	.72	.27			

*Note:* N = 8

\* mean difference is significant at  $p < 0.05$

**Results from analyses of participant characteristics and survey**

**outcomes.** Various participant characteristics were reviewed to determine whether there were substantial differences between different subgroups of participants. The subgroups compared here are between paired samples of male and female administrators, experienced and inexperienced principals, and principals and district administrators. Independent sample *t*-tests were run to

measure the difference in participant groups based on their survey responses on the four constructs. Significant differences were found in several instances. Female administrators were significantly more in agreement that the teacher evaluation process is engaging for teachers. As seen in Table 4, female administrators' mean level of agreement on this construct was 3.58 ( $SD = 0.42$ ), falling in the *agree* range. This compared to male administrators at 2.50 ( $SD = 0.43$ ), falling in the *disagree* range. The difference between the two means (1.08) was statistically significant at the  $p < 0.01$  level. A similar difference was found for the same construct when comparing new principals with more experienced. In this case, the experienced principals had significantly higher agreement scores on questions that addressed the level of teacher engagement in the teacher evaluation process. Experienced principals averaged 3.42, falling in the *agree* range on the survey, while the less experienced principals averaged 2.44, falling in the *disagree* range on the survey. Finally, when comparing principals with district office administrators, two constructs had significantly higher agreement scores for district administrators than for principals. District administrators were more likely to agree that the evaluation system engaged teachers (3.83 in the *agree* range), compared to principals (2.78 in the *not sure* range). District administrators were also more likely to see the evaluation system as helpful for teachers (4.06 in the *agree* range), compared to principals (3.30 in the *not sure* range). Table 4 provides additional details for these results.

Table 4

*Participant survey response difference by participant sub-group*

Construct	Participant Sub-Group	<i>M</i>	<i>SD</i>	$m_2 - m_1$	<i>p</i>
Teacher engagement with the TES	Male Admin.	2.50	0.42	1.08	.01*
	Female Admin.	3.58	0.43		
Teacher engagement with the TES	New Principal	2.44	0.19	0.91	.01*
	Exp. Principal	3.42	0.42		
TES potential to help teachers	Principal	3.30	1.8	0.76	.00*
	District Admin.	4.06	0.08		
Teacher engagement with the TES	Principal	2.78	0.58	1.06	.02**
	District Admin.	3.83	0.24		

\* mean difference is significant at  $p < 0.01$

\*\* mean difference is significant at  $p < 0.05$

**Correlations among the four constructs.** The four constructs from the survey are shown in Table 5 below to highlight significant correlations found in participant responses between constructs (Smith & Glass, 1987). It is clear that rating of constructs 2 and 3 (evaluation system's potential to help teachers and teacher engagement with the evaluation system) are positively correlated. As shown in the Table 5, this correlation is significant at the .05 level, meaning there is a 95% certainty in the correlation. Additionally, there is a similarly significant positive correlation between construct one (administrators' confidence in inter-

rater reliability), and construct two, (the evaluation system's potential to help teachers).

Table 5

*Teacher Evaluation Survey: Correlations Between Four Constructs on the Post-Survey*

		C1	C2	C3	C4
C1: Perception of inter-rater reliability	Pearson Correlation	1			
	Sig. (2-tailed)				
C2: TES potential to help teachers	Pearson Correlation	.821*	1		
	Sig. (2-tailed)	.023			
C3: Teacher engagement with the TES	Pearson Correlation	.832*	.811*	1	
	Sig. (2-tailed)	.020	.027		
C4: Positive perception of TES tools	Pearson Correlation	.396	.579	.480	1
	Sig. (2-tailed)	.379	.173	.276	

\* Correlation is significant at the 0.05 level (2-tailed).

### Results of the Stages of Concern Questionnaire

The pre- and post-survey results of the SoCQ were analyzed according to methods prescribed in *Measuring Implementation in Schools: The Stages of Concern Questionnaire* (George et al., 2008). The reliability and validity of the SoCQ has been well established for over 30 years under prescribed conditions (George et al., 2008) that were adhered to in this study. The alterations that I made to the SoCQ are within the recommendations prescribed by George et al. (2008).

Scoring the instrument required calculating raw scores for each of the seven stages on scales of implementation and plotting the results on a SoCQ chart to match to the percentiles calculated by the original stratified sample from the 1974 standardization group of 830 individuals. An electronic spreadsheet designed by the authors for this purpose (George, et al, 2008) was utilized. The seven stages of concern were reviewed in Chapter 2, and a summary can be seen in Figure 1.

All nine participants were also invited to participate in the SoCQ pre- and post-surveys. Eight responded to the pre-survey and seven responded to the post-survey. Five participants' pre-surveys could be reliably matched to their post-survey. All participants were included in the full group results where individual matching was not relevant. As a first step, results from the SoCQ were analyzed in terms of peak and second highest stage scores. This analysis was followed by profile interpretations to analyze individual and group patterns of concerns measured by the instrument. The third step was to analyze patterns of change from pre- to post-survey for individual cases and group averages. Interpretation of the results of the SoCQ survey data was closely guided and analyzed in accord with the methods outlined in *Measuring Implementation in Schools: The Stages of Concern Questionnaire* (George et al., 2008).

**Peak and secondary stage score interpretation.** This method of analysis is relevant for both individual and group data, and the results from this study yielded a pattern of results between and among users about the intensity of their concern scores for each area. According to George et al. (2008), "The higher the

score, the more intense the concerns are at that stage” (p. 32). This analysis was conducted by listing each participant’s pre- and post-survey percentile scores on each stage of concern on a matrix. This matrix included survey results for each individual participant, and the mean of the participant group as a whole. The matrix also included mean scores of various paired participant groups, including principals/district office administrators, more experienced/less experienced evaluators, and principals new to the district/principals veteran in the district.

The matrix in Figure 8 shows the peak and secondary stage scores for the five respondents whose pre- and post-survey results were matched. The full matrix can be referenced in Appendix F. The peak score pattern on this sample is similar to what was seen on the matrix as a whole. The highest level of concern was remarkably consistent at Stage 0 (with the exception of a shared peak on the first respondent). Stage 0 scores provide an indication of the degree of priority the respondent is placing on the innovation and the relative intensity of concern about the innovation. According to George et al. (2008), “Stage 0 addresses the degree of interest in and engagement with the innovation in comparison to other tasks, activities and efforts of the respondent” (p. 33). According to these survey results, respondents do not appear to be placing the implementation of the new teacher evaluation system as a primary concern.

Because of the developmental aspect of the stages of concern, George et al. (1986) predict that secondary peak scores are often adjacent to the peak score. In the case of the five participants’ results shown in Figure 8, the secondary peak scores were not adjacent, but had relatively higher concern levels at S2 and S3.



The data reveals instead that secondary peak scores were clustered heavily at S3. All participants scored secondary highs in the post-survey on S3, and all but one had this as their high secondary score in the pre-test. A high score in S3 indicates strong concerns about the management and logistics of implementation of an innovation. Four participants' level of concern on S3 became noticeably more elevated at the post-survey, and eight had their levels drop noticeably. S2 was another area where there were other high secondary peak scores, and even a peak score. Concerns on S2 are related to more personal concerns about adequacy to implement the system and conflicts with other innovations.

PARTICIPANT	S0	S1	S2	S3	S4	S5	S6
PRE-1	91	84	92	88	33	40	57
POST-1	94	45	52	56	9	31	26
PRE-2	91	51	55	56	21	40	22
POST-2	91	43	48	52	16	31	30
PRE-3	99	63	55	27	7	64	57
POST-3	91	43	52	65	16	31	26
PRE-4	94	16	25	34	11	22	20
POST-4	96	51	59	60	19	40	30
PRE-5	99	43	35	65	59	68	65
POST-5	87	45	55	56	16	25	17

XX =Peak Stage Score  
XX =Second Highest Stage Score

Figure 8. SoCQ questionnaire peak and secondary scores on pre- and post-survey

**Profile interpretation.** Individual profiles were also analyzed using the pre- and post-SoCQ survey results to track individual progressions in the concerns

profiles during the implementation of the innovation. George et al. (1986)

hypothesized the following about an ideal progression of individual growth in adopting a new innovation:

...as individuals move from nonuse and scant awareness to an innovation to beginning use and, eventually, more highly sophisticated use, their concerns move through the defined stages. They begin with their concerns being most intense at Stages 0, 1, and 2, then shift to Stage 3, and ultimately register their highest levels of concern at 4, 5, and 6. If the innovation is appropriate and well designed and if there is adequate support for its implementation, an individual's concerns profile plotted over time should look like a wave moving from left to right. (p. 37)

Figure 9 shows this progression visually in a line graph. Similar line graphs were created for each participant to compare pre- to post-intervention profiles. Although certain features of individual change could be seen and described, growth from pre- to post-intervention was not generally patterned or predictable among the participants of this study. Change in concern levels of individual were much less clear or patterned for participants in this study and showed no discernable evidence of a regular progression displayed in the hypothesized development of SoCQ line graph.

A line graph for one of this study's participants is shown in Figure 10 to illustrate a sample progression for one candidate. Participant 1 is identified as a new principal with less than five years experience evaluating teachers. Like most other participants, this individual had a high score in S0. Concerns in

informational, personal and management stages dropped markedly on the post-survey, while concerns in the later stages remained lower, indicating that this individual was not yet reaching the levels of an experienced or renewing user. Diagrams of the other four matched participants can be seen in Appendix G.

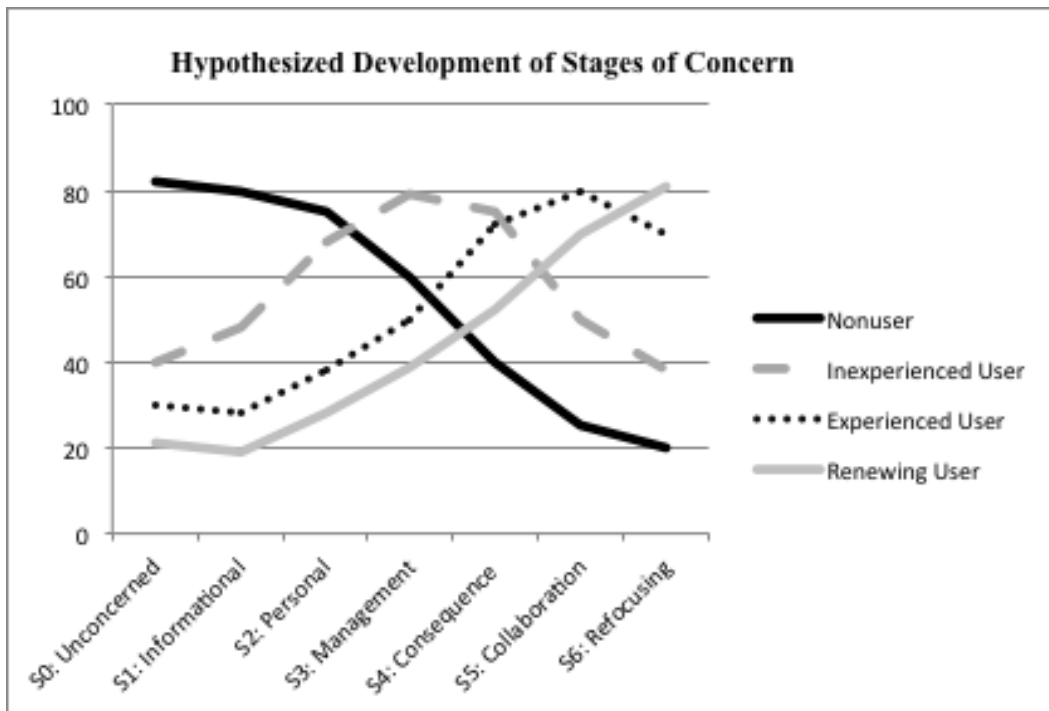


Figure 9. Hypothesized development of stages of concern

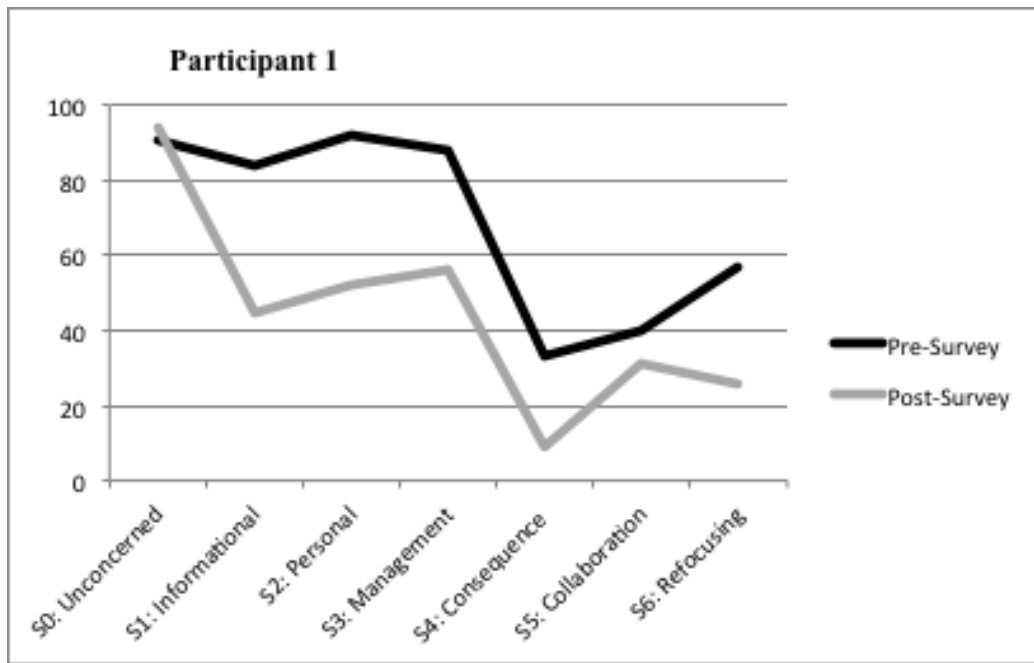


Figure 10. Stages of concern profiles for participant 1

## Results of Qualitative Data Analysis

Qualitative data collected from the open-ended survey questions, interviews, recorded meetings, researcher's journal, and meeting notes were analyzed using grounded theory. HyperRESEARCH v. 3.0.3 (Researchware, 2011), software tools were utilized to aid in organizing textual materials gathered in the course of the intervention. Open coding was used to gain a deeper understanding of the text that was collected (Anderson et al., 2007; Glaser & Strauss, 1967; Ryan & Bernard, 2000). Following an inductive approach, several readings of the textual data were made before beginning to seek codes. The open coding approach was used to capture the detail, variation, and complexity of my qualitative data. As I coded and during the entire process I constantly compared data instances, cases, and categories for conceptual similarities and differences

using the constant comparison method. Lincoln and Guba (1985) described four distinct stages of the constant comparison method: “comparing incidents applicable to each category, integrating categories and their properties, delimiting the theory, and writing the theory.” (p. 339).

As I moved through the analysis process I sampled new data and cases on theoretical grounds. Theory was extended and enriched my emerging ideas. In the process I wrote memoranda to link concepts to theory. Next coding phases were more focused. I employed focused, axial, or cross-referenced coding. I looked for repeated patterns or properties that made connections between categories, and ultimately used this analysis to understand how study participants helped answer the identified research questions (Glaser & Strauss, 1967). Coding, comparisons, and theory sampling continued until saturation (no new relevant insights) was reached.

**Participant concerns.** The first of these themes, participant concerns, addressed issues that study participants had about the innovation of adopting a standards-based teacher evaluation. Data from interviews and survey comments were the sources of this information. Figure 11 presents an overall view of nine areas of concern that emerged from all of the data, and the frequency with which they were found. Figure 12 shows the same data disaggregated into pre- and post intervention categories for comparative purposes.

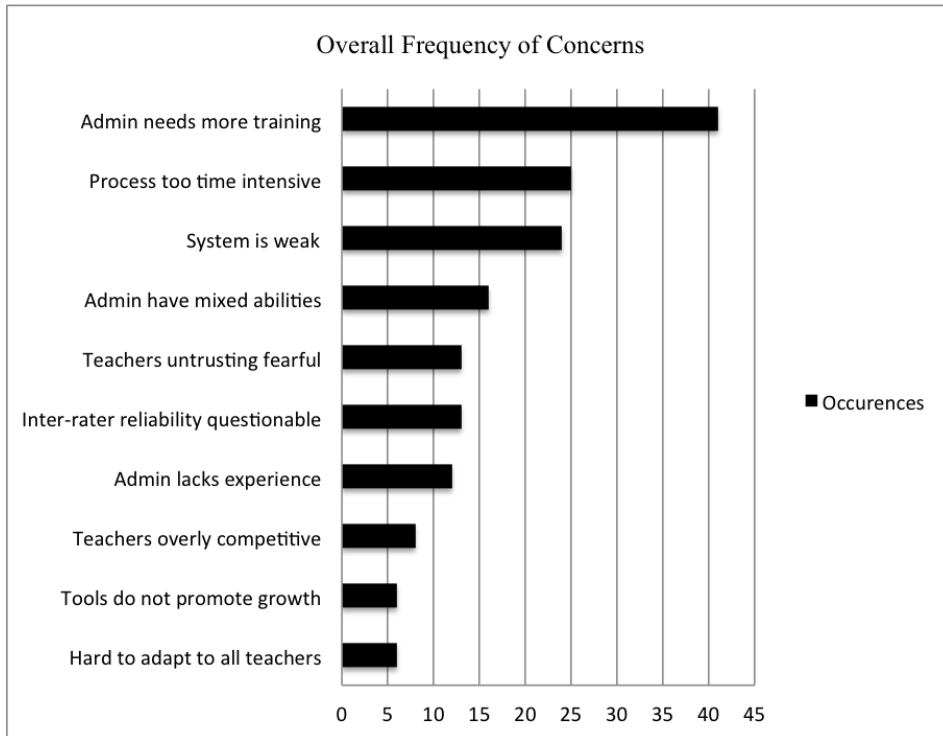


Figure 11. Overall frequency of participant concerns

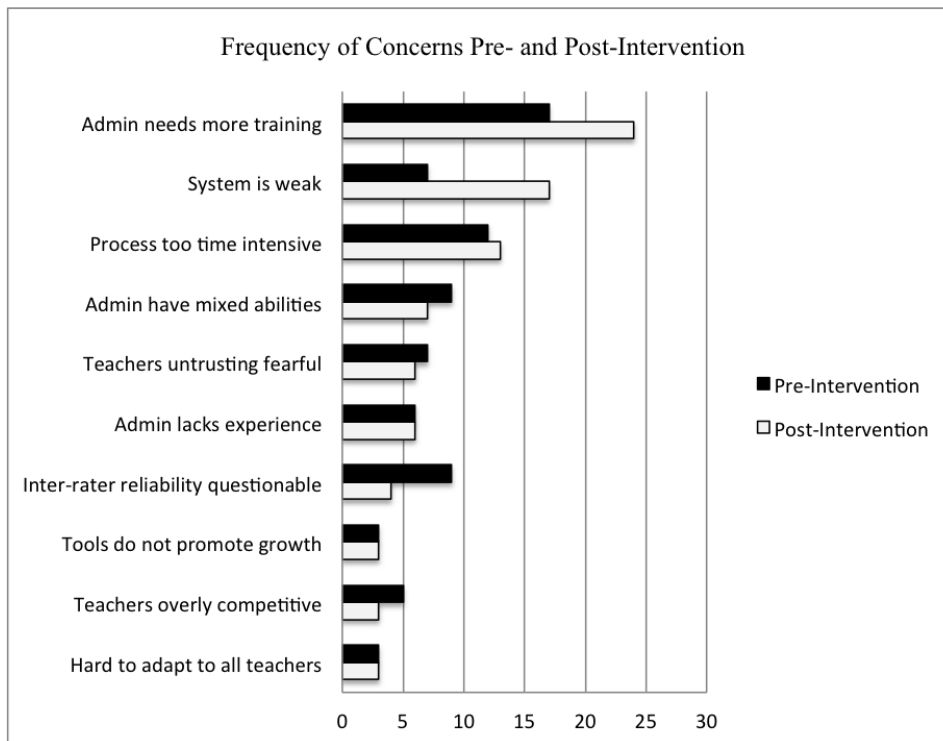


Figure 12. Frequency of participant concerns pre- and post-intervention

Concerns were most frequently noted about the need for additional training. Administrator concerns in this area were cited most frequently in both the pre- and post-intervention, but there was a notable increase in frequency of concerns in this area in the post-intervention surveys and interviews (24 occurrences), compared to pre-intervention (17 occurrences). It appears that after the learning experiences that administrators' expanded awareness may have led to a greater need for training in this area. One administrator from the district office noted in a post-intervention interview,

I think that it's a work in progress; that it's definitely not perfect in any regard. We have a lot of work to do, but I think that we have brought the administrators that are using it closer in their understanding of the rubrics themselves, and they are somewhat overwhelmed by the complexity. I don't know if we've clearly defined specific practices, more so we've generalized and maybe defined some of them, but we have not necessarily pinpointed exactly what you might see from a teacher.

A principal with less than five years experience evaluating teachers said, I get worried about getting lost in all the procedures and the how of it, compared to utilizing it as a tool and maybe that's just me being new to it, but there's so many components to get used to which has to do with my experience. I don't exactly know if I can use it as well as I have with other systems. I don't know if I'm as successful with this valuable tool.,

Another new principal said when asked about current concerns,

I'm concerned whether I am doing it justice, that I want to be doing it the right way. I want to be utilizing it to meet all the reasons it was put into place, because I know there are a lot of factors in this; like teacher retention, professional development, and teacher improvement...

Another more experienced principal said,

I think there needs to be more conversation around certain things. I think we need to continue to do what you started out doing, but looking at some lessons or going into more classroom together, and then coming out and debriefing about that, because the conversations that take place – that's how we'll get the value in getting everybody on the same page.

Yet another new principal said, "This (the evaluation system) worries me a little bit, wish we had more seat time."

Another new principal commented on the need for additional professional development in terms of the difficulties of aligning reform efforts of the adoption of the teacher evaluation system and similar teaching standards presented through a district-wide grant initiative:

We're still aligning with [the grant initiative] teaching standards; it's gotten to be a lot more challenging—we're trying to norm that out as a district and as administrators, and we all have varying degrees of experiences and so forth in education and professional development so that we get it balanced out - and we're all kinda looking at it generally the same.



The second most frequently cited concern in the post-intervention qualitative data was relative to doubts about the quality of the teacher evaluation system adopted by the district. Similar to the concerns about administrators needing more training, this concern was cited more frequently on data from the post-intervention (17 occurrences) than pre-intervention (7 occurrences). Much of this increase in critical comments about our current adoption were in the context of comparing the currently adopted materials to other evaluation systems to which participants became exposed in the course of the innovation.

Many of these comments were critical of the rubrics, and stemmed from perceived differences between the teacher evaluation system we adopted and other pedagogical measures that were part of the district-wide reform efforts briefly referenced in Chapter 1. There was a lively discussion at one of the CoP meetings where administrators were critical of the evaluation system in many instances because it didn't align well with these reform initiatives and seemed to stretch them in different directions depending on what reform project they were using as a frame of reference. In another meeting, administrators became exasperated trying to re-define the rubrics in the teacher evaluation to make them fit better with the professional development of the reform initiative. They argued about ways to change the evaluation system to fit the new initiatives but struggled with the practicality of making these changes directly. They also struggled with conflicts about re-interpreting the standards and rubrics as they were written, and many expressed caution about being overly liberal in their interpretation of the rubrics. Others expressed concern that the lack of alignment is ultimately

confusing for teachers unless we constantly provided crosswalks to help everyone understand when selected rubrics could be considered the same or similar in both initiatives.

Expressing frustration with interpreting the rubrics, one district administrator with more than five years experience as an evaluator stated, “Sometimes some of the rubrics are very big. So, for some of the more detailed things you might be looking for, they aren’t always specifically spelled out in the rubric, you have to really make that connection yourself.” Another experienced district administrator contrasted the evaluation system with evaluation rubrics that were being used in a district-wide reform effort by saying, “I think it’s hard to judge people on the criteria when the criteria are kind of stretched and your stretching to find that in a teacher. It’s hard to do.” An experienced principal expressed some regret about the selection by saying, “We could have spent a little more time looking at some other [evaluation systems], but... hindsight’s 20-20 and I’m OK with this one, but having seen other instruments now, I realize how general and broad, I think, ours is versus some of the other ones.” Another experienced principal criticized the clarity of the system by saying, “some of those elements and some of those descriptors within the elements, they’re very general and very hard to really pin down exactly what it means. I don’t think it’s explicit enough but I think we’re heading in the right direction.”

A new principal to our district who came to us with over five years experience evaluating teachers went in to greater depth about the shortcomings of

the system in terms of alignment of the evaluation rubrics with other resources used as part of district-wide reform efforts:

No, I don't feel like there's a common understanding of our evaluation rubrics or the teaching standards based on the dialogue that I've had with my colleagues around the walkthrough instrument and a lot of the... nuts and bolts of teaching and very heavy, direct instruction types of things. So... it creates a starker contrast between other coaching/evaluation strategies that we're working on now and what our evaluation instrument might be trying to measure.

Another strongly noted concern was the intensive amount of time necessary to properly evaluate teachers. Frequency of concerns expressed about the intensive amount of time was similar in the pre-intervention and the post-intervention measures. A new principal, experienced in evaluation, commented on the time it takes and the expectations of other reforms, "It is the amount of time this particular thing takes. At this point, I'm new, but the district from my perspective is involved in so many things at once." This principal went on to say, "I'm just concerned with getting in there and doing it the way I want to do. And so I get a good sense of what they are doing... I thought I would be able to get into classrooms two days a week, but it is really hard." An experienced district administrator said, "I think the challenge is finding time to do it."

An experienced principal commented,

I think the time constraint gets in the way of me supporting [more effective teachers] more than I do with the ineffective. So I spend more

time with my less effective teachers than I do my more effective, and I battle with that all the time because really, I feel like I should be spending more time with the more effective because it would increase their efficacy I guess.

Another experience principal commented on this primary concern, “The time. I think it’s the most important part of our job and sometimes the part we have the least amount of time to do.”

As can be seen in Figures 11 and 12, inter-rater reliability was less of a concern. Figure 12 shows that there were less than half as many notations of this concern in the post-intervention data than for the pre-intervention. Comments from the pre-intervention data include this from an experienced principal,

I would have a hard time defending it among teaching staff because they already know and because they already talk to each other. My campus talks to other campus and other staff members talked about what we do... I don’t think I was hard, I think I was fair. I didn’t overinflate anybody’s ratings so, I think a lot of people already feel that was done on some campuses and so if that’s – that’s tough... ultimately a lot of subjectivity comes into it and for evaluators, it’s very hard to stay away from that. So, there is some subjectivity and that’s a good thing, and it’s not a good thing sometimes.

Although there were fewer instances of concern in this area post-intervention, the concerns were strong. A new principal with experience as an evaluator said the following when asked about what was a primary concern,

I still believe we all use it differently... when I think about the different personalities in the room (administrators), I think... there's some people that might, that would be more analytical and really look at it deeper... I'm thinking that some of the schools are receiving different trainings and so therefore we have different expectations for what we're looking for and what we're seeing. So, just knowing that, and because we have different expectations because of different professional development, I'm not sure we would rate somebody the same because of that. Because there's those extenuating factors. Like for instance, at [one] school there's all that additional training... that colors [the principal's] lens when he goes in to evaluate.

As can be seen in Figures 11 and 12, additional concerns were found throughout the data. One relatively frequently mentioned concern, administrators have mixed abilities, was cited often in reference to the fact that three of the six principals were in their first year with the district. These newcomer principals had to quickly catch up with the others in learning about the teacher evaluation system. As one experienced principal said,

We have very little common understanding of the teaching standards and evaluation rubrics – especially true since we have three new people on board. We definitely need more chance to get on the same page.

Some concerns were raised about the unintended consequence of teachers becoming untrusting or fearful of the evaluation system. Some blamed this effect on the use of the evaluation system as a measure for reduction in force decisions

from the year before, and also the looming requirements to of Senate Bill 1040 that would enforce more serious consequences as a result of teacher evaluations.

One of the new principals complained,

I had a teacher express they are afraid to try new things because their scores are going to take a dip which will cause them to be evaluated poorly. It's almost as if she sees herself as constantly being under the evaluation umbrella – which should or could be a positive thing. You know sometimes people get nervous about evaluations and what we have is continuous.

Another new principal stated, “And when you don’t know them that well there is more fear, which inhibits the learning, the conversation, the whole process I think.” A district administrator summed up the dichotomy between our intentions to help teachers and the unintended consequences by saying, “We think it’s meant to be more of a growth instrument, but because accountability is attached you can lose your job over it if it comes out terrible.” An experienced principal shared how it’s framed,

I tell my teachers it shouldn’t be an instrument of destruction; you can look at it that way, and a lot of people freak out about it because of the reduction in force rubric or the achievement and all of that, but ultimately it should be to improve instruction in the classroom. Most are ok, but not all.

Additional concerns that appeared less frequently revolved around various topics. Administrators lack experience is an area where administrators expressed

concern about how new this process is for everyone and the challenges of such a steep learning curve. There were other concerns noted that the evaluation system was not designed well enough to guide teacher growth. Some noted that teachers became overly competitive, and rather than settle in to work on becoming more proficient in areas, became obsessed with less than a perfect score. Finally, concerns were raised about adapting the evaluation system for teachers with less traditional job responsibilities, such as instructional coaches and counselors. Administrators expressed frustration trying to make the teacher evaluation system fit these less traditional teaching assignments.

**Participant learning.** A second area of qualitative data analysis centered on participant learning that took place in the context of this innovation. Data from interviews, survey comments, and the research journal informed an understanding of the learning processes of the nine study participants. Rather than using grounded theory, the first review of data was conducted using a priori codes developed before examining the current data (Johnson & Christensen, 2012). The a priori codes were synthesized from the key transitional elements from Vygotsky Space theoretical frame: appropriation, transformation, publication and conventionalization (see Chapter 2, Figure 3). Figure 13 shows the frequency of occurrences of each of these phases of Vygotsky Space that were coded in the qualitative data. The frequencies illustrated demonstrate a general pattern of diminishing frequency of qualitative evidence for each of these categories as they move from the more basic learning stage of appropriation to the more advanced learning stage of conventionalization (Gavelek & Raphael, 1996). Of the nine

study participants, three principals were newly introduced to this innovation at the beginning of the school year. The rest began implementation of this system just one year earlier.

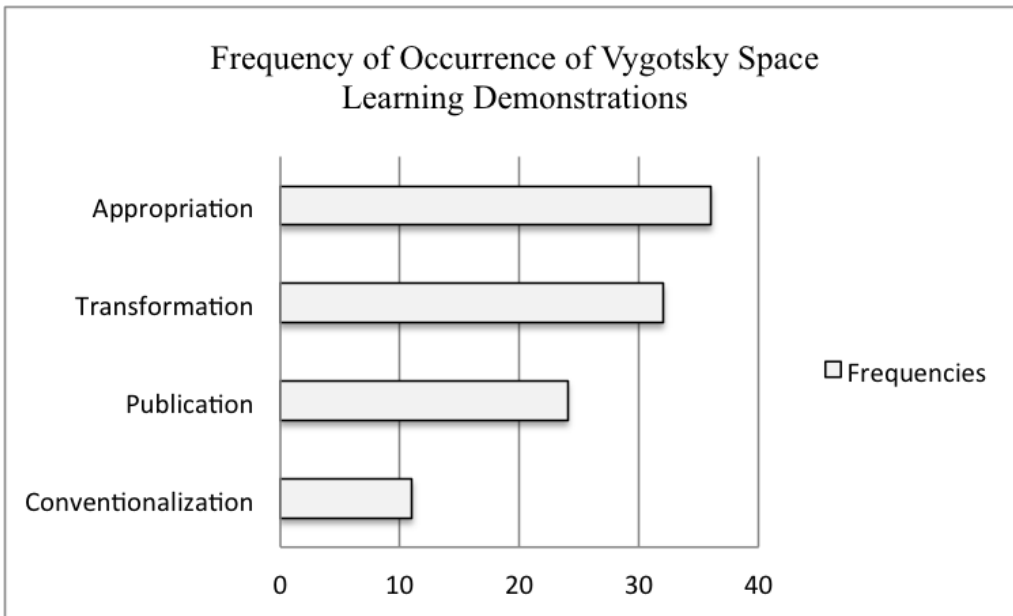


Figure 13. Frequency of occurrence of Vygotsky Space learning demonstrations

Comments about and evidence of appropriation – collective learning in our CoP – were noted in many instances. A director reflected on learning taking place by stating:

I do, I think that they are developing a better understanding as we continue our sessions: taking apart the standards and sharing observations from the videos. There's a learning piece there, there's some philosophical beliefs too that kind of pull us in a couple different directions. For the most part,



though, I think we have many starting places with the contributions (from participants) and are moving forward... they have a starting place to get a middle ground measure.

A new principal stated his perspective on being new and learning in CoP by stating that,

I'm intimidated by it to be honest. My background with evaluation systems was simplistic in nature... this tool is more directional. I do I think that as we go through all the pieces and just make sure 'what it is' and 'what its not' is a way to calibrate together to work on understanding – and building inter-rater reliability.

Another new principal commented on learning in the CoP by reflecting,

Coming together is important I think just because I am so new to it. Just making sure that I am implementing it with fidelity is my greatest concern right now. So I'm learning like how we sat down a couple weeks ago and just talked about what those next steps would be in the progression of if X happens you do this, if Y happens you do this? It's really my path to learning the fidelity of the program.

Yet another new principal reflected on learning in the CoP and the social experience of mediating understanding in the group by stating,

I think the principals that have been here and are familiar with it – they seem to have a good handle on it. They understand it, what it seems to be looking at and assessing: how to utilize it. Comments I've heard (in our group) they pay particular attention to certain areas more than others in

their mind they are weighted more. For the new people, it's good to know this and mediate with them.

The same principal stated that, "...the level of understanding of the various elements of it: I think it has really improved since the beginning of the year and just the opportunity of using it several times and reviewing it and having discussions about it has really helped." A director made a statement that supported the idea that the presentations and learning in the CoP are shared by saying, "It seems like it's getting better, that the meetings are helping: It looks like they're thinking about things more... I think in the end it benefits us because at least we're talking and you find who has different ideas about different things." Another more experienced principal said, "[This] is really going to make evaluation more dynamic for us because it's bringing our group together... so we do have kind of a common standard, common vision, common language to do the evaluation on the instructional practice." Another more experienced principal stated,

I think our conversation about what's actually happening when you are going into a classroom, and then coming out and debriefing about that, because the conversations that take place: that's the value in getting everybody on the same page.

This principal went on to say that, "some of those conversations led to a better understanding for me, for certain parts, and for other people, too... [to] make sure everybody's on the same page."

My research journal makes many references to the learning that took place in the CoP. In my journal I noted:

We talked about the ‘messiness’ of coming to terms with the complexities of evaluation and the richness of the ‘knowledge in the room’.

Participants were very enthusiastic about learning from each other—and coming to terms with the complexities of the teaching standards. Everyone was pretty engaged and there was a high pitch of activity in trying to flesh out agreement on standard two.

Frustration was evident after we crept through the first couple of elements in standard 4. Participants vented, but ended up acknowledging that the work was relevant, and essential. We pushed on and got through all but one element.

In terms of the sequence of Vygotsky Space, appropriation is followed by transformation, the phase where individual ownership of thinking comes into a personal context. This concept was observed in the comment from a director,

Ok... I look at it in two different ways. The first way is in respect to my job and how it affects the curriculum... and the instructional pieces that apply... and just everything through my department that fits into the evaluation tool. When I’m making decisions and policies, I consider how I can tie it back to [the professional teaching standards].

My research journal noted,

[New Principal] was positive about the opportunity to use the new teacher evaluation system, but expressed considerable concern about learning all of the facets of the new program and was concerned about how it might align with other initiatives being pursued.

Another entry noted, “Additional discussion revolved around the use of the teacher evaluation tools and how this fits in with individual resources at the sites.” A related comment by a new principal demonstrated the transformation concept with participants thinking about the learning in their own personal context, “I’m asking myself how I align our instrument with the requirements of common core so that I can understand how to support and inject that language into informing teachers about their instruction and professional growth.”

Another more veteran principal demonstrated his transition to a personal understanding when he described how the evaluation system was fitting into his conceptual view:

As I’ve looked at those things, the other standards inform one another. So I don’t see them as separate; they’re almost a little bit of overlay, kinda like a Venn diagram type of thing, if I were to visualize it. I see those five domains as kind of overlapping in certain areas. Certainly there are separate entities... and I think that’s how our teachers see it, so some of us are saying, “Well, we’ll just focus on the instructional part,” and then others are saying, “I need to look at the whole big picture,” and then still others are saying, “How do I move through the volume of the instrument,

in terms of how many teachers I need to get done?” I am rather just making my focus on the actual evaluation.

The Vygotsky Space concepts of publication and conventionalization were evidenced in the qualitative data as participants commented about how they managed their evaluations in the field and by my observations of them working with the instruments in classrooms and dialoging with teachers and with each other about their evaluations. One director showed evidence of publication when she talked about how her interactions with principals were improving as participants became more conversant in the language of the evaluation tools:

I think it's helped in the conversations I've had with principals, and especially with the teachers, because we have more of a common language... because we're looking at the same tool and saying, "Okay, this is where they could improve or where they need help." And the evaluation did kind of help that because it lays out what they're supposed to be doing and I'm using it, the principals are using it, so when we're talking about what teachers need, or how teachers are doing well, we have that common tool and language.

Another more experienced principal showed a conventionalization of practice with the evaluation system when he talked about his practice,

I really target good instructional practices and use it to clearly communicate to the staff, to score it and to create plans under it to increase the quality and focus of instruction... Each year I pick up a little bit more and little bit more. It really is a different culture or mind set because it's

just not coming to work and doing the same things over and status quo – there is no status quo.

This principal was observed in my researcher's journal to move confidently through the evaluation cycles, and to talk comfortably with his teachers individually in through emails about the expectations and procedures. Another director summed up perceptions of conventionalized changed practice in teacher evaluation,

Now it's interactive, it gives teachers a voice through the year, showing their growth, showing their efforts so they can get the full points or acknowledgment for what they're doing versus just kind of a yes or no – or nothing at all.

**Analysis of leadership.** The third area of the analysis of qualitative data addressed my leadership of this intervention to improve administrator's practice in teacher evaluation. Data from interviews, survey comments, and my research journal helped to inform my understanding of the leadership processes that I put in place during this action research. I began by filtering my observations through the theoretical lenses of Wenger et al. (2002) and Kotter (1996). Since these two theoretical lenses helped inform my work with this innovation, I developed a priori coding based on key concepts from these theories. Wenger et al. (2002) was a key reference in the literature on CoP's, and I used the concept of cultivating the CoP from this literature as an a priori code to capture elements where these strategies appeared in the body of qualitative data. Kotter (1996) was another key reference in the literature, informing my strategies for leading this change to a

standards-based teacher evaluation. I used Kotter's eight steps to leading change to identify additional codes that emerged from the body of qualitative data collected during this innovation. Specific codes derived from Kotter's eight steps were: develop guiding coalition, develop sense of urgency, create short-term wins, and communicate vision. Figure 14 shows these a priori codes derived from Wenger et al. (2002) and Kotter (1996), along with other codes that emerged using grounded theory. The codes shown in Figure 14 appear in the order of frequency with which they appeared in my analysis.

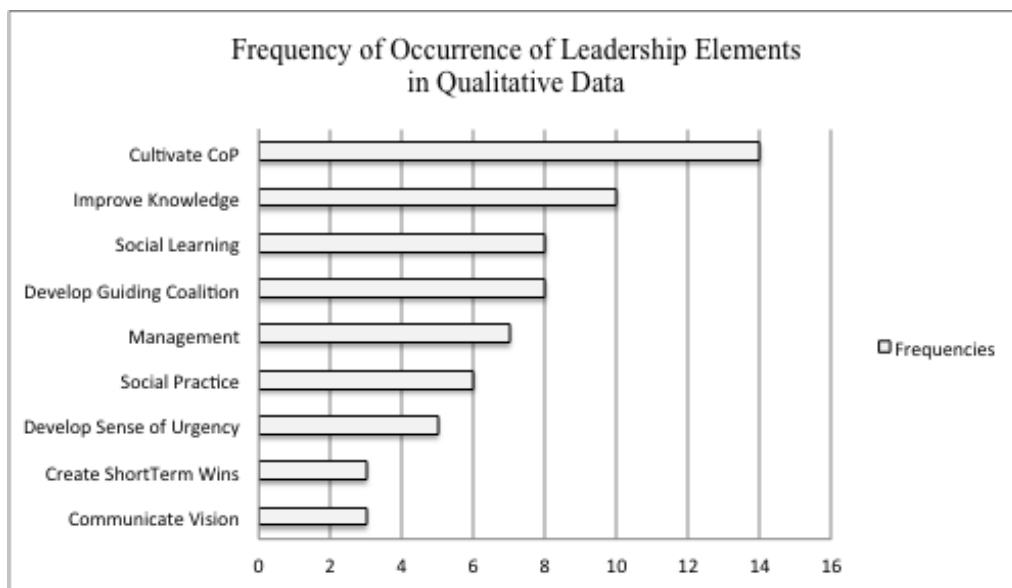


Figure 14. Frequency of occurrence of leadership elements in qualitative data

Figure 14 shows that leadership elements reported in the qualitative data were most frequently categorized under cultivating the CoP. As discussed in Chapter 3, Wenger et al. (2002) described the importance of being conscious of

the elements of the CoP: domain, community, and shared practice. They also outlined an argument for organizations to make efforts to actively and systematically cultivate CoP's in specific ways to improve the benefits for the individuals and the organization itself. The qualitative data from this study documents instances of cultivating the CoP. Data refers to my effort to "design for evolution" (Wenger et al., 2002) in such a way that encouraged the cohesiveness of the CoP. This was accomplished, for example, by bringing the topic of the group's existence to important existing forums, such as the Superintendent's Advisory Council and the Administrative Leadership Team. This helped to establish the CoP's legitimacy, create time for the group to meet during the business week, and ensure that the right people would be attracted to participate regularly.

Wenger's concept of "open a dialogue between inside and outside perspectives" was also evident in the leadership data (Wenger et al., 2002), as was Wenger's concept of boundary brokering (Wenger, 2000). As the committee faced major stumbling points, especially when trying to meet the data requirements established by SB 1040, opportunities were created to bring in a mix of outside resources to help the group get new ideas, fresh perspectives, and much-needed technical assistance. A particular example in this intervention was when the group began to stall in their progress on meeting data requirements for the entire spectrum of teacher roles. At that time, I led the group to invite in resource people from within our professional circles, but not a regular part of the CoP, to participate in our learning activities.



The concept of improving knowledge addressed my efforts to provide direct instruction to the participants throughout the process of formally convening the CoP. Data recorded showed that much of our meeting time was devoted to sharing knowledge. Although the researcher convened these meetings, the discussion was purposefully directed toward drawing on the knowledge in the room. Though much of the presentations about teacher evaluation were planned and presented by the researcher, participants were actively encouraged to present the knowledge they brought to the setting. This type of direct instruction seemed necessary as a means to bring this diverse group of participants up to a similar level of knowledge about the intricacies of teacher evaluation and especially the new level of complexity required by SB 1040's data reporting requirements.

Many examples were identified and coded in my data of leadership efforts to engage the participants in social learning. Activities at regularly scheduled meetings were designed to encourage small and full group participation in learning. Participants were also encouraged, though with only limited engagement of members, to participate in shared editing of documents designed to guide the teacher evaluation process. Within our large group meetings, participants were led to confront each other's perceptions of how to define the professional teaching standards and how to interpret the rubrics that measured attainment of these. This process was reiterated throughout the series of meetings and was extended through electronic sharing of work through emails and shared documents. These shared documents, hosted with Google Docs provided a forum for administrators to edit shared documents between meetings.

Kotter (1996) also included “develop a guiding coalition” in his eight steps to leading change, and emphasized the importance of gathering key stakeholders to play important guiding roles in the change process. This type of activity was demonstrated in my qualitative data, for example, in several meetings with the superintendent, directors, and more influential principals. Each of these meetings was planned to build support for this initiative and to ensure that key leaders were either supporting or openly endorsing the planned activities. The Superintendent’s Advisory Council and the Administrative Leadership Team venues were also used to publicize the involvement of this guiding coalition to help move the CoP forward in their activity.

Management was another leadership factor that was documented in my qualitative data. These data instances were records of my activity to organize the mechanics of the teacher evaluation system and arrange the many opportunities for participants to interact as a CoP and in smaller groups. The high frequency of these activities indicated a sense of the many details that were necessary for me to address in bringing physical and human resources together so that this CoP could function effectively enough to meaningfully address the teacher evaluation initiative. Additionally, management items addressed the activity of solving many short-term problems that were impeding participants’ progress in adopting the teacher evaluation system. Most of these could be addressed relatively quickly through my administrative interventions. Kotter (1996) in his work on change leadership (see Figure 2) referred to “removing obstacles” as the fifth step in facilitating change. He said that it is important for leaders to remove obstacles to

empower people to execute vision. As the data shows, basic management tasks and removing obstacles were a substantive part of my leadership efforts.

The code “social practice” refers to events in this initiative where participants were led to practice teacher evaluation while peers shadowed their work and helped to negotiate the outcomes with the supervising administrator. Although I originally planned for these practice sessions to include the supervising administrator, myself, and one other administrator, time and circumstances allowed for me to be the only shadow in these experiences. Principals were especially hesitant to have this occur, but when reflecting on their learning, they often cited these experiences as positive means to improving their skills as teacher evaluators. One principal stated, “Requesting visits to our campuses to work on evaluations together and conduct some joint observations was what most helped me get a better feel about how to rate teachers and run the process.” Another principal stated a contrary opinion that these paired evaluations were not enough to make a difference.

Two other elements from Kotter’s (1996) work on leading change were also part of the data collected about leading this initiative. “Create short term wins” and “communicate the vision” were concepts that were evidenced in the record of my leadership. There were several instances, mostly in the Administrative Leadership Team meetings, where I either personally announced, or arranged for others to announce, milestones (short term wins) that were reached by the CoP in moving to the next level of teacher evaluation and/or making progress in meeting the requirements of SB 1040. I also arranged several

occasions where the vision of teacher evaluation could be positively communicated to the CoP and to the greater school community. Again, I worked within the Superintendent's Advisory Council and the Administrative Leadership Team to craft messages for administrators and teachers designed to positively express our vision for positive outcomes from our work with the new teacher evaluation systems.

## **Chapter 5 Findings**

Throughout the previous chapter, I presented data from survey instruments, my research journal, and notes and transcripts from meetings and interviews. In this chapter I will use the analysis from Chapter 4 to synthesize assertions that respond to the research questions originally posed in Chapter 1. The assertions presented in this chapter are the result of triangulation of findings from those data sources that best informed each research question. These assertions were also informed by member-checking techniques used to provide confirming as well as disconfirming results to build on the reliability of each assertion (Lincoln & Guba, 1985).

### **Data Triangulation**

Triangulation of findings was used as a method to identify patterns of data convergence and corroboration from sources of data. Gay et al. (2009) define triangulation as “a process of using multiple methods, data collection strategies, and data sources to obtain a more complete picture of what is being studied and to cross-check information” (p.377). By triangulating the data I balanced the weakness and strengths of the quantitative and qualitative instruments and their consistency (Fraenkel, & Wallen, 2005). Figure 15 shows my research questions and how I used data triangulation to establish reliability in answering my research questions. Complementarity among data was used to find instances of elaboration, enhancement or illustration and clarification of the results (Johnson & Onwuegbuzie, 2004).

Research Questions	Data Sources				
	SoCQ & AES Survey Ratings	SoCQ & AES Survey Essays	Interviews	Meeting Transcripts	Reflective Journal
1. What concerns do administrators within my district have about effective implementation of our new teacher evaluation system and how and to what extent will those concerns change as our CoP learns, practices and engages in discourse?	QUAN	QUAL	QUAL	QUAL	QUAL
2. What will administrators say, do and feel as our CoP attempts to a. develop a common understanding of the professional teaching standards, b. increase inter-rater reliability on teacher evaluation instruments c. understand the purpose of our TES to improve professional practice	QUAN	QUAL	QUAL	QUAL	QUAL
3. How do I lead this process of change?		QUAL	QUAL	QUAL	QUAL

QUAN=quantitative      QUAL=qualitative  
*Figure 15.* Research questions and data sources

### Validity of the Data

The primary tests for confirmability or trustworthiness of the data were triangulation, thick description of qualitative data, and member checks (Erickson, 1986; Miles & Huberman, 1994). Data were compared among and between sources as a means to enrich, challenge, and validate findings as they occurred. Several data analysis strategies were applied to strengthen the validity of the study’s findings. As noteworthy findings emerged from this study, they were validated through a member check process (Lincoln & Guba, 1985). Findings were taken back to participants to compare my interpretation of the information to their own understandings about the innovation (Miles & Huberman, 1994).

During this process confirmations and disconfirmations of my findings were recorded and further discussed with each participant. This dialog helped lead to adjustments in the findings to more accurately reflect actual opinions and perceptions.

### **Changes in Administrators' Concerns: Research Question #1**

The first research question I posed was, "What concerns do administrators within my district have about effective implementation of our new teacher evaluation system, and how, and to what extent, will those concerns change as our CoP learns, practices and engages in discourse." Data about participant concerns were gathered during the entire course of this innovation. Quantitative analysis of the SoCQ and AES surveys, along with qualitative analysis of survey essays, interviews, meeting notes, and my research journal informed this topic.

**Concerns frequent and only slightly changed.** In spite of the extensive time our CoP spent learning, practicing, and engaging in discourse about teacher evaluation, there appeared to be little or no change in participants' number or level of concerns. The SoCQ results pre- and post-survey showed changes in some areas for each participant, but overall levels of concern about the adoption of teacher evaluation remained relatively unchanged. Utilizing the analytical methods prescribed by George et al. (2006) for group analysis of concerns profiles, there was very little in the way of consistent or predictable patterns of change in the concern levels for the group as a whole, or for any of the individual profiles when comparing pre- to post-innovation scores. Review of the qualitative data also failed to show any lessening in the numbers or levels of concern about

adoption of the teacher evaluation system. The exception to this finding, lessening of concerns about inter-rater reliability, will be addressed later in this chapter.

Review of the qualitative data revealed confirming evidence of relatively unchanged levels of concern from pre- to post-innovation, but also provided more detailed information about the number and level of concerns. Concerns about teachers' attitudes, time for teacher evaluation, and the difficulty of adapting evaluation to different teacher roles remained relatively the same over the course of this innovation. Two notable exceptions were revealed. One was a marked increase in concerns about the need for additional administrator training, and the other was a similar marked increase in concerns over the quality of the instrument. These exceptions are important to this discussion and will be addressed next.

**Concerns about the evaluation system.** Concerns about the evaluation instrument itself appeared to grow over the course of this innovation. The AES pre- and post-survey responses about the construct that measured the extent of positive perception about the teacher evaluation tools were elevated just slightly from pre- to post-survey, going from 3.29 to 3.39 (both in the *unsure* range of opinion). Information from the qualitative data, however, shows strong levels of increased concern about the teacher evaluation instruments. The number and levels of concerns noted in their voices about the teacher evaluation instruments in the qualitative data grew from seven instances in the pre-innovation to 17 in post-innovation measures. Participant comments were characterized by comments that as administrators had more opportunities to learn, use, and share the meaning



of the indicators in our CoP, doubts about their quality emerged. Because new instruments were introduced in the middle of the innovation, doubts about the quality of our adopted system became elevated. As administrators became exposed to instruments that were easier to understand and more readily aligned to other reform efforts taking place in the district, their concerns about our current adoption grew.

Member checks about these findings revealed that participants' perceptions changed over time. One said that when first looking at the instruments, "It all looked great and relatively easy to use," but that when it came to interacting with teachers and colleagues on the certain indicators, "...it starts to fall apart – looks good in theory, but in practice you start to feel a little shaky in making claims about teacher performance in some of the areas."

Other participants referred to the cause of their growing doubts about the instrument by referring to other instruments to which they were recently introduced. One administrator said, "When we bought into this, we looked at what was available, but now we are seeing other possibilities." This administrator went on to say, "The [other evaluation system] is better aligned with what we actually see in the classroom – and what we're looking for [our reform agenda]."

**Concerns about inter-rater reliability.** Concerns about inter-rater reliability were the exception as the only area of concern that dropped. The AES survey data revealed a significant change in survey ratings on this area from pre- to post-survey and was the only construct to show a significant positive change. The level of agreement on this construct of survey questions increased

significantly, with a large effect size, from a mean of 2.82 on the pre-survey to 3.29 on the post-survey. This clearly indicated administrators felt more confident in the inter-rater reliability because of being a part of a CoP and having experiences learning, practicing and engaging in discourse about teacher evaluation. Although there was positive growth in confidence in our inter-rater reliability, the results were not overwhelming. Both pre- and post-survey ratings fell in the *not sure* range.

The increase in confidence about inter-rater reliability was confirmed in the analysis of interview and observational data recorded in my qualitative analysis. Member check interviews also explored the finding that concerns about inter-rater reliability were somewhat relieved after the first phase of the innovation was completed. Participants did not experience a complete reduction of their concerns in this area, however. They acknowledged that the time spent working with their colleagues helped make them less concerned about problems with inter-rater reliability. One participant said, “We pretty well worked through the rubrics and should be much better in agreement now.” On the other hand, another participant stated, “It’s still a pretty big stretch to say we’re going to see things the same way on some of these indicators.”

**Concerns about level of training.** The qualitative data revealed high frequency of concerns that administrators were not feeling sufficiently prepared to adequately implement the teacher evaluation system. Specifically, administrator concerns about the need for additional training were documented as the most frequently mentioned concern on both in the pre- and post-survey. These concerns

were also documented 17 times on the pre-innovation interviews and 24 times on the post-innovation interviews. I used member checks to learn more about this phenomenon and found confirming evidence when conducting interviews with the participants. About half of the participants, and mostly less experienced administrators, reported that once they began to learn and practice more with the evaluation tools, they realized the complexities of teacher evaluation and became more overwhelmed with management and logistics of the task. Three of the new principals were also especially concerned whether the high stakes teacher ratings they made would match that of their predecessors from just one year ago. Some of the more experienced administrators, on the other hand, maintained a sense of confidence in their use of the evaluation system. While expressing concern that some of the rubrics were not clear enough, they expressed confidence in their ability to use the ratings fairly and consistently to help teachers get good feedback on their teaching practice. These more experienced administrators expressed less need for additional training.

**Teacher evaluation concerns about alignment and integration.** The SoCQ data showed evidence of a relatively unusual and similar pattern among all administrators. Their concern patterns were consistently elevated at “Stage 0,” the indicator showing how interested and engaged an individual is with the innovation when compared to the respondent’s universe of professional concerns. This finding suggests that participants did not place teacher evaluation high among all of the other initiatives and challenges presently being pursued (George et al., 2006). I conducted extensive member checks on this facet of my findings in

the form of interviews with each participant. I began by asking how they were doing with evaluations, and then asked about the level of importance they assigned to the new teacher evaluation system. They all agreed that the evaluation system was important, saying things like, “it is extremely important that we get this right,” “the stakes are high,” and, “there is a lot of attention to these scores: it can be stressful with some teachers.” When I followed-up by asking about how teacher evaluation aligned with their larger universe of strategic initiatives for school reform, I got much deeper into the concerns they were facing with teacher evaluation.

Many of the administrators, and especially the principals, expressed that however important teacher evaluation was, it most often ranked below more pressing concerns. One principal said, “we have a lot going on, and as much as I would like to do this (teacher evaluation) right, I have to make choices about what I do with my time, and there are other things I can do that will move the needle.” When asked to elaborate, this principal talked more about pedagogic coaching, training teachers on lesson plan development, and learning new strategies to help teachers analyze and act on formative achievement data. Other concerns revolved around doubts about how well matched our teacher evaluation system was with current district-wide education reform initiatives. This theme originally emerged in the qualitative data, especially in my research journal entries where I recorded summaries of lively discussions taking place in our CoP meetings, and was also evidenced to a lesser degree in the post-innovation interviews and survey comments. When I explored this topic further during member check interviews,

participants confirmed concerns that the teacher evaluation system and the reform initiatives were sometimes difficult to coordinate. Most felt that it would be easier to manage if the adoption of standards-based teacher evaluation was better coordinated with our overarching district reform initiatives. They noted instances where the same or similar professional development activities were repeated and framed in varying ways between the initiatives. They also referenced other evaluation instruments they were exposed to through our reform initiative and cooperative projects with other districts. They cited these as alternatives that could possibly be more easily integrated with our current reform initiatives.

### **Administrator Learning: Research Question #2**

The second research question, “What will administrators say, do, and feel as our CoP attempts to a) develop a common understanding of the professional teaching standards, b) increase inter-rater reliability on teacher evaluation instruments, and c) understand the purpose of our teacher evaluation system to improve professional practice,” was focused on administrator learning in the sociocultural context of our CoP. The question was informed by quantitative analysis of the AES and SoCQ surveys, and qualitative analysis of survey essays, interviews, meeting notes, and my research journal.

Much of the scheduled work of our CoP centered around our common understanding of the professional teaching standards and how this understanding informed our ability to reliably rate teachers on these standards, regardless of which administrator is rating. Out of this comes our potential ability to affect positive change through using our teacher evaluation tools to for the purpose of

improving professional practice. This process was planned and subsequently analyzed using the theoretical lens of Vygotsky Space.

Our CoP spent four sessions averaging just less than three hours each reviewing the five standards and the twenty-five elements distributed among the standards (see Appendix A). The AES survey construct on inter-rater reliability was one measure of how our administrators perceived their learning progress in this area was reflected positively with a significant improvement in the level of administrator confidence about our inter-rater reliability from pre- to post-innovation. There were 2.6 times as many *agree* and *strongly agree* responses on inter-rater reliability confidence questions on the post-survey compared to the survey given before our CoP engaged in these activities. As was noted earlier, however, this measure improved only to the extent that it still averaged in the *not sure* range of the survey opinion scale. The qualitative data confirmed these survey results and were addressed in the findings in Chapter 4. There were less than half as many comments questioning inter-rater reliability in the post-innovation interview as there were in the pre-innovation interview.

Administrators expressed reasons why they were feeling more confident in the inter-rater reliability among their peers, and these were captured in the qualitative data with short essays from the surveys, interviews, and entries in my research journal that summarized and reflected on our group and individual learning sessions. Participants that made expressions of confidence in the improvement of our inter-rater reliability attributed improvement to the learning done within our CoP primarily at the meetings scheduled for that purpose. There

were also several administrators who noted that the district-wide reform initiative as having an effect on their growing level of common understandings about teacher evaluation, citing elements of the reform initiative that helped administrators and teachers come to common understandings about effective lesson planning and pedagogy in the context of our setting.

On the other side of the growth in confidence shown by the survey, there still existed a sizeable contrary opinion still unsure about our progress to learn what the standards mean and to improve our inter-rater reliability. This level of doubt is reflected in the AES survey where even with the improvements from the pre-survey, the post-survey results still averaged *unsure*. The qualitative data also reflected occurrences of doubt about our CoP's ability to move forward productively to come to a common understanding. Again, this lower level of confidence decreased post-innovation compared to pre-, but remained a strong minority opinion. In one individual circumstance, both qualitative and quantitative data sources point to less confidence in our CoP's capacity to learn and become consistent with the standards and rating. In this particular case, the individual expressed discouragement in the perception that peers may be "shortchanging the teachers by minimizing the process." Another questioned the fidelity of evaluation if there is not time to adequately train administration. As already discussed in reviewing participant concerns, the adequacy of time for training is a prevalent issue. There was little doubt about participants' overwhelming sense of need for our CoP to continue to pursue this learning

process together, and considerable concern over the lack of time to make that happen.

On the SoCQ survey, four participants' levels of concern on S3, the indicator associated with concerns about management, time and logistical aspects of the innovation (George et al. 2006) became noticeably more elevated at the post-survey. This was perhaps as a result of learning more about the complexity of the systems. These higher levels in S3 may be a reflection of participants' struggle with how well the teacher evaluation changes meshed with other newly introduced innovations, or perhaps were a reflection of their greater knowledge of the complexities of the teacher evaluation system creating greater concerns about managing these intricacies for each teacher. Other indications in the data appeared to confirm this notion that as participants learned more about the innovation, some became more concerned about the level of skill and knowledge necessary to become proficient in standards-based teacher evaluation. As discussed earlier, qualitative data revealed increased frequency of concerns indicating that administrators needed more training, and some administrators cited the concern that the evaluation was much more complex than it appeared at first glance, and that the more they discussed the standards and the rubrics, the more they felt they needed to learn and practice.

I conducted member check conversations on this topic that yielded a roughly equal proportion of confirming and disconfirming evidence about whether increased discussion and learning led to an awareness that more learning was necessary. When participants were asked whether our learning activities



made them feel like we were just scratching the surface four agreed and made comments such as, “We have a long way to go,” and “There are complexities that can get a little overwhelming.” Five others were more optimistic and expressed confidence that our CoP learning activities helped prepare them to create accurate and meaningful evaluations. One administrator stated, “Our teacher evaluation system supports professional practice by specifying expectations for the teachers through the evaluation rubrics.”

In an attempt to better understand administrators’ range of perceptions of the progress made in learning about the teaching standards and inter-rater reliability – from confident and positive to negative and doubting – I conducted member checks on my findings. I asked for opinions about my findings regarding inter-rater reliability among peers, and how inter-rater reliability was affected by the learning experiences over the course of this innovation. Similar to the results of data already gathered, participant responses showed appreciation of the learning experiences. One participant said, “I don’t worry about our inter-rater reliability. We’ve spent time together to work out mutual expectations. We’ve compared that to lessons and situations and we seem to be agreeable – on the same page with all of this pretty much.” Administrators ranged from very confident and positive to doubtful and negative about the potential for learning and success in this pursuit of learning and development. There was a strong sense that as more learning opportunities transpired, administrators become more confident and trusting of the evaluation systems they reviewed with their peers. There was also a strong sense in the data that the number and degree of positive

perceptions of this process are growing over time. Member checks revealed that some participants that were more likely to question our progress wanted more time and experience with administrative peers to negotiate common understandings and build confidence in our CoP's ability to navigate through the substantial body of learning necessary. Clearly, all participants confirmed that there was still much work to be done.

As discussed in Chapter 4, I analyzed participant learning through the lens of Vygotsky Space (Harre, 1984; McVee et al., 2005) as a means to better understand participant learning on this complex topic. Qualitative analysis of my research journal, interviews, and short essays from both surveys were used to quantify frequencies of learning behaviors matched to the sociocultural learning stages of Vygotsky Space. The frequencies revealed in this data demonstrated a general pattern of diminishing level of qualitative evidence for each of these categories as they move from the more basic learning stage of appropriation to the more advanced learning stage of conventionalization. This pattern is what I expected, based on my assumption that participants' relatively new exposure to learning about standards-based teacher evaluation would result in a higher concentration of activities in the appropriation and transformation phases of Vygotsky Space (Gavelek & Raphael, 1996). In fact, the pattern of data appeared to illustrate participants' relatively early stages of learning and adoption of this teacher evaluation system, and again reinforced the notion that this learning process is complex and not easily accomplished in a brief series of learning activities.

### **How I Led This Process of Change: Research Question #3**

My third research question, “How do I lead this process of change?” was answered by analyzing multiple sources. Most of the data collected to answer this question were taken from my research journal and reflects my personal understanding of this innovation. Data from interviews, survey essays and meeting notes were also used as supplementary means to answer this question. In order to validate these findings I reviewed the body of research data to make two broad assertions that follow. These assertions were validated by follow-up member check conversations with several individuals who formed the guiding coalition of our CoP (Wenger et al., 2006).

My first assertion is that one key determinant of success of my leadership of this initiative was to bring attention and a sense of urgency to the mission of our CoP to help administrators become more effective with teacher evaluation. The initial challenge for this initiative was whether or not these activities would even occur. Our teacher evaluation was already somewhat established and the district was at a critical point in launching a comprehensive district-wide reform initiative. This reform agenda was already underway with our district administrators and our three veteran principals, but in consideration of three new principals coming in we made the decision to focus the first part of the school year on intensive administrator training to move the reform initiative forward.

The reform agenda was a necessary objective for the district to pursue, but its existence created a formidable obstacle for my plans to focus this innovation for our teacher evaluation CoP. This forced me to think about strategies to

proceed. Using Wenger's important concept of boundary brokering, the process that "introduces elements from one practice to another" (Wenger, 2000, p. 236), I assumed the role for our teacher evaluation CoP of a boundary broker by making connections with other administrative CoPs to help establish our objectives in the context of their interests and activities. I brought conceptual information about the urgency to tackle issues of teacher evaluation learning, inter-rater reliability, and compliance with legislative requirements (SB 1040), and negotiated priorities with the other administrative CoP's to ensure that some time would be carved out, by necessity, for the teacher evaluation CoP. I began with the Superintendent's Advisory Council CoP (SAC), our group of district administrators that report directly to the superintendent. I brokered priorities to get the teacher evaluation CoP on the agenda for the Administrative Leadership Team CoP (ALT), a group with broader membership that includes all of our administrators. I found, however, that it was difficult to get full support and buy-in to the priority of the teacher evaluation CoP without brokering with the CoP that led our reform initiative. At the time this occurred, however, I was too far outside of the reform CoP to act in the role of a boundary broker. Instead, I worked with another district administrator from our teacher evaluation CoP to act in the role of boundary broker between these two communities of practice. I met with less success with this transaction and did not have as strong an alignment as I would have liked among all three of these CoPs. It was enough, however, to successfully move our CoP's agenda from SAC to ALT, and resulted in enough of a commitment for our CoP to pursue this agenda.

These brokering activities involved the use of several of Kotter's (1996) strategies to lead change and to create the opportunity for the evaluation CoP to have the time and support to continue. My work with SAC was an opportunity to build a guiding coalition to help carry this innovation forward. This guiding coalition began within the superintendent's advisory, but evolved to include one key district administrator and two principals who were articulate in the language of teacher evaluation and helped carry positive messages forward within and outside of our CoP. There were also naturally occurring opportunities to build a sense of urgency. The requirements of SB 1040 were indeed looming, and it was no exaggeration for me to advertise there was little time to prepare to meet these ambitious requirements.

My second assertion is that there was far too little time to move our CoP's agenda forward to the extent that we had anticipated. Even though the group's legitimacy was well established and time was set aside for regularly occurring meetings and fieldwork, it was not enough. With the time we did have, administrators were enthusiastic participants in the social learning activities and their dialog with each other was spirited and meaningful, but even so there was an overarching sense that we weren't making the progress needed to come to a strong consensus of understanding about the process of teacher evaluation. Additionally, there was not enough time for fieldwork. It took longer than anticipated to get administrators comfortable doing evaluations with supervisors and peers present. As a result these weren't done to the extent I would have liked. Participants commented in member check conversations that however beneficial

the activities were on the whole, they were just a start. One principal said, “We’ve barely started, so it’s really hard to say what we’ve accomplished.” Many agreed that that the group needed more time to arrive at better levels of mutual understanding, but all also acknowledged they could not afford much more time outside of their day-to-day responsibilities to attend to these activities.

Overall, leadership of this learning initiative was difficult to initiate in the face of competing interests and the limited resources of time and energy to focus on our goals. There were huge challenges to bring the initiative far enough to the fore to legitimize and initiate the process. Once those hurdles were surmounted, I was faced with the challenge to make a meaningful impact on participant growth and development within the constraints of limited time and competing interests. Overall, I was buoyed by the perception that success was enjoyed to the extent that the CoP was positively moving in the right direction. I was disappointed only to the extent that due to time and competing interests, we didn’t get as far as hoped.

## **Chapter 6: Conclusion**

I began this action research with the purpose to ensure and improve my school district's ongoing implementation of standards-based teacher evaluation. Our district had already made a good start in the first year of our adoption of this innovation, but faced many challenges going into our second critical year. We had new members on our administrative team, newly legislated requirements for teacher evaluation that would come into effect soon, and an array of competing interests and activities that were looming as potential obstacles or opportunities for progress for this ongoing implementation. My research purpose was to sustain and enhance my administrative CoP to plan and implement strategies to ensure our new evaluation tools were used fairly, consistently, and effectively.

Throughout this process, and in my role as an action researcher, I pursued answers to research questions by focusing on identification and clarification of participant concerns. I sought to see how these would change as we learned, practiced and engaged in discourse about teacher evaluation. I also sought to understand what administrators would say, do, and feel as we pursued a common understanding of the professional teaching standards, attempted to increase inter-rater reliability among our CoP, and understanding of the purpose of our teacher evaluation system in the context of a mission to improve professional practice. Finally, I sought to learn more about my leadership role relative to this innovation.

## **Achieving the Purpose of This Action Research**

Whether or not this action research innovation achieved my purpose of improving teacher evaluation is more a question of determining what incremental progress was made, rather than that of declaring, “yes we did it,” or “no we didn’t.” I began to answer this question by comparing my initial vision of what we could accomplish against later assessments of what was actually achieved during the course of this action research innovation.

During the initial planning stages of this work with our teacher evaluation CoP, I envisioned a set of accomplishments that in retrospect were not entirely achieved. In short, clear progress was made, but much work remains. I initially expected that we would have moved more quickly in developing our common understanding of the professional teaching standards, and would have been left with more time to for the advanced work of social learning in the field. We lingered on the professional teaching standards, however, as a matter of choice and necessity. The group began to deconstruct the standards and the rubrics and found the work more challenging than expected. Long and engaging dialogs ensued and individuals were hesitant to simply move on without clear consensus about meaning for each of the many standards, elements, and rubrics in our teacher evaluation system.

The group was also sidetracked, in some sense, by exposure to other teacher evaluation methods. Our CoP participated in extensive professional development provided through a district-wide reform initiative focused initially on instructional leadership. These activities dovetailed to some extent with our



teacher evaluation CoP activities, and participants credited this as helping improve their confidence in our common understanding of effective pedagogy. Within these reform initiative activities, members of the teacher evaluation CoP were exposed to other evaluation instruments that many found easier to understand than what we had adopted. Another teacher evaluation instrument, developed by a cooperative of school districts and designed to meet the requirements of SB 1040, was also introduced to our teacher evaluation CoP when we were invited to join that cooperative. Administrators also perceived measures of professional teaching standards on this instrument more closely aligned to our reform initiative than our currently adopted teacher evaluation system. This realization caused our CoP to engage in extensive discussions regarding the potential to redesign our new teacher evaluation system to something that would more closely align with our overall reform efforts.

Another factor that changed the course of our progress was the need to focus on important questions about compliance with SB 1040. Provisions of this legislation intended to balance summative teacher performance assessments with a prescribed portion of student achievement outcomes would soon become a requirement for every teacher's statutorily required annual summative evaluation. Strategic discussions about how our district would meet these requirements were a necessarily recurring item of discussion. Since this requirement was set to go into effect for the coming school year and new developments on this topic were brought to our CoP through several different channels, this topic occupied a big part of our CoP's formal and informal agendas.

Regardless of the pace of the innovation, participants made incremental progress in mapping out common understandings of our teaching standards. As documented in Chapter 4, this work resulted in increased confidence in inter-rater reliability among our CoP. Some fieldwork was done with small groups in the CoP and participants reported that these field activities helped them grow in their professional practice. They were engaged in these opportunities to work together in the field, and noted that it helped build their common understanding of targeted pedagogy. This progress was hindered to some extent, however, by an initial hesitancy from principals about having other administrators shadowing them in classrooms. They did, however, seem to get more comfortable with the practice after doing shorter walk-through evaluations with just me present, and will probably be ready to do more extensive fieldwork with their peers in the future.

Overall, our administrative team took positive steps forward in their intellectual understanding and practice with standards-based teacher evaluation. They engaged enthusiastically in the group discussions about evaluation and challenged themselves and their peers about using these tools to improve professional practice. The process of gaining a common understanding in this area was complex and may lead our CoP to reconsider or modify our current adoption to achieve closer alignment with our current reform initiatives.

### **Answering the Action Research Questions**

The action research activities yielded a sufficient body of data to address the research question posed for this study. Data from pre- and post-innovation surveys and a repeated set of interview questions posed to all participants was

applied as planned to the research questions posed. Additionally, my research journal was used as a means to record qualitative data throughout the innovation. Notes were kept of meeting records and observations of participants as they interacted with me and with each other. These notes were invaluable and were also applied to the research questions.

The sum of this data was triangulated to inform each of the research questions and to begin to reveal findings. Once those answers began to develop, I followed with a liberal use of member check conversations. These member checks, framed as casual conversations, were very helpful as a means to put participants enough at ease that we could comfortably talk about our work in a less formal setting. Though there were limitations that will be discussed later in this chapter, the data appeared to adequately address the overall scope of my research questions.

### **New Learning About Implementing Teacher Evaluation**

The process of engaging our teacher evaluation CoP in this action research innovation helped me develop some insights about learning to use a teacher evaluation system. During our initial adoption of our standards-based teacher evaluation system, a group of stakeholders reviewed about a dozen different packages of teacher evaluation instruments. Individuals were provided with copies of three finalist evaluation systems, and after the opportunity for individual review and group discussions, the list was narrowed to the adoption we presently have.

Since the initial adoption, teachers and administrators trained with the new system over the summer and fall, and then practiced with it for one year. This was followed the next school year with our teacher evaluation CoP spending three intensive months in this action research innovation to explore the system in greater detail, while at the same time we began the second year with the system. I anticipated the result of these intensive learning experiences would have resulted in a greater reduction of concerns from administrators about teacher evaluation. Instead, it appeared that the more the administrator group delved into the details of the evaluation rubrics, the more frustrated they became with the instruments.

I found two distinct reasons for these elevated concerns. First, administrators stated they looked at some of the indicators and rubrics in more simple ways after first being introduced to the evaluation system, and were satisfied in these less complex interpretations made when first reading and using the evaluation system. After discussing these further with their peers, they found that some of these less complex interpretations were more difficult to define in professional dialog with other administrators. As this was discovered, our CoP felt the need to embellish the descriptors to arrive at something that could more reliably be measured and understood. This process was difficult and made the evaluation system seem far more complex and difficult to use with teachers who had not been part of these same discussions.

The other reason concerns became elevated as our CoP explored our evaluation system more deeply revolved around issues of alignment. Participants were grappling with how to effectively use our evaluation system at the same time

they were deeply immersed in a district-wide reform initiative that addressed effective pedagogy and instructional leadership. They worked hard to make connections between the teacher evaluation system and what they were learning in the reform initiative. Although our CoP found ways to build crosswalks between the two initiatives, participants began to express growing concerns that translation efforts were confusing and counterproductive. When exposed to an evaluation system that was already much more closely aligned with the professional development of our district-wide reform, they were immediately attracted to this.

### **Leadership Insights**

In my continuing work administering teacher evaluation system, I learned the power of applying theoretical frameworks to leadership initiatives. Concepts from Kotter's change theory (Kotter, 1996) helped me plan strategically to ensure that implementation was properly attended to throughout the implementation. These strategies were applied before this action research began and were continued throughout this action research innovation. It was tempting to fall into a more usual pattern of managing change leadership, simply by trusting my experience and instincts. By referring to change leadership theory, however, I was able to gain additional insight and made strategic choices that otherwise would not have occurred to me.

Change theories also helped me understand and plan for individual adaptations to change. Rogers' diffusion of innovations theory (Rogers, 1995) was helpful as a reference to informally assess and better understand individual

behaviors. I was able to reflect of various behaviors, including occasional resistance, or sudden acceleration of progress, in terms of progression along Rogers' ordered attributes. If a member was not attending to our work, or was resistant, I had the opportunity to review the administrator's progress in terms of the Rogers' categories of adopters and see if that classification system could help me better understand the individual, and create groupings among administrators.

The Concerns-Based Adoption Model (Hall et al., 1973) served as a more formal means of assessing how each individual progressed in the adoption of the teacher evaluation system, and also provided data for groups and subgroups. As seen in Chapter 4, these data provided some useful insight about how teacher evaluation fit in the larger perspective of our CoP. This data helped inform my thinking, and led me to conclude that alignment might be a primary issue for this adoption. The limitation of applying this theoretical tool, however, is that two data points are not really enough to understand group progression through the seven stages of development. The opportunity to administer the SoCQ with greater frequency over a longer period of time may have produced more distinguishable patterns of concerns for interpretation.

My leadership of this innovation was also informed by social learning theory (Vygotsky, 1978, Daniels, 2001). Application of these principles helped me create active, contextualized learning experiences where the adults, all very well-accomplished instructional leaders, could take advantage of the knowledge and experience already in the room and participate in constructing deeper levels of knowledge through social negotiation (<http://www.learning-theories.com/>

2011). Vygotsky's social learning theory helped me understand the importance of planning dynamic learning contexts where our administrators played an active role in their own learning and the learning of others in a collaborative learning environment (Wertsch & Sohmer, 1995). Wenger et al. (2002) also played a key role in helping me understand sociocultural learning and the power of working within an existing CoP, and creating conditions that helped to cultivate and strengthen the foundation of our teacher evaluation CoP. The concept of Vygotsky Space was also useful as a means to understand progression of administrators as they brought ideas to the table, adjusted them in the context of what they learned from their peers, made new ideas their own, practiced these in their own setting, and began to conventionalize new practices to create new norms that will continuously inform and refresh this cycle of social learning.

### **Limitations of This Action Research**

Smith and Glass (1987) review major threats to validity of research. Based on their analysis, I have identified several validity threats that may have impacted this study. Because of the small sample that included all administrators of our district that conduct teacher evaluations, the results of this study may not be generalized. Participants were members of a specific setting and do not represent a broad cross-section of the population of school administrators. Because of the limitations of this sample, findings are not meant to be representative of the general population of school administrators.

Experimenter effect may have also affected this study. During the action research innovation I served in the role of a district administrator and held a

degree of supervisory oversight over the principals. I was also responsible to oversee the successful implementation of all teacher evaluation activities for the district. My strategy to overcome this effect was to confront the issue openly by stating my purpose to get the most honest and frank responses on all the topics we covered. I also took full advantage of principals' expertise and liberally assigned leadership roles in the process to mitigate a sense of hierarchy in these activities. Embedding our activities in social learning contexts may have been a means to minimize my role and emphasize practitioner's knowledge and experience in using the evaluation tools. The experiment effect may also have been minimized by extensive member checks where I had the opportunity to engage in candid follow-up with individual participants and gave direct permission for open feedback, positive or negative.

Finally, instrumentation threats could have been in play with this study since I was the sole analyst for the data collected from this study. Although I tried to take special care to code responses objectively, and to use member checks as a means to confirm or disconfirm findings, I acknowledge that my personal perspectives may have influenced how the data was analyzed. In an attempt to mitigate this effect, I used member checks liberally as I begin to develop themes in my findings. The results of those member checks helped me check the validity of my findings as they emerged.

### **Implications for Future Research**

This action research covered a slice in time with administrators grappling with the challenge of becoming expert users of a new standards-based teacher



evaluation system. The study covered the scope of a single semester spent learning together with this administrator CoP. My next steps as an action researcher will be to review the findings of this study and use the outcomes in such a way to frame new research questions for our continuing work to improve our teacher evaluation innovations. These new research questions will continue to be framed by new and evolving challenges that have grown out of our work. Maintaining my role as action researcher I may choose to pursue the following research questions as my work continues with this administrator CoP:

- How, and to what extent, will administrators' understanding of a misaligned teacher evaluation system affect their efforts to reform this system?
- How will our administrator CoP design ongoing professional development for its members, and in what form will those professional development themes be advanced for use with teachers?
- How will teacher engagement in the process of teacher evaluation reform affect the administrator teacher evaluation CoP?
- What concerns do administrators have about their ongoing work with teacher evaluation, and how will those concerns change as our learning and development continue?

Ideas from this action research could also be expanded to other settings, and with larger groups of participants. Findings from this study demonstrated an array of administrator concerns and insights about this work with teachers. A better understanding of these perspectives could be important to inform future

research and policy decisions. School administrators have the unique vantage point of ongoing experience implementing various teacher evaluation systems in a variety of settings while balancing responsibility for their schools and districts. This body of knowledge and experience does not appear to be well documented in the literature. Direct observation and feedback from practitioners may help inform a better understanding of the realities of making teacher evaluation an effective element of school reform strategies.

### **Closing Thoughts**

This action research innovation ended with the production of this dissertation. The learning process of developing a common understanding of our professional teaching standards, increasing inter-rater reliability on teacher evaluation instruments, and understanding the purpose of our teacher evaluation system to improve professional practice will by necessity continue. This action research will move forward outside the bounds of this dissertation as our CoP will inevitably continue to grapple with creating a teacher evaluation system that meets our needs by helping teachers improve their effectiveness with students.

## REFERENCES

- Anderson, G. L., Herr, K., & Sigrid Nihlen, A. (2007). *Studying your own school: An educator's guide to practitioner action research* (2nd ed.). Thousand Oaks, CA: Corwin Press.
- Borman, G., & Kimball, S. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close student achievement gaps? *Elementary School Journal*, 106(1), 3-3.
- Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models*. Princeton, NJ: Educational Testing Service.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Crawford, K. (1996). Vygotskian approaches to human development in the information era. *Educational Studies in Mathematics*, (31), 43-62.
- Creswell, J. W., & Plano Clark, V. L. (2006). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Daniels, H. (2001). *Vygotsky and pedagogy*. New York, NY: Routledge Falmer.
- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. (2010). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35-39.
- Darling-Hammond, L. (1990). Teacher evaluation in transition: Emerging roles and evolving methods. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers* (pp. 17-32). Newbury Park, CA: Sage.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1), 1-44.

- Davis, D., Ellett, C., & Annunziata, J. (2002). Teacher evaluation, leadership and learning organizations. *Journal of Personnel Evaluation in Education*, 16(4), 287-301.
- Erickson, F. (1986). Qualitative methods on research in teaching. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3<sup>rd</sup> Ed.). New York, NY: MacMillan.
- Fraenkel, J. & Wallen, N. (2006). *How to design and evaluate research in education*. New York, NY: McGraw Hill Companies, Inc.
- Fuller, F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 6(2), 207-226.  
doi:10.3102/00028312006002207
- Gallagher, H. (2004). Vaughn elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education*, 79(4), 79-107.
- Gallucci, C. (2007). *Using sociocultural theory to link individual and organizational learning processes: The case of Highline School District's instructional improvement reform*. Seattle, WA: University of Washington, Center for the Study of Teaching and Policy.
- Gallucci, C. (2008). Districtwide instructional reform: Using sociocultural theory to link professional learning to organizational support. *American Journal of Education*, 114(4), 541-581.
- Gallucci, C., Van Lare, M. D., Yoon, I. H., & Boatright, B. (2010). Instructional coaching: Building theory about the role and organizational support for professional learning. *American Educational Research Journal*, 47(4), 919-963.
- Gamoran, A., Porter, A., Smithson, J., & White, P. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis*, 19(4), 325-338. doi:10.3102/01623737019004325
- Gavelek, J.R. & Raphael, T.E. (1996). Changing talk about text: New roles for teachers and students. *Language Arts*, 73(3), 182-192.
- Gay, L.R., Mills, G.E., & Airasian, P. (2009). *Educational research: Competencies for analysis* (9th ed.). New Jersey, NJ: Pearson.
- George, A., Hall, G.E., & Stiegelbauer, S. (2008). *Measuring implementation in schools: The stages of concern questionnaire*. Austin, TX: Southwest Educational Development Laboratory.

- Glaser, B. G., & Strauss, A. L., (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine Publishing Company.
- Hall, G.E. (1979). The concerns-based approach to facilitating change. *Educational Horizons*, 57(4), 202-208.
- Hall, G. E., George, A. A., & Rutherford, W. A. (1998). *Measuring the stages of concern about the innovation: A manual for the use of the SoC questionnaire*. Austin, TX: Southwest Educational Development Laboratory.
- Hall, G.E., & Hord, S. (1987). *Change in schools: Facilitating the process*. SUNY series in educational leadership. Albany, NY: SUNY Press.
- Hall, G. E., & Hord, S. M. (2006). *Implementing change: Patterns, principles and potholes* (3<sup>rd</sup> ed.). Boston, MA: Pearson.
- Hall, G.E., Wallace, R. C, & Dossett, W. D. (1973). *A developmental conceptualization of the adoption process within educational institutions*. Austin, TX: Research and Development Center for Teacher Education, The University of Texas.
- Harre, R. (1984). Problems and prospects in developmental psychology. *Oxford Review of Education*, 10(3), 251-304.
- Hinchey, P. H. (2008). *Action research*. New York, NY: Peter Lang.
- Iarossi, G. (2006). *Power of survey design: A user's guide for managing surveys, interpreting results, and influencing respondents*. Herndon, VA: World Bank Publications.
- James, R. (1981). Understanding why curriculum innovations succeed or fail. *School Science and Mathematics*, 81(6), 487-495.
- Johnson, B., & Christensen, L. (2012). *Educational research: Quantitative, qualitative, and mixed approaches* (4<sup>th</sup> ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Kelly, M., & Staver, J. (2005). A case study of one school system's adoption and implementation of an elementary science program. *Journal of Research in Science Teaching*, 42(1), 25-52.

- Kimball, S., White, B., Milanowski, A., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education*, 79(4), 54-78.
- Kochendorfer, L. (1997). Active voice. Types of classroom teacher action research. *Teaching and Change*, 4(2), 157-174.
- Kong, A., & Pearson, P. D. (2003). The road to participation: The construction of a literacy practice in a learning community of linguistically diverse learners. *Research in the Teaching of English*, 38(1), 85-124.
- Kotter, J. (1996). *Leading change*. Boston, MA: Harvard Business School Press.
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage Publications
- Lave, J., & Wenger E. (1991). *Situated learning: Legitimate peripheral participation (learning in doing: Social, cognitive and computational perspectives)*. New York, NY: Cambridge University Press.
- Lincoln, Y., & Guba, E. (1985). *Research, evaluation, and policy analysis: Heuristics for disciplined inquiry*. Retrieved from <http://www.eric.ed.gov/PDFS/ED252966.pdf>
- McLaughlin, M. W. (1990). Embracing contraries: Implementing and sustaining teacher evaluation. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation* (pp. 403-415). Newbury Park, CA: Sage.
- McVee, M. B., Gavelek, J. R., & Dunsmore, K. L. (2005). Revisiting schema theory. *Review of Educational Research*, 75(4), 531-566.
- McVee, M. B., Gavelek, J. R., & Dunsmore, K. L. (2007). Considerations of the social, individual, and embodied: A response to comments on "schema theory revisited". *Review of Educational Research*, 77(2), 245-248.
- Medley, D. M., & Coker, H. (1987). How valid are principals' judgments of teacher effectiveness? *Phi Delta Kappan*, 69(2), 138-140.
- 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., & Kimball, S. (2003, April). *The framework-based teacher performance assessment systems in Cincinnati and Washoe*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

- Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Mills, G.E. (2007). *Action research: A guide for the teacher researcher* (3<sup>rd</sup> ed.). Upper Saddle River, NJ: Pearson.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington D.C.: United States Department of Education.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 §1425 (2002).
- Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257.
- Odden, A. (2004). Lessons learned about standards-based teacher evaluation systems. *Peabody Journal of Education*, 79(4), 126-137.
- Odden, A., Borman, G., & Fermanich, M. (2004). Assessing teacher, classroom, and school effects, including fiscal effects. *Peabody Journal of Education*, 79(4), 4-32.
- Peterson, K. D., & Peterson, C. A. (2006). *Effective teacher evaluation: A guide for principals*. Thousand Oaks, CA: Corwin Press.
- Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York, NY: Free Press.
- Rogoff, B., Matusov, B., & White, S. (1996). Models of teaching and learning: Participation in a community of learners. In D. Olson & N. Torrance (Eds.), *The handbook of cognition and human development* (pp. 388-414). Oxford, UK: Blackwell.
- Ryan, G.W. & Bernard, H.R. (2000). Data management and analysis methods. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 769-802). Thousand Oaks, CA: Sage Publications.
- Sanders, W., & Horn, S. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247-256.
- Schlechty, P. C. (2009). *Leading for learning: How to transform schools into learning organizations* (1st ed.). San Francisco, CA: Jossey-Bass.

- School district employment report*. (2011). Retrieved from the Arizona State Department of Education website:  
<http://www.ade.az.gov/sder/PublicReports.asp>
- Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. New York, NY: Doubleday/Currency.
- Smith, M. L. & Glass, G. V. (1987). Experimental studies. In M. L. Smith & G. V. Glass, *Research and evaluation in education and the social sciences* (pp. 124-157). Needham Heights, MA: Allyn and Bacon.
- Stringer, E. T. (2007). *Action research* (3<sup>rd</sup> ed.). Thousand Oaks, CA: Sage.
- Suzuki, L. A., Ahluwalia, M. K., Arora, A. K., & Mattis, J. S. (2007). The pond you fish in determines the fish you catch: Exploring strategies for qualitative data collection. *The Counseling Psychologist*, 35(2), 295-327.
- Vygotsky, L.S. (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.
- Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, 7(2): 225-246.
- Wenger, E., McDermott, R. A., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Boston, MA: Harvard Business School Press.
- Wertsch, J. V. & Sohmer, R. (1995). Vygotsky on learning and development. *Human Development*, (38), 332-37.
- Westbury, I. (1993). American and Japanese achievement...again: A response to Baker. *Educational Researcher*, 22(3), 21-25.
- Wilhelm, J. D., Baker, T. N., & Dube, J. (2001). *Strategic reading: Guiding students to lifelong literacy, 6-12*. Westport, CT: Heinemann.
- Wooley, C.M. (2009). Meeting the mixed methods challenge of integration in a sociological study of structure and agency. *Journal of Mixed Methods Research*, 3, 7-25



APPENDIX A

PROFESSIONAL TEACHING STANDARDS ADOPTED IN 2010

## ***SCHOOL DISTRICT PROFESSIONAL TEACHING STANDARDS***

### **Standard I: Teachers Demonstrate Leadership**

- *Teachers lead in their classrooms.*
- *Teachers demonstrate leadership in the school.*
- *Teachers lead the teaching profession.*
- *Teachers advocate for schools and students.*
- *Teachers demonstrate high ethical standards.*

### **Standard II: Teachers Establish a Respectful Environment for a Diverse Population of Students**

- *Teachers provide an environment in which each child has a positive, nurturing relationship with caring adults.*
- *Teachers embrace diversity in the school community and in the world.*
- *Teachers treat students as individuals.*
- *Teachers adapt their teaching for the benefit of students with special needs.*
- *Teachers work collaboratively with the families and significant adults in the lives of their students.*

### **Standard III: Teachers Know the Content They Teach**

- *Teachers align their instruction with the state standards and their district's curriculum.*
- *Teachers know the content appropriate to their teaching specialty.*
- *Teachers recognize the interconnectedness of content areas/disciplines.*
- *Teachers make instruction relevant to students.*

### **Standard IV: Teachers Facilitate Learning for Their Students**

- *Teachers know the ways in which learning takes place, and they know the appropriate levels of intellectual, physical, social, and emotional development of their students.*
- *Teachers plan instruction appropriate for their students.*
- *Teachers use a variety of instructional methods.*
- *Teachers integrate and utilize technology in their instruction.*
- *Teachers help students develop critical thinking and problem-solving skills.*
- *Teachers help students work in teams and develop leadership qualities.*
- *Teachers communicate effectively.*
- *Teachers use a variety of methods to assess what each student has learned.*

### **Standard V: Teachers Reflect on Their Practice**

- *Teachers analyze student learning.*
- *Teachers link professional growth to their professional goals.*
- *Teachers function effectively in a complex, dynamic environment*

APPENDIX B  
INSTITUTIONAL REVIEW BOARD APPROVAL

**To:** Debby Zambo  
4701 West

**From:** Mark Roosa, Chair  
Soc Beh IRB

**Date:** 07/12/2011

**Committee Action:** **Exemption Granted**

**IRB Action Date:** 07/12/2011

**IRB Protocol #:** 1107006618

**Study Title:** Implementing a Standards-Based Teacher Evaluation System: Learning Experiences for Administrators in an Urban School District

The above-referenced protocol is considered exempt after review by the Institutional Review Board pursuant to Federal regulations, 45 CFR Part 46.101(b)(1) .

This part of the federal regulations requires that the information be recorded by investigators in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects. It is necessary that the information obtained not be such that if disclosed outside the research, it could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.

You should retain a copy of this letter for your records.

APPENDIX C  
PARTICIPANT CONSENT LETTER

July 18, 2011

Dear Administrative Colleagues:

I am a graduate student under the direction of Dr. Debby Zambo, Associate Professor in the College of Education at Arizona State University. I am conducting an action research study about administrator learning of standards-based teacher evaluation and the development of meaningful professional growth plans for teachers.

I am inviting your participation in this study and the accompanying dialogues and professional development which will occur from August through December 2011, during two to three professional development sessions a month for about two to three hours each. This study will involve professional development in all aspects of standards-based teacher evaluation. Participating administrators in the study will interpret standards and rubrics from the Teacher Evaluation System and will participate in exercises to improve inter-rater reliability. Your participation in this study is voluntary. You must be 18 or older in order to participate. If you choose not to participate or to withdraw from the study at any time, there will not be a penalty and it will not affect your participation in district professional development. You have the right not to answer any question, and to stop participation at any time. There are no known risks from taking part in this study, but in any research, there is some possibility that you may be subject to risks that have not yet been identified. The benefits of your participation in this research study includes professional development that provides time to work with your peers to improve teacher evaluation, and an opportunity to help others learn how administrators think and act in this type of professional development.

All information obtained in this study will be confidential. I will be collecting data in the form of: pre and post survey responses, meeting and interview transcripts, and my researcher's reflective journal. I would like to audiotape your participation in the professional development sessions and individual interviews; however, if you do not want to be recorded, you have the right to ask not to be recorded at anytime. You can also change your mind once the recording starts, just let me know.

All data collection measures will be analyzed and described in my final dissertation. These data will be kept confidential, and anonymity of each participant will be maintained. No identifying information will be gathered. I will not know who you are when I collect data. Additionally, our school and district names will not be identified in my final dissertation study. The audiotapes will be stored in a secured cabinet in my office. The tapes will be destroyed on June 1, 2012 at the conclusion of my study.

If you have any questions concerning the research study, please contact the research team at:

Dr. Debby Zambo, Principal Investigator  
4701 W. Thunderbird Ave  
Glendale, AZ 85306  
623-543-6334

Chris Canelake, Co-Investigator  
6539 E. Eugie Terrace  
Scottsdale, AZ 85254  
602-284-0881

If you have any questions about your rights as a participant in this research, or if you feel you have been placed at risk, you can contact the Chair of Human Subjects Institutional Review Board, through the ASU Office of Research Integrity and Assurance, at 480-965-6788. Please let me know if you want to be part of the study.

Sincerely, Chris Canelake

APPENDIX D  
SURVEY INSTRUMENTS

Dear Administrative Colleagues,

As a part of my graduate studies at Arizona State University I am involved in conducting an action research study that investigates our process of adopting and learning how to use our new teacher evaluation system (TES).

I am inviting you to participate in this study by helping me field-test survey instruments designed to gauge your current attitudes toward our new TES.

The items for the first part of this survey were adapted from the Stages of Concern Questionnaire (copyright 2006) and developed from typical responses of school and college teachers who ranged from no knowledge at all about various programs to many years experience using them.

As a result, many of the items on this questionnaire may appear to be of little relevance or irrelevant to you at this time. For the completely irrelevant items, please mark "0" on the scale.

Other items will represent those concerns you do have, in varying degrees of intensity, and should be marked higher on the scale.

Please respond in terms of your present concerns, or how you feel about your involvement with the new TES. We do not hold to any one definition of the new TES so please think of it in terms of your own perception of what it involves.

Remember to respond to each item in terms of your present concerns about your involvement or potential involvement with the innovation.

The results of your individual responses to this survey will be entirely confidential. Your responses will not be associated with your identity or the identity of your campus or department.

Please let me know if you have any questions or concerns.

I appreciate your help with this project!

Chris Canelake  
Director of Human Resources  
602-629-6400



Memorable Identifier: Please enter your mother's birthday, month and day only using four digits:  
\_\_\_/\_\_\_

(You will be asked to enter this same information in a follow-up survey to allow matching of pre- and post-survey responses.)

The items for the first part of this survey were adapted from the Stages of Concern Questionnaire (copyright 2006) and developed from typical responses of school and college teachers who ranged from no knowledge at all about various programs to many years experience using them.

As a result, many of the items on this questionnaire may appear to be of little relevance or irrelevant to you at this time. For the completely irrelevant items, please mark "0" on the scale.

Other items will represent those concerns you do have, in varying degrees of intensity, and should be marked higher on the scale.

Please respond in terms of your present concerns, or how you feel about your involvement with the new TES. We do not hold to any one definition of the new TES so please think of it in terms of your own perception of what it involves.

Remember to respond to each item in terms of your present concerns about your involvement or potential involvement with the innovation.

The results of your individual responses to this survey will be entirely confidential. Your responses will not be associated with your identity or the identity of your campus or department.

**The following thirty-five questions can be rated 1 through 7, on a graduated scale, or you can rate "0" if the question is irrelevant to your perspective.**

1. I am concerned about teachers' attitudes towards our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

2. I know of some other TES that might work better than what we have adopted.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

3. I am more concerned about other strategies that support improvement of teacher quality than I am about our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

4. If I am required to continue to implement our new TES, I am concerned that I might not have enough time to organize myself each day.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

5. I would like to help other administrators use our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

6. I have a very limited knowledge of our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

7. I would like to know the effect of adopting our TES on my professional status.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

8. I am concerned about conflict between my interests and my responsibilities for our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

9. I am concerned about revising my use of teacher evaluation with our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

10. I would like to develop working relationships with both our administrative team and outside administrators using similar standards-based TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

11. I am concerned how our new TES affects the teachers at my school.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

12. I am not concerned about our new TES at this time.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

13. I would like to know who will make decisions regarding our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

14. I would like to discuss the possibility of expanding my use of our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

15. I would like to know what resources are available if I decide to move toward an expanded use of our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

16. I am concerned about my inability to manage all that our new TES requires.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

17. I would like to know how my administrative duties are supposed to change as a result of our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

18. I would like to familiarize other people with the progress of our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

19. I am concerned about evaluating my impact on teachers while implementing our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

20. I would like to revise the approach of our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

21. I am preoccupied with things other than our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

22. I would like to modify our new TES based on the experiences of our teachers.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

23. I spend little time thinking about our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

24. I would like to excite my teachers about their part in our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

25. I am concerned about time spent working with nonacademic problems related to our new TES.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

26. I would like to know what our new TES will require in the immediate future.

Very true of me now		Somewhat true of me now			Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

27. I would like to coordinate my efforts with others to maximize our new TES effects.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

28. I would like to have more information about time and energy commitments required by our new TES.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

29. I would like to know what other district administrators are doing in regards to our TES.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

30. Currently, other priorities prevent me from focusing my attention on our new TES.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

31. I would like to determine how to supplement, enhance or replace our new TES.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

32. I would like to use feedback from teachers to change our new TES.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

33. I would like to know how my role will change when using our new TES.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

34. Within our new TES process, coordination of tasks and people is taking too much of my time.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

35. I would like to know how our new TES is better than what we had before.

Very true of me now	Somewhat true of me now				Not true of me now		Irrelevant
7	6	5	4	3	2	1	0

36. Do you have any comments about the questions asked in this section of the survey?

Memorable Identifier: Please enter your mother's birthday, month and day only using four digits:  
\_\_\_/\_\_\_

(You will be asked to enter this same information in a follow-up survey to allow matching of pre- and post-survey responses.)

The items for this part of this survey are designed to learn more about your perceptions and ideas about our new TES. There are twenty-five questions divided among five specific areas. There is also an open response area following each section.

Your frank and candid responses are appreciated for each section of this survey.

**The following twenty-five questions can be rated 1 through 5, on a graduated scale.**

### Inter-rater reliability

1. Our TES is likely to result in consistent ratings between administrators.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

2. I am confident that most administrators' ratings of the same teacher would be similar to my own.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

3. I am not overly generous or overly severe in rating teachers on the TES.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

4. My administrative peers are not overly generous or overly severe in rating teachers on the TES.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

5. Teachers can be confident of the consistency of evaluations of teachers on our campuses.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

6. Do you have any comments about the status of inter-rater reliability among our administrative team?

**Potential to improve teacher performance**

7. Our new TES is an effective tool to influence teacher development.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

8. Our new TES assists poorly performing teachers to make significant changes.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

9. The growth plans in our new TES are effective tools to help teachers improve their performance.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

10. I use our TES growth plans to effectively guide teachers to improve performance.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

11. Do you have any comments about our TES ability to influence teacher performance?

**Teacher involvement**

12. Teachers I supervise are actively involved in using our TES to improve their professional practice.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

13. Teachers use our TES self-assessment to improve their professional practice.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

14. The Teaching Standards in our new TES help teachers improve their practice.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

15. Teachers use the TES professional growth plans as an effective tool to improve their professional practice.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

16. Do you have any comments about teacher involvement with our TES?

**Satisfaction with instrument**

17. Our new TES Summary Evaluation Form is an effective tool for evaluating teacher performance.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

18. Overall, our new TES is relatively easy to manage.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

19. The teaching standards measured by our new TES are focused on what is necessary to raise student achievement.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

20. The rubrics to measure the teaching standards in our new TES are adequately descriptive.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

21. Do you have any comments about your satisfaction with our TES instruments?

**Professional growth plans**

22. I am able to facilitate development of professional growth plans based on information from our TES.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

23. Professional growth plans developed with our TES help make my best teachers even better.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

24. Professional growth plans developed with our TES help my worst performing teachers make progress.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

25. The TES is an effective tool for creating professional growth plans.

Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree
5	4	3	2	1

26. Do you have any comments about TES professional growth plans?

**Demographics:**

27. Gender: M/F

28. Are you a school or district office administrator? \_\_\_ School \_\_\_ District Office

29. How many years classroom teaching experience do you have? \_\_\_

30. What percentage of your K-12 teaching experience falls in these categories?

K-3 \_\_\_\_, 4-6 \_\_\_\_, Middle/Jr. High School: \_\_\_\_, High School: \_\_\_\_

31. What percentage of your K-12 teaching experience falls in these categories?

Regular Ed. \_\_\_\_, Special Ed. \_\_\_\_, Special Area: \_\_\_\_,  
Non-classroom Student Support: \_\_\_\_, Non-classroom Teacher Support: \_\_\_\_

32. How many years experience do you have evaluating teachers? \_\_\_

33. How many total years of experience do you have as a certificated educator? \_\_\_

34. How many total years experience do you have in your current school district? \_\_\_

Thank you! Do you have any additional comments or questions?



APPENDIX E  
KEY INFORMANT INTERVIEW

Interview Prompt	Primary Research Question(s) Addressed
1. Tell me about how you are doing with our new TES.	<ul style="list-style-type: none"> <li>•What concerns do administrators within my district have about effective implementation of our new TES and how will those concerns change as our CoP learns, practices and engages in discourse about the TES?</li> <li>•How do I lead this process of change?</li> </ul>
2. Do you feel our administrators have a common understanding of our teaching standards and how to use our new TES?	<ul style="list-style-type: none"> <li>•What concerns do administrators within my district have about effective implementation of our new TES and how will those concerns change as our CoP learns, practices and engages in discourse about the TES?</li> <li>•What will administrators say, do and feel as our CoP attempts to develop a common understanding of the professional teaching standards?</li> <li>•How do I lead this process of change?</li> </ul>
3. What do you see as ideal outcomes for the TES?	<ul style="list-style-type: none"> <li>•What concerns do administrators within my district have about effective implementation of our new TES and how will those concerns change as our CoP learns, practices and engages in discourse about the TES?</li> <li>•What will administrators say, do and feel as our CoP attempts to understand the purpose of TES to improve professional practice?</li> <li>•How do I lead this process of change?</li> </ul>
4. How would you feel about defending our inter-rater reliability on the TES?	<ul style="list-style-type: none"> <li>•What concerns do administrators within my district have about effective implementation of our new TES and how will those concerns change as our CoP learns, practices and engages in discourse about the TES?</li> <li>•What will administrators say, do and feel as our CoP attempts to increase inter-rater reliability?</li> <li>•How do I lead this process of change?</li> </ul>
5. Tell me about your experiences with evaluation and professional growth with your most effective, average, and least effective teachers.	<ul style="list-style-type: none"> <li>•What concerns do administrators within my district have about effective implementation of our new TES and how will those concerns change as our CoP learns, practices and engages in discourse about the TES?</li> <li>•What will administrators say, do and feel as our CoP attempts to understand the purpose of TES to improve professional practice?</li> <li>•How do I lead this process of change?</li> </ul>
6. How would you describe your overall perception of our TES adoption?	<ul style="list-style-type: none"> <li>•What concerns do administrators within my district have about effective implementation of our new TES and how will those concerns change as our CoP learns, practices and engages in discourse about the TES?</li> <li>•How do I lead this process of change?</li> </ul>

---

7. What do you see as your greatest challenge for improving your practice with our TES?

- What concerns do administrators within my district have about effective implementation of our new TES and how will those concerns change as our CoP learns, practices and engages in discourse about the TES?
- How do I lead this process of change?

---

8. What are your key concerns about effective implementation of our TES?

- What concerns do administrators within my district have about effective implementation of our new TES and how will those concerns change as our CoP learns, practices and engages in discourse about the TES?
  - How do I lead this process of change?
-

APPENDIX F

SoCQ QUESTIONNAIRE PEAK AND SECONDARY SCORES

PARTICIPANT	S0	S1	S2	S3	S4	S5	S6
PRE-1	91	84	92	88	33	40	57
POST-1	94	45	52	56	9	31	26
PRE-2	91	51	55	56	21	40	22
POST-2	91	43	48	52	16	31	30
PRE-3	99	63	55	27	7	64	57
POST-3	91	43	52	65	16	31	26
PRE-4	94	16	25	34	11	22	20
POST-4	96	51	59	60	19	40	30
PRE-5	99	43	35	65	59	68	65
POST-5	87	45	55	56	16	25	17
PRE-A	99	63	67	88	13	36	26
POST-B	94	48	59	69	19	31	30
PRE-C	99	43	63	95	27	52	52
POST-D	91	48	52	56	16	36	30
PRE-E	99	88	96	90	86	64	30
POST-F	91	48	57	56	21	31	22
PRE-ALL	98	57	63	73	27	48	38
POST-ALL	91	45	55	60	16	31	26
PRE-DIST	98	57	55	43	13	52	38
POST-DIST	91	43	52	60	16	31	30
PRE-PRINC	98	54	67	80	33	48	42
POST-PRINC	94	48	57	60	16	31	26
PRE-EXP PRINC	99	34	41	69	27	48	42
POST-EXP PRINC	94	48	57	60	19	31	22
PRE-EXP EVAL	87	45	55	56	16	25	17
POST-EXP EVAL	75	40	45	43	9	22	20
PRE-NEW	97	75	87	88	38	48	38
POST-NEW	94	48	57	65	13	31	30
PRE-INEXP EVAL	99	57	63	73	27	55	57
POST-INEXP EVAL	91	45	52	60	13	31	26
	XX	=Peak Stage Score		XX	=Second Highest Stage Score		

Dist=District Administrators, New=Principals new to district, Exp Princ=Principals with four or more years as principal, Exp Eval=Administrator with five or more years experience evaluating principals, New=Principals new to district, Inexp Eval=Administrator with less than five years experience evaluating teachers.

APPENDIX G

SoCQ INDIVIDUAL PROFILE LINE GRAPHS

