Developing a culture of evaluative inquiry: A case study

Happy Miller

Follow this and additional works at: https://digitalrepository.unm.edu/oils_etds

Recommended Citation
Happy Lee Miller
Candidate

Educational Leadership and Organizational Learning

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Patricia Boverie, Chairperson

Charlotte Gunawardena

Mark Salisbury

Arlie Woodrum
DEVELOPING A CULTURE OF EVALUATIVE INQUIRY: A CASE STUDY

By

Happy Lee Miller

B.S., Foreign Service, Georgetown University, 1987
M.A., English as a Second Language, University of Hawaii-Manoa, 1994

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Organizational Learning and Instructional Technology

The University of New Mexico
Albuquerque, New Mexico

May 2012
Dedication

I dedicate this dissertation to my wonderful family and friends who have graciously supported me throughout my doctoral studies. I especially would like to thank my patient husband, Charles Bryant, and precious daughter, Katerina Retwaiut, who have put up with my diverted attention for many years now. Yes, Katerina, you may get a cat now. I also would like to thank my mother, Gay Regan, for her unwavering support of my educational endeavors throughout my life. Finally, I would like to thank my friends, Denise Mainardi, Dorothy Stewart, and Maria Cordova, for watching my daughter after school so that I could attend class. This dissertation truly would not have been possible without the support of these people.
Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr. Patsy Boverie, for her guidance and support throughout the entire research process. Her caring and support have been steadfast from the first class in which I enrolled until now. The other members of my dissertation committee, Dr. Mark Salisbury, Dr. Lani Gunawardena, and Dr. Arlie Woodrum have generously given their time and expertise to better my work. I thank them for their contribution and good-natured support. I also would like to thank Dr. Bruce Noll for his guidance and support in the early stages of this research.

I would like to thank my classmates for providing valuable feedback throughout the process including Laura Sullivan Gallegos, Pamela Verstynen, and Michael Stanton.

Finally, I would like to thank the participants from my research site who very generously allowed me to access to their world. They are not named here in order to protect their anonymity. However, I am eternally grateful to them because this research would not have been possible without their help.
DEVELOPING A CULTURE OF EVALUATIVE INQUIRY: A CASE STUDY

By

HAPPY LEE MILLER

B.S., Foreign Service, Georgetown University, 1987
M.A., English as a Second Language, University of Hawaii-Manoa, 1994
Ph.D., Organizational Learning and Instructional Technology, The University of New Mexico, 2012

ABSTRACT

Research has shown that few educational reform efforts have achieved significant, long-lasting improvement in student achievement. In order to accelerate growth in student achievement, many researchers have advocated the development of cultures of evaluative inquiry. In a culture of evaluative inquiry, teachers use data to identify instructional areas in which their students struggle; collaboratively explore the problem, their current practices, and possible solutions; and test the solutions through active experimentation. Through this iterative process, teachers develop their mental representations of their students’ needs and their own instructional practice which leads to sustained changes in instructional strategies.

This case study documented the actions and perceptions of members of the Instructional Leadership Team and five professional learning community teams as they endeavored to develop a culture of evaluative inquiry from July of 2008 through May of 2010 at North Mesa Elementary School. Observations, interviews, and document reviews were utilized to answer the central question of this study: How do staff members of a public elementary school develop a culture of evaluative inquiry?
The results indicated that instances of significant learning and changes of practice were evident but the staff was not successful in fully developing and sustaining a culture of evaluative inquiry. Staff members acquired some evaluative inquiry skills through the interdependent use of data in order to provide grade-level wide interventions although explicit training efforts faltered. A number of challenges were encountered including a lack of shared sense of urgency, a lack of shared vision, a lack of psychological safety within some teams, insufficient communication, and incomplete implementation of standards-based instruction.

The study concluded that a shared vision, a focus on standards, learning and results, a collaborative culture, aligned systems and structures, and a knowledge management system are essential components of a culture of evaluative inquiry. Furthermore a five-stage iterative data-based problem solving process was prescribed with the goal of providing both professional development for teachers and improving student achievement. Finally, the district and school leadership need to create a system of defined autonomy which balances the need for alignment across the system with teacher empowerment.
# Table of Contents

List of Figures ......................................................................................................................... xi

List of Tables ........................................................................................................................... xii

**Chapter I: Statement of the Problem** ................................................................................... 1

  - Opening Vignette ........................................................................................................ 1
  - Problem Statement ..................................................................................................... 2
  - Purpose of the Study ................................................................................................. 12
  - Guiding Questions ..................................................................................................... 12
  - Significance of the Study ........................................................................................... 13
  - Delimitations ............................................................................................................... 13
  - Definitions .................................................................................................................... 13
  - Summary ..................................................................................................................... 15

**Chapter II: Literature Review** ............................................................................................ 16

  - Philosophical Paradigms .......................................................................................... 17
  - Individual Learning .................................................................................................... 31
  - Group Learning .......................................................................................................... 32
  - Organizational Learning ............................................................................................. 39
  - Culture of Evaluative Inquiry .................................................................................... 49
  - Collaborative Use-Focused Approaches to Evaluation .............................................. 52
  - Evaluation Capacity Building ..................................................................................... 56
  - Elements of Sustainable Practice .............................................................................. 58
  - Barriers to Developing Cultures of Evaluative Inquiry .............................................. 63
  - Using ECB to Develop Cultures of Evaluative Inquiry .............................................. 65

viii
Chapter III: Methods ........................................................................................................ 68
Restatement of the Problem and Research Questions ........................................... 68
Research Site ........................................................................................................... 69
Role of the Researcher ......................................................................................... 76
Research Design .................................................................................................... 77
Research Procedures ............................................................................................ 79
Participants ......................................................................................................... 79
Data collection methods .................................................................................... 82
Trustworthiness and consistency of the data ..................................................... 84
Data analysis ....................................................................................................... 85

Chapter IV: Results .................................................................................................. 86
Summary .............................................................................................................. 118

Chapter V: Discussion and Implications ................................................................. 120
Discussion ........................................................................................................... 121
The Miller Evaluative Inquiry Model ................................................................. 134
Purpose and goals of the Miller Evaluative Inquiry Model ......................... 134
The big picture ................................................................................................... 135
The evaluative inquiry process ......................................................................... 141
Questions for Further Research ................................................................. 150
Conclusion ........................................................................................................ 151

References ......................................................................................................... 153
Appendices

Appendix A: Consent Form

Appendix B: The Case
List of Figures

Figure I-1: Conversation topics related to assessment in grade-level PLC meetings .......... 11
Figure II-1: Kochanek’s Process Model of Trust Building in Schools .............................. 63
Figure III-1: Suburban School District Student Enrollment 1994-2010 .......................... 71
Figure III-2: Student enrollment at North Mesa Elementary School, 2002-2010 ............ 72
Figure III-3: Distribution of Ethnicities within Suburban School District, 1994-2010 ....... 73
Figure III-4: Ethnic composition of North Mesa Elementary School and Suburban School District ......................................................................................................................... 74
Figure V-1: The Big Picture of the Miller Evaluative Inquiry Model ............................... 136
Figure V-2: Evaluative Inquiry Process ............................................................................ 142
List of Tables

Table II-1: Comparison of Two Approaches to Professional Development ......................... 27
Table II-2: Decision-Making Styles ....................................................................................... 38
Table III-1: Population growth in Suburbia ........................................................................... 70
Table IV-1: Percentage of students proficient on the State Accountability Assessment at
           North Mesa Elementary School, 2006-2010.............................................................. 117
Table IV-2: Adequate Yearly Progress Trends at North Mesa Elementary School, 2008-2010
           ................................................................................................................................... 118
Table V-1: Types of Questions to use in the Evaluative Inquiry Process ............................ 146
Table B-1: Adequate Yearly Progress Trends at North Mesa Elementary School, 2006-2008
           ................................................................................................................................... 173
Table B-2: Percentage of Students Proficient on the SBA, 2006-2008 .............................. 173
Chapter I: Statement of the Problem

Opening Vignette

After a long day of teaching, eight elementary school teachers wearily trudge into their colleague’s classroom. As they shift around on the hard, student-sized seats trying to get comfortable, they try to focus on the agenda for their meeting. Wishing that they were back in their own classroom catching up on grading or preparing for the next day’s lesson, one of the teachers asks:

I don’t want to sound mean or whatever, but what are we accomplishing in our PLCs [professional learning communities]? What is our goal? Absolutely nothing?

One teacher immediately responds, “A lot of times, no.” Another teacher chimes in, “We’re supposed to learn to work together. This group is never going to work together.” A third teacher says, “I think we’re supposed to share ideas on how to teach better, on what’s working.” One of the teachers adds “Last year we read books about teaching practices and discussed them,” and then asks, “Did we ever discuss students?” to which her colleagues respond negatively. One teacher clarifies, “People sometimes informally brought students up when they had some issues or something but that certainly wasn’t the focus.”

This conversation, drawn from an observation at an elementary school in Suburban School District1 in central New Mexico in late September, 2007, illustrates that the process of creating effective organizational cultures focused on improving student learning requires more than mandating collaborative meeting time.

---

1 Pseudonyms have been used for the name of the district, the school, and all of the participants in order to protect their anonymity.
Problem Statement

Many people perceive American public schools as static bureaucracies which are unable to adapt to changing environmental conditions and demands for high performance. Despite decades of school reform initiatives, very little has actually changed in the academic achievement of students (Garmston & Wellman, 1999). In reaction to this, two trends in school improvement have emerged in recent years: externally-mandated accountability initiatives and internally-driven reform efforts.

Public distrust in schools has translated into renewed demands for accountability and bottom-line performance as evidenced by the 2001 ‘No Child Left Behind’ Act (PL 107-110). This externally-mandated school reform emphasizes “scientifically-based research” and sanctions for schools which do not meet Adequate Yearly Progress. These tactics are grounded in both empirical-rational and power-coercive strategies for change. Both strategies for change share the assumptions that good ideas are best developed outside the organization, and organizations will not change unless pressured from the outside (Chin, 1967).

The history of externally-mandated school reform has produced mixed results at best. The Rand Corporation (2006) states that some comprehensive school reform initiatives have produced modest improvement; others have had no impact. Further investigation shows that externally-mandated reforms are rarely implemented with fidelity. A survey of 250 schools in Florida and Texas that nominally implemented comprehensive school reform models showed that none had adopted all of the specified changes (Rand Corporation, 2006). The study found that, “schools were most likely to adopt the curriculum prescribed by the model developer, but were less likely to adopt the recommended instructional practices (p. 2).”
This finding lends credence to the concern of some researchers that teachers’ opinions have been ignored in the suggested formulas for reform (Wildman, Niles, Magliaro, & McLaughlin, 1990).

In contrast to the empirical-rational view of teaching as a technical enterprise, many organizational developers believe that lasting organizational change is most likely to occur when the initiative originates from within the organization. Through reflective inquiry, teachers should become increasingly empowered and skilled at conducting inquiry into their own practice. Over time, the reflective inquiry will lead to the evolution of instructional practices which should lead to improved student achievement (Wildman, Niles, Magliaro, & McLaughlin, 1990).

Through inquiry into their tacit assumptions and beliefs, teachers affect the organizational culture. Organizational culture is, “a learned pattern of unconscious (or semiconscious) thought, reflected and reinforced by behavior, that silently and powerfully shapes the experience of people (Owens & Valesky, 2007, p. 192).” Owens and Valesky (2007, p. 201) state that there is overwhelming agreement that, “organizational culture is powerful in determining the course of change in an organization. Not a few believe that it is often the most powerful determinant.”

The tension between external and internal forces for change is also reflected in changing models of effective professional development for teachers. Traditional professional development for in-service teachers has reflected an objectivist view of learning. Staff developers followed the ADDIE (Analyze, Design, Develop, Implement, Evaluate) model of instructional systems design based on the belief that they could provide professional knowledge to teachers during workshops and that the teachers would then be able to transfer
the newly acquired knowledge to their instructional practice. Although a pull-out workshop model is efficient, research has cast doubt on the effectiveness of this staff development model (Sparks & Loucks-Horsley, 1989).

Research has demonstrated, however, that some professional development practices are effective. Some of the practices that have been shown to be effective are (Sparks & Loucks-Horsley, 1989):

- Programs conducted in school settings and linked to schoolwide efforts;
- Teachers participating as helpers to each other and as planners, with administrators, of in-service activities;
- Emphasis on self-instruction, with differentiated training opportunities;
- Teachers in active roles, choosing goals and activities for themselves;
- Emphasis on demonstration, supervised trials, and feedback; training that is concrete and ongoing over time; and
- Ongoing assistance and support available upon request.

Consequently, the focus of staff development has been shifting from the delivery of workshops and training to developing collaborative inquiry in school settings (Killion & Harrison, 1997).

The concept of teacher collaboration in professional learning communities has gained great support as a means for developing professional capacity (DuFour & Eaker, 1998; Marzano, 2003; McLaughlin & Talbert, 2001; Schmoker, 2005). Schmoker (2005) has stated, “[there is] a broad, even remarkable concurrence among educational researchers and organizational theorists who have concluded that developing the capacity of educators to function as members of professional learning communities is the best-known means by which we might achieve truly historic, wide-scale improvements in teaching and learning.” In professional learning communities, teachers work in teams to collectively inquire into the
learning of their students and how their own instructional practices impact the learning of their students. DuFour describes the nature of the advocated collaboration:

The powerful collaboration that characterizes professional learning communities is a systematic process in which teachers work together to analyze and improve their classroom practice. Teachers work in teams, engaging in an ongoing cycle of questions that promote deep team learning (2005, p. 36).

Through collaboration, teachers are expected to both improve practice and generate knowledge.

While various definitions of “professional learning community” or “teacher learning community” exist, they all cite working collaboratively to reflect on practice, examining evidence about the relationship between practice and student outcomes, and making changes that improve teaching and learning for the particular students in their classes (McLaughlin & Talbert, 2006). These communities serve as a “community of explanation” where members develop shared understandings of evidence and expectations (Freeman, 1999). These communities foster more coherent practices across the community and more productive practices because they are situated in particular instructional settings. Since day-to-day practices are inherently local, site-based communities are often most effective in supporting sustained changes in practice as compared to off-site communities of practice (McLaughlin & Talbert, 2006).

In order to analyze and improve their classroom practice, it is necessary for teachers to develop evaluation skills. Researchers such as King (2007), Cousins (2004), Preskill and Torres (1999), and Patton (2004) suggest that the most appropriate approach to evaluation for the context of developing the evaluation capacity of staff members would be one in which evaluation is ongoing and iterative, integrated into regular work routines, and is performed
primarily by organization members. The long-term goal of Evaluation Capacity Building (ECB) is to cultivate a spirit of continuous organizational learning, improvement, and accountability and to create awareness and support for evaluation as a performance improvement strategy (King, 2007).

Evaluation researchers have noted that many challenges exist in the development of a culture of evaluative inquiry. For example, Shulha states, “it is rare to find the structures and resources necessary to support teachers and administrators in intentional and continual inquiry into programs and practices (2000, p. 39).” Some of the barriers identified by Preskill and Torres (1999) include reactive organizational cultures, fear of making mistakes, and the belief that evaluative activities cost too much in terms of money, time, and personnel resources. King (2007) has also observed that it is easier to begin ECB than to sustain it over time.

Interest in Evaluation Capacity Building and the development of cultures of inquiry has grown in recent years (Preskill, 2008). However, there is little empirical evidence detailing the process by which schools have successfully created this shift in culture (Joyce, 2004). Indeed, some research has questioned the effectiveness of many induced collaboration efforts (Lam, Yim, & Wing-hong Lam, 2002; Little, 1990). Hargreaves (1991) suggests that mandated collaboration risks the possibility that teachers will approach the joint work with a project mentality. Rather than forming a truly collaborative culture with meaningful interdependent work, the teachers may just adopt the rituals of collaborative practice and form a culture of contrived collegiality.

Suburban School District has been striving for the past several years to develop cultures of evaluative inquiry within schools through the creation of professional learning
communities (PLCs). Prior unpublished pilot study research has shown that the nature of the ensuing collaborative teacher inquiry has varied across Suburban School District. Data collected through thirty-one observations, thirteen interviews, fifty-seven surveys, and document reviews at two elementary schools and four secondary schools between spring of 2007 and spring of 2008 indicate a wide variety of practices and varying degrees to which conversations focus on student achievement and instructional practices (Miller, 2007, 2008a, 2008b).

The prior research shows that a shared vision of the collaborative evaluative inquiry process embedded in the definition of professional learning communities does not yet exist in Suburban School District. One PLC leader stated:

Teachers really don’t understand what it [the concept of PLCs] is. Unless you had an opportunity to go [on a site visit] or attend these [PLC] workshops, it is just a buzzword.

Some people equate PLCs with the sharing of resources and teaching strategies and not with collaborative evaluative inquiry as the following comment implies:

We already do it. They just started calling it PLCs this year, but we’ve been getting together and helping each other out and talking about stuff on our own…Plus, by after five minutes, we’re done. We don’t have anything else to say really because we’ve covered everything we need to know.

The creation of a shared vision and changing the organizational culture is more complex than mandating meeting time as this comment from a teacher indicates:

I think forcing people to have a PLC is a bad idea because then…you feel like you’re micromanaged. Well, you are forcing people into something that they don’t really have the time, they don’t really care about. They are already working on what they need to work on. People are already overburdened. And if you add one more thing in there. Oh, well now you have to meet with your colleagues. It’s like well, I already have everything. I have already shared that with my colleagues during the day. Now I have to meet with them. I think that forcing a PLC would be a bad idea. At least within our department because everybody is already essentially doing it on their own
time and they don’t have time to do it after school or whenever they would have it in a formal setting.

Furthermore, some teachers appear to view PLC activities as an expansion of their role as an educator and believe that there should be additional financial compensation. One teacher stated:

If they [the district leadership] really want to see the short cycle assessment data and have a tight curriculum map, you know, a really good syllabus, then we would need more time, more paid time outside of the classroom. Because, I mean, nobody is really in charge of doing that except the content leader which is me [laughs] but you don’t want to do that as an individual because you want to include the entire department…It would be nice if they could be paid because that’s not really part of their job description to develop this curriculum. They’re the teacher.

Thus, the expectation that teachers need to meet to collaboratively inquire into the learning of their students has not yet become a mainstay of the organizational culture.

Even with teachers who valued the collaborative inquiry process, challenges were observed. In many of the observations of professional learning community meetings, business topics took the majority of the formal meeting time rather than discussions about student learning. Business topics included discussions of joint field trips, ordering instructional materials, providing technical support, preparing committee reports, and resolving the logistics of other joint activities such as community-related activities. One elementary grade-level chair explained:

We are trying to address assessment issues. You know, any changes that need to be made. We, our intentions always is to talk about student achievement but we never ever seem to get to do that. A lot of it seems to be more business than anything as of late…We don’t have a set time where it is just let’s have a grade level meeting to go over some quick business stuff that you need to know and then we don’t have a set time to go over how the kids are doing and share ideas. Even though that is what we would all really like to do but it just seems like the time isn’t there.
The degree to which business topics dominated meetings varied across teams. The pressure of external sanctions has created some motivation to change teacher practices as one secondary school described:

There’s more emphasis [on discussing benchmarks] this year because… last year most of the things that we did in our department were nuts and bolts. And, um, unfortunately that still has to happen but we’re talking more about benchmarks this year in the PLC definitely. I’m not sure if it’s the PLC or because we’re required to make up that short-cycle assessment or because of the emphasis on the SBA [the state accountability exam] or…AYP [Adequate Yearly Progress].

However, the degree to which the teachers have consistently examined, analyzed, and used data about student achievement is inconsistent.

For example, one of the PLC teams analyzed summative assessment information from the prior year to create goals for their required “Plan-Do-Study-Act” (PDSA) inquiry plan but none of their conversations between October and the end of the observations in March focused on their PDSA goals; a review of their PDSA plan in March confirmed that no revisions or updates had been made. For some teams that did review their PDSA plan on a regular basis, the process still had limited effectiveness. One teacher at a different school admitted:

I’m not even really ready to talk about the PDSA for third quarter even though I know that we probably should. But because we meet so rarely…we’re still not getting to the meat and now we have to talk about it again and it’s like ‘eh’.

The data from formative assessments are also being erratically used. One secondary school department Chair stated in an interview, “Everybody pre- and post-tests, but what they do with that data I’m not sure.” Observations of elementary school PLC meetings showed that examination of common assessments such as the math Short Cycle Assessment (SCA), the Developmental Reading Assessment (DRA), the Dynamic Indicators of Basic
Early Literacy Skills (DIBELS), the Northwest Evaluation Association’s Measures of Academic Progress test (NWEA), and the Standards Based Assessment (SBA) led to a variety of topics. The majority of conversations focused on the logistics of testing. For example, the kindergarten teachers agreed to facilitate the individual assessment of students on the DIBELS by taking a portion of other teachers’ classes for several days while those teachers administered the test to their students. This topic was discussed in two meetings. Once the assessment had been administered, the teachers discussed how to synthesize contrasting assessment scores, how to use the conversion chart, and how to record the SCA data on the Intranet.

Figure I-1 indicates the breadth of comments observed in elementary grade-level discussions regarding assessments:
Comparisons of student scores frequently led to comparisons of how the tests were administered and scored. The logistical discussions were necessary to ensure that differences in scores were reflective of differences in teaching practices and not differences in assessment practices. However, by the time the logistical conversation had taken place, there was often little time remaining in formal meetings for teachers to discuss the implications and plan a course of action. One teacher stated:

People want other ideas from each other and they want to know what people are doing in their classrooms…it’s just the time factor that we sit down to share information and then people end up feeling like ‘Oh, my class is at the bottom’ but there’s never any more discussion…We’re at the initial phase all year long.
This trend of not being able to fully and effectively implement a cycle of evaluative inquiry was consistent across the teams observed.

**Purpose of the Study**

The purpose of this single case study is to describe the processes used to develop a culture of evaluative inquiry at one public elementary school in Suburban School District. The boundaries of this case are defined as school improvement-related processes undertaken by the school leadership team and teachers at the school between July, 2008 and May, 2010.

**Guiding Questions**

**Central question:**
How do staff members at a public elementary school develop a culture of evaluative inquiry?

**Subquestions:**
- How do members of the school leadership team increase their knowledge and skills related to evaluation and data analysis?
- How do members of the school leadership team apply their evaluative inquiry knowledge and skills to emergent issues within their school community?
- How do members of the school leadership team consciously build the evaluation capacity of teachers within their professional learning communities?
  - What training activities are held?
  - What roles do leadership team members play?
  - What learning processes are included in the evaluation capacity building activities?
  - How does the leadership team provide support in terms of communication, systems and structures and organizational culture?
  - How does the leadership team address the information needs of differing MBTI personality types (function theory)?
  - What challenges are encountered in the evaluation capacity building process?
- How does this initiative change the collaboration processes among teachers in the professional learning communities? Are there unintentional consequences in addition to the goals of the initiative?
- Do changes in demographics, student achievement, perceptions, and processes occur over the course of the case study?
Significance of the Study

Researchers from the fields of organizational development, school improvement, and program evaluation all are concurrently advocating the development of cultures of reflective, evaluative inquiry as a means of changing teachers’ instructional practices. The goal of the instructional changes is to enhance student learning. There is convergent research from several fields that claim that professional community is important in both instructional improvement and school reform (Andrews & Lewis, 2007; DuFour & Eaker, 1998; Marzano, 2003; McLaughlin & Talbert, 2001; Preskill & Torres, 1999; Schmoker, 2005). However, there is relatively little research that shows how effective professional learning communities and cultures of evaluative inquiry are created and sustained (Mitchell & Sackney, 1998; Stoll, Bolam, McMahon, Wallace, & Thomas, 2006).

This case study provided data about the process of developing a culture of evaluative inquiry and some of the challenges encountered. The results are of interest not only to the larger educational community but also to the school district in which the research site is located since the goal is for this school to serve as a model for other schools to undergo a similar process in future years.

Delimitations

This study examines the process of developing a culture of evaluative inquiry at an elementary school. Since it is a case study of a specific organization, the results are not necessarily generalizable to other populations.

Definitions

- Attitudes: “Acquired internal states that influence the choice of personal action.”
  (Gagne, 1985)

• Evaluation: The process of determining the merit, worth, or value of something, or the product of that process. (Scriven, 1974)

• Evaluation Capacity Building: “The intentional work to continuously create and sustain overall organizational processes that make quality evaluation and it uses routine.” (Baizerman, Compton, & Stockdill, 2002, p. 109)

• Evaluative inquiry: “An ongoing process for investigating and understanding critical organizational issues.” (Preskill & Torres, 2000)

• Group: “Two or more people who interact with each other to perform a task…Members have a shared objective, feel connected to the group, and recognize others as members. They coordinate their behavior, resources, and tools toward the project and tasks, and they share responsibility for the group’s collective outcomes.” (London & Sessa, 2006, p. 304)

• Organizational culture: “A learned pattern of unconscious (or semiconscious) thought, reflected and reinforced by behavior, that silently and powerfully shapes the experience of people.” (Owens & Valesky, 2007)

• Organizational learning: “The systems, principles, and characteristics of organizations that learn as a collective entity.” (Marquardt, 1996, p. 230)

• Paradigm: “A basic set of beliefs that guide action.” (Guba, 1990, p. 17)

• Process use: “Relating to and being indicated by individual changes in thinking and behaving that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process.” (Patton, 1998, p. 225)
• Professional Learning Community: “A group of people, operating as a collective enterprise, sharing and critically interrogating their practice in an ongoing, reflective, collaborative, inclusive, learning-oriented, growth-promoting way.” (Stoll, Bolam, McMahon, Wallace, & Thomas, 2006, p. 223)

Summary

Decades of school reform initiatives have failed to produce sustained improvement in student achievement. In response, two different trends in school improvement with contrasting philosophical bases have competed for dominance in recent years. While externally-mandated school reforms have been imposed by the U.S. Congress through the ‘No Child Left Behind’ Act of 2001, many researchers hypothesize that internally-driven collaborative inquiry processes are more likely to create long-lasting change.

In order to analyze and improve their classroom practice, it has been posited that teachers need to integrate evaluation into their regular work routines in an ongoing and iterative manner. Furthermore, it has been posited that schools should create cultures of evaluative inquiry which are focused on continuous organizational learning, improvement, and accountability. Researchers have noted that many challenges exist in the development of cultures of evaluative inquiry. This study sheds light on the process of developing and sustaining a culture of evaluative inquiry in an elementary school and the challenges that school members encountered.
Chapter II: Literature Review

For generations many citizens have felt that schools have not adequately prepared students for the challenges that they will face as adults. New educational reform movements have emerged on a regular basis over the past hundred years. Graham (2005) describes how changing external events and conditions have led to expectations for public education that evolve every twenty to thirty years. For example, the wave of immigration largely from eastern and southern Europe from 1890 to 1920 led to demands that the school system assimilate students into the “American culture” and develop the immigrants into patriotic citizens for the good of the nation. From 1920 to 1954, the primary focus was on the psychosocial needs of students and helping them fit into their “expected niche in society” (Graham, 2005, p. 160). The Civil Rights movement led to demands of improved access to educational opportunities. However, Graham states (2005, p. 101) that, “the energy was focused on access, not on the quality of the program”. Since the publication of *A Nation at Risk: The Imperative for Educational Reform* in 1983, there has been widespread concern that student achievement as measured by standardized assessments is insufficient.

While there is widespread agreement that reform is necessary, there is not widespread agreement about how reform is best achieved. This chapter describes two philosophies that affect one’s approach to learning, instruction, staff development, evaluation, and consequently, reform initiatives: post-positivism and social constructivism. The chapter further describes how individuals, groups, and organizations learn. The chapter continues by detailing factors that affect organizational cultures and how evaluation practices can be incorporated into those cultures to create cultures of evaluative inquiry.
Philosophical Paradigms

Two different philosophical paradigms underlie current educational reform initiatives. A paradigm is “a basic set of beliefs that guide action” (Guba, 1990, p. 17). These paradigms shape educational theorists’ views on the nature of reality, the proper relationship between the researcher and that which is being researched, the role of values in research, and appropriate research design. While other paradigms such as critical research do exist, the two dominant paradigms currently are postpositivism and social constructivism.

The pervasive assumption in Western cultures, until at least the mid-twentieth century, has been that the world was characterized by underlying patterns of logic, system, and order (Cresswell, 2007). Reality was considered to be objective, reductionistic, and knowable through rigorous methods of data collection and analysis. Postpositivist theorists hold this assumption as a central tenet of their philosophical beliefs today. Postpositivist researchers believe that they can objectively test hypotheses based on *a priori* theories through a series of logically related steps. The “gold standard” of postpositivist research is the use of experiments with randomly assigned participants, comparison groups, and carefully controlled conditions (Creswell, 2007).

Under the postpositivist philosophy, education is considered to be an object or phenomenon to be studied (Merriam, 1998). Knowledge and effective instructional practices can be classified into relative parts and placed into hierarchical schema (e.g., Bloom’s Taxonomy) (Cresswell, 2007). Education is considered to be orderly and organized. This worldview is found in many common educational practices including teaching for objectives and standardized educational assessment. Furthermore, this philosophical paradigm serves as
the basis for the emphasis on “scientifically-based research” in the ‘No Child Left Behind’

The postpositivist philosophy also underpins the traditional bureaucratic, or “factory
model”, approach to organizational theory. The bureaucratic approach tends to emphasize
five mechanisms to control and coordinate the behavior of people within an organization
(Callahan, 1962). The first mechanism is to maintain firm hierarchical control of authority
and close supervision of those in lower ranks. Within schools, administrators serve as
inspectors and evaluators of teachers’ classroom instructional practices. The second
mechanism is to establish and maintain adequate vertical communication. Third, along with
vertical communication, the bureaucratic model calls for clearly written rules and procedures
to set standards and guide actions. Examples include curriculum guides, policy handbooks,
standard forms, and rules and regulations. Fourth, the bureaucratic approach promulgates
clear plans and schedules for participants to follow. These include pacing guides and bell
schedules among others. Finally, as the organization encounters changing conditions, the
response in bureaucratic models is frequently the addition of supervisory and administrative
positions such as program coordinators.

Several of the educational reform efforts in the past three decades have utilized these
widely accepted mechanisms for exercising control within school systems. After the
National Commission in Excellence on Education stated in A Nation at Risk: The Imperative
for Educational Reform that the public educational system was jeopardizing national security
in 1983, a number of educational researchers called for a common core curriculum with clear
objectives, extended school days and school years, and increased amounts of homework
(DuFour & Eaker, 1998). According to Wiliam (2011), virtually all states in the 1980s used
bureaucratic assumptions as the basis for the flurry of resulting school improvement efforts which collectively became known as the Excellence Movement. The efforts were characterized by standardization, reliance on rules and regulations, and detailed specifications of school practices (DuFour & Eaker, 1998). This movement did not produce measurable improvements in student achievement and was replaced by the Restructuring Movement.

The Restructuring Movement, in contrast to the Excellence Movement, suggested comprehensive redesign and systemic transformation of the public school system. The Restructuring Movement reform efforts in 1989 produced eight national goals for education, collectively known as Goals 2000, to be accomplished within the following ten years. Within six years the focus of standards development was shifted from the federal level to the states. Moreover, the decision-making power for how to implement the standards was shifted to site-based management teams. While not standardized, the restructuring efforts commonly included site-based management with meaningful authority over staffing, program, and budget; shared decision making; staff teams with frequent, shared planning time and shared responsibility for student instruction; multi-year instructional or advisory groups; and heterogeneous grouping in core subjects (Newmann et al., 1996). It was hoped that local control would lead to improved teaching and learning. However, studies have consistently shown that practitioners elected to focus on non-academic, administrative issues in school improvement such as student discipline, parental involvement, and faculty morale (DuFour & Eaker, 1998). Graham (2005) points out that educational research has often focused on the periphery of schooling (e.g., the demographics of the students, the teachers, the funding) but not on the essence of schooling (i.e., what makes the students learn).
Similar to the Excellence Movement, the Restructuring Movement had little impact on classroom practices.

The recent ‘No Child Left Behind’ Act of 2001 is also based on the postpositivist philosophy and bureaucratic assumptions. Furthermore, the legislation typifies an empirical-rational strategy of change. The legislation assumes that researchers can use rigorous methods of scientific inquiry to ascertain what products and techniques produce significantly increased student achievement levels. As “rational and reasonable” beings, educators are assumed to want to adopt superior products and techniques (Chin, 1967). However, some people (e.g., Chin & Benne, 1969) also believe that organizations generally emphasize stability over change and are, therefore, resistant to change. To ensure that everyone will implement the products and techniques that are “scientifically” deemed to be superior, the legislation also takes a power-coercive approach and threatens sanctions in order to obtain compliance. Owens and Valesky (2007, p. 242) state, “both empirical-rational and power-coercive strategies for change share two assumptions: 1) that good ideas are best developed outside of the organization and 2) that the organization is the target of external forces for change.”

The nature of traditional professional development of in-service teachers has reflected an objectivist view of learning. Staff developers followed the ADDIE model of instructional systems design based on the belief that they could provide professional knowledge to teachers during workshops and that the teachers would then be able to transfer the newly acquired knowledge to their instructional practice. “ADDIE” stands for the first letter of each step in a traditional instructional systems design process: Analysis, Design, Development, Implementation, and Evaluation (Rothwell & Kazanas, 2004). Within the
instructional systems design process, instructional designers convert the results of task analyses into specific performance objectives based on classification hierarchies. Gagné, Briggs, and Wager (1992) distinguish among intellectual skills, cognitive skills, verbal information, motor skills, and attitude in a classic classification scheme. By analyzing tasks and identifying the hierarchy of skills needed to accomplish the tasks, instructional designers can identify and remediate specific skills lacking in teachers’ instructional repertoires. Performance objectives, however, are focused on the demonstration of new skills at the end of the instruction and not on the application of the skills in the context of teachers’ classrooms (Rothwell & Kazanas, 2004).

This pull-out workshop model is efficient; however, research has cast doubt on the effectiveness of this staff development model (Sparks & Loucks-Horsley, 1989). Additional concerns regarding traditional professional development practices include (Diaz-Maggioli, 2004):

- Top-down decision making;
- The idea that teachers need to be “fixed”;
- Lack of ownership of the professional development process and its results;
- The technocratic nature of professional development content;
- Universal application of classroom practices regardless of subject, student age, or level of cognitive development;
- Lack of variety in the delivery modes of professional development;
- Inaccessibility of professional development opportunities;
- Little or no support in transferring professional development ideas to the classroom;
- Standardized approaches to professional development that disregard the varied needs and experiences of teachers;
- Lack of systematic evaluation of professional development, and;
- Little or no acknowledgment of the learning characteristics of teachers among professional development planners.

The postpositivist emphasis on objectives extends beyond learning to assessment and evaluation. In fact, Ralph Tyler (1942), the father of modern educational evaluation, is
known for objectives-oriented evaluation whereby intended outcomes (objectives) are compared to well-measured actual outcomes. Alkin and Christie (2004, p. 19) state, “Objectives-oriented evaluation has had a strong continuing influence on education for many decades.” Evaluators whose approaches are consistent with the postpositivist philosophy such as Campbell (1966), Weiss (1991), and Cronbach (1982) focus on methodological issues such as the necessary conditions for experimental studies, internal validity, external validity, construct validity, reliability, generalizability, and “scientific rigor” in general.

In the past forty years, more research has been based on a different philosophical paradigm, social constructivism. The social constructivist, or interpretivist, paradigm developed from a convergence of several 19th- and 20th-century intellectual traditions, including German idealist philosophy, phenomenology, hermeneutic philosophy, and American pragmatism (Lindlof & Taylor, 2002). Some of the perspectives that have been labeled constructivist include Kant’s notions of knowledge and experience, feminist theorists’ views on knowledge construction, Kuhn’s work on scientific paradigms and revolutions, Piaget’s theory of cognitive development, and Dewey’s assumptions about knowledge and experience (Merriam & Caffarella, 1999).

Consistent with socioconstructivism are several related concepts from social psychology which lay a foundation for viewing cognition as a fundamentally social activity: distributed cognition, socially shared cognition, and situated cognition. These concepts portray cognition as an interpersonal as well as an intrapersonal process (Thompson, Levine, & Messick, n.d.).

Distributed cognition holds that information is not necessarily located in individual brains, but is distributed across networks of humans, artifacts, environment, culture, and
social interactions (Dillenbourg, 1996). While the aim of distributed cognition is to contribute to system design and implementation (Rogers & Ellis, 1994) and is most frequently discussed in examinations of the dynamics of interaction with technology, the concept of distributed cognition actually underlies the justification for collaboration in most constructivist-based instructional design theories. Since information is shared among the members and artifacts of a learning community, the learner must trust and rely on others in order to construct a robust understanding of the learning target.

Socially shared cognition also looks at the interaction between the learner and the environment and the subsequent impact upon learning. Grounded in the ideas of Dewey, Vygotsky, Durkheim, Piaget, and Mead, socially shared cognition posits that thinking and cognitive development involve participating in forms of social activity constituted by systems of shared rules which have to be grasped and voluntarily accepted (Cole & Nicolopolou, 1990). The shared rules will vary with each cultural context. Therefore, diversity must be embraced in order to sustain cognitive development.

Another aspect of the interpersonal nature of cognition is the concept of situated cognition. Situations are considered to “co-produce knowledge through activity”. In other words, the activity, context, and culture in which knowledge is developed and deployed are not separable from or ancillary to learning and cognition. The concept of situated cognition highlights the importance of enculturation and authentic activity (Brown, Collins, & Duguid, 1989).

Cognitive apprenticeship methods try to enculturate students into authentic practices through activity and social interaction in a manner similar to craft apprenticeship. Collins et al. (1989, p. 456) define cognitive apprenticeship as “learning-through-guided-experience on
cognitive and metacognitive, rather than physical, skills and processes.” In the first phase of cognitive apprenticeship, coaches make tacit processes visible to learners so the learners can observe and practice them (Collins et al., 1989). Methods that support making tacit processes explicit include modeling, explaining, coaching, scaffolding, reflecting, articulating, and exploring (Enkenberg, 2001). Then the coaches scaffold the activities and support the students’ attempts at doing the task. Finally, the coaches empower the students to continue independently (Brown, Collins, & Duguid, 1989). The ultimate goal of the cognitive apprenticeship is to create self-regulated learners.

The assumptions that underlie qualitative research are more consistent with the social constructivist paradigm and they differ significantly from those assumptions that underlie postpositivist-based quantitative research. The first assumption is ontological in nature. Unlike postpositivism where there is considered to be one truth whether people are aware of it or not, under social constructivism reality is considered to be subjective. People are believed to develop subjective meanings of their experiences through interactions with others and through the historical and cultural norms in which they have been socialized. The intent of research is to report the multiple realities. This is often achieved through the use of numerous quotes from varied individuals in order to present different perspectives (Creswell, 2007).

The second assumption is epistemological in nature. Under postpositivism, the researcher tries to remain distanced from that which is being researched in order to maintain objectivity. In contrast, social constructivist researchers try to get as close as possible to the participants who are being studied in order to better interpret the meanings that they have
constructed. This means that context is important and that researchers frequently need to conduct studies in the field (Creswell, 2007).

The third assumption is axiological. Under postpositivism, researchers are considered to be able to keep their values out of their research through the use of tightly controlled research methods. Social constructivists, however, believe that no one, including researchers, can function independently of their values. Therefore, researchers need to explicitly report their own values and biases as well as the value-laden nature of the information gathered in the field (Creswell, 2007).

These assumptions affect both the methodology and the language of research. Rather than starting with an *a priori* theory, social constructivist researchers inductively develop a pattern or meaning. The research can be characterized as inductive, emerging, and shaped by the researcher’s experience in collecting and analyzing the data. Since the logic is developed from the ground up, research questions and data collection strategies may evolve over the course of the study (Creswell, 2007). Unlike quantitative researchers who focus on generalizability, qualitative researchers are most concerned about issues such as credibility and validation. In order to validly represent the multiple points of view, researchers frequently use metaphors and stories with first person point of view in their writing.

Under the social constructivist philosophy, education is considered to be a process and school is defined as a lived experience (Merriam, 1998). Educational researchers should use interviews, observations, and document reviews to understand the process or experience the participants engage in within the context of “school”. The researchers document the experience of participants and the meanings that the participants construct.
Social constructivism underpins the human resources development approach to organizational theory. According to Owens and Valesky (2007, p. 117), “human resources development emphasizes using the conscious thinking of individual persons about what they are doing as a means of involving their commitment, their abilities, and their energies in achieving the goals of the organization.” While the bureaucratic model advocates firm hierarchic control and vertical communication, the human resources model posits that coordination can be achieved through the socialization of employees to organizational goals and values (Doyle & Hartle, 1985). Through an emphasis on expert power, lateral communications, and flexible structures, collaborative inquiry groups can be viewed as open systems that are capable of adapting to changing environmental conditions (Robbins, 1976).

Instead of reifying educational organizations, proponents of the human resources development perspective believe that organizations are invented social realities. Osterman and Kottkamp (2004, p. 22) state, “while organizations certainly exert powerful influences on the people who inhabit them, they remain human creations guided by human intentions and decisions, and individuals have the potential to shape them to their purposes.” Therefore, employees at all levels have the potential for improving organizational performance through their own growth and development.

McGregor (1960) in his Theory Y elaborates on assumptions about employees:

1. If it is satisfying to them, employees will view work as natural and as acceptable as play.
2. People at work will exercise initiative, self-direction, and self-control on the job if they are committed to the objectives of the organization.
3. The average person, under proper conditions, learns not only to accept responsibility on the job but also to seek it.
4. The average employee values creativity—that is, the ability to make good decisions—and seeks opportunities to be creative at work.
Adherents of the human resources development model posit that lasting organizational change is most likely to occur when the initiative originates from within the organization.

Lasting change occurs when employees identify, assess, and change underlying beliefs and assumptions. In contrast to the empirical-rational and power-coercive strategies for change, the normative-reeducative strategy posits that (attitudes, beliefs, and values can be deliberately shifted to more productive norms by collaborative action of the people who populate the organization (Greenfield, 1973). Through organizational self-renewal processes, an organizational culture can be created that actively supports the view that much of the knowledge needed to plan and carry out change in schools is held by the people within the schools.

The social constructivist philosophy and human resources development model call for a reflective practice model of professional development. Table II-1 contrasts the reflective practice model with the traditional, postpositivist model of staff development.

<table>
<thead>
<tr>
<th>Table II-1: Comparison of Two Approaches to Professional Development</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
</tr>
<tr>
<td><strong>Assumptions</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Strategies</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

*Source: Osterman & Kottkamp, 2004, p. 16.*
These assumptions and strategies draw on the principles of experiential learning, and situated cognition as well as social constructivism (Osterman & Kottkamp, 2004). The most important of these principles is that learning is an active process which requires the learner to construct knowledge based on prior knowledge and experiences. Furthermore, learning is more effective when socially constructed and grounded in a context relevant to the learner. One implication of these principles is that teachers must be respected as self-directed learners (York-Barr, Sommers, Ghere, & Montie, 2006). Given the importance of prior knowledge, experience, and context, a differentiated approach to professional development is essential.

Some of the social constructivist-based staff development practices that are currently advocated include:

- Programs conducted in school settings and linked to schoolwide efforts;
- Teachers participating as helpers to each other and as planners, with administrators, of in-service activities;
- Emphasis on self-instruction, with differentiated training opportunities;
- Teachers in active roles, choosing goals and activities for themselves;
- Emphasis on demonstration, supervised trials, and feedback; training that is concrete and ongoing over time; and
- Ongoing assistance and support available upon request (Sparks & Loucks-Horsley, 1989).

Consequently, the focus of staff development has been shifting from the delivery of workshops to more reflective inquiry into teachers’ own instructional practices (Killion & Harrison, 1997).

Given that knowledge is socially constructed, many educational researchers have advocated the use of collaborative inquiry teams known as professional learning communities as a form of professional development. The term ‘community’ implies a shift from the view of schools as a rational organization featuring linear lines of communication, chain-of-command decision making, differentiation of tasks, hierarchical supervision, and
formal rules and regulations to a different view of how schools should function (Scribner et al., 1999). Community, in general, can be defined as common experiences which foster a sense of shared identity, connectedness, trust, belonging, and mutual dependence (Furman, 1998). The concept of ‘professional community’ is a distinct strand of community (Scribner et al., 1999). Professional community refers to the more inclusive support of a whole school in order to promote collaboration among staff members, break down barriers that isolate teachers in their work, and support improved professional practices (Louis et al., 1996).

Other terms that are frequently used in the educational literature are ‘professional learning community’ and ‘teacher community’. Seashore, Anderson, and Riedel (2003, p. 3) state,

By using the term professional learning community we signify our interest not only in discrete acts of teacher sharing, but in the establishment of a schoolwide culture that makes collaboration expected, inclusive, genuine, ongoing, and focused on critically examining practice to improve student outcomes…The hypothesis is that what teachers do together outside of the classroom can be as important as what they do inside in affecting school restructuring, teachers’ professional development, and student learning.

Some of the key characteristics of professional learning communities (PLCs) are shared values and vision, collective responsibility, reflective professional inquiry, collaboration, and both individual and group learning (Hord, 2004; K.S. Louis, Kruse, & Bryk, 1995; Stoll, Bolam, McMahon, Wallace, & Thomas, 2006). Teachers discuss their practices related to students, subjects, and classroom settings within the professional learning communities.

Simons, Kushner, & James (2003) characterize the collective interpretation of practices as “situated generalization”. Teachers’ knowledge of their practice within that context complements and interacts with research-based knowledge. The collaborative discussions of teachers can help build knowledge by “providing a social life for information”
(Brown & Duguid, 2000). In this manner data and knowledge of practice can be transformed into new understandings through discussion and reflection.

Most studies of community examine the behavior of existing groups. Grossman, Wineburg, & Woolworth (2001), however, have empirically identified markers of teacher community formation. They identified continuums of behavior along four dimensions. Members of mature teacher communities identify with the group as a whole rather than with a subgroup, value the diverse perspectives of members of the community, and take communal responsibility for and regulation of group behavior. Furthermore, members of mature teacher communities deal with differences of opinion openly and honestly. They recognize that teacher learning and student learning are intertwined, and they are committed to the professional growth of all the community members.

Just as philosophical assumptions affect the design of professional development, they also affect the design of evaluation. Alongside postpositivist-influenced, methodological-focused approaches to evaluation, there are also social-constructivist-influenced approaches. Value-driven evaluators like Scriven (1974), Stake (1975), and House (1993) believe that the values of the evaluators should be made explicit and that evaluators should place value both on the findings and on the evaluation process (Alkin & Christie, 2004). Furthermore, other approaches to evaluation emphasize how participants interact with the evaluation process in order to maximize the use of the evaluation findings. Researchers who are concerned with the meaning created for participants include Patton (2004), Alkin (2004), Cousins (2004), Preskill (1999), and King (2004).
While post-positivism and social constructivism stem from different philosophical assumptions, both approaches have merit. Furthermore, it is possible to create systems that balance both approaches.

**Individual Learning**

In order to create long-lasting change, schools must unite disparate individuals to create and implement a shared mission and a vision that focuses on student achievement. Teachers, administrators, and other school staff members are individuals. As a disparate cadre of individuals, teachers, administrators, and other staff members have their own constellation of innate learning style preferences. These learning style preferences include, but are not limited to, preferred modality (e.g., visual, kinesthetic, verbal, etc.), preferred degree of social interaction, and preferred environmental conditions for learning. School personnel also come to the school from diverse backgrounds, bringing a variety of acquired knowledge and skills with them. Throughout their lives, these individuals acquire content and procedural knowledge in both informal and formal learning situations. Furthermore, they develop performance skills such as speaking foreign languages or playing a sport. Individuals organize their knowledge in the form of mental schemas, or categories and scripts. These schemas are components of a tacit mental model which enables them to quickly and efficiently interpret events and respond appropriately.

Just as mental models help individuals to parse information efficiently, attitudes and beliefs influence attention to particular data. Gagné (1985) defined attitudes as “acquired internal states that influence the choice of personal action.” As individuals receive information, they try to assimilate the data by connecting the data onto established anchoring
ideas (Driscoll, 2000). Attitudes and beliefs act as a filter, or prism, through which incoming information and responses must pass.

There is a second filter through which the information must pass. Each individual has a unique or personal life dream, a purpose for working, and ensuing career goals. Many individuals have also made commitments to themselves to take actions which will help them actualize their personal goals and vision. When an action is consistent with an individual’s personal vision and goals, he or she will be more motivated to carry out that action.

When people “critically examine their tacit values, assumptions, beliefs, and mental frameworks within which they understand and solve problems” it can lead to radical generative learning known as double loop learning (Argyris, 1982). This type of learning contrasts with incremental single loop, learning where people do not question their mental models and assumptions. Since mental models and attitudes influence to which data a person attends, changes to mental models and attitudes can greatly change the way that a person responds.

**Group Learning**

Collaboration requires that teachers work together as a group or a work team. London and Sessa (2006, p. 304) define a group as

> two or more people who interact with each other to perform a task…Members have a shared objective, feel connected to the group, and recognize others as members. They coordinate their behavior, resources, and tools toward the project and tasks, and they share responsibility for the group’s collective outcomes.

Hackman and Wageman’s (2005) definition of work teams holds many similarities to London and Sessa’s definition of groups. They define work teams as full-fledged groups that perform tasks in social system contexts. Hackman and Wageman (2005) articulate three
distinguishing features of work teams. First, they are real groups with intact social systems that include boundaries, interdependence, and differentiated member roles. Second, work teams have collective responsibility for accomplishing a group task. Finally, work teams operate in a social system context and manage relationships with other individuals or groups in the larger social system together as a collective. Work teams frequently display characteristics of living systems. London and Sessa (2006) posit that work teams maintain and renew themselves using energy and resources from their environment. They further state that living systems hold forms or structures stable as information, materials, or viewpoints are transformed.

While organizational learning is the primary goal of school reformers, learning is a by-product of work team processes, not the primary purpose. London and Sessa (2006, p. 305) state that the purposes of a work team are to complete group projects, to fulfill members’ needs, and to maintain the structure and integrity of the group as a system. Hackman and Wageman (2005) evaluate team effectiveness in a similar manner. They also put primacy on the work product. Their first attribute of effectiveness is that the productive output of the team meets or exceeds the clients’ standards. Their second attribute is related to the group’s need to maintain the structure and integrity of the group as a system. They state that to be effective, “the social processes the team uses in carrying out the work enhance members’ capability of working together interdependently in the future (p. 272).” Finally, the group experience fulfills members’ needs by contributing positively to the learning and personal well-being of individual team members. Interestingly, Hackman and Wageman relegate learning to an individual satisfaction category.
In contrast, London and Sessa (2006) articulate a nested hierarchy of living systems with learning occurring at the individual, group, and organizational levels. Focusing on group learning, they define group continuous learning as, “a deepening and broadening of the group’s capabilities in (a) (re)structuring to meet changing conditions; (b) adding and using new skills, knowledge and behaviors; and (c) becoming an increasingly high-performing system through feedback and reflection about its own actions and consequences (p. 305).”

Individuals influence the learning of the group through their insights and motivation. Organizations may also influence the learning of the group through pressure, challenges, and opportunities. Whether the group responds to the stimuli from individuals or responds to the organization is determined by the group’s readiness to learn. London and Sessa (2006) state that readiness depends on the tools that the group has in its repertoire, its boundary permeability, its stage of development, and its sensitivity to other systems.

In order for work teams to be effective, Wageman (2001) posits that four general structural and contextual conditions are required. First, the group must function as a real team with clear membership that is reasonably stable over time, thereby providing the capability for members to behave as a collective. Second, the work team must have clear objectives to be achieved. Third, the team must have an “enabling structure”. In other words, there must be an appropriate team size, optimal skill diversity, task interdependence, challenging task goals, and articulate strategy norms. Finally, there must be a supportive organizational context which includes a collective reward system, an information system that facilitates communication, an education system to provide training or technical consultation, and sufficient material resources to carry out the work.
The needs of work teams, however, are not static over their life span. Many researchers have proposed conceptual frameworks of fixed stages of group development with each successive stage contingent upon the successful completion of the prior stage. One of the most prominent models is the “forming-storming-norming-performing” model proposed by Tuckman (1965). Recent research has cast doubt about the generality and validity of stage models (e.g., Ancona & Chong, 1999; Gersick, 1988). Gersick (1988) found that groups developed a distinctive approach toward their task from the beginning and stayed with that approach until precisely halfway between the first meeting and the project deadline. At the midpoint, teams underwent major transitions and became focused on task execution until close to the project deadline. Finally, termination processes dominated the attention of team members.

The impact of personality preferences is felt in every element of team dynamics. When team members have different preferences, they are more likely to misinterpret each other’s actions. However, they are also more likely to consider multiple perspectives when solving problems. Smylie (1995) suggests that creativity and innovation may be constrained if teachers only have access to others with similar ideas and experiences.

One framework for analyzing team dynamics is the Myers-Briggs Type Inventory. During World War II, Isabel Myers and her mother, Katherine Briggs, began to develop an assessment of Carl Jung’s theory of Psychological Types with the goal of improving self-awareness, communication, and trust. The Myers-Briggs Type Indicator Assessment became available to the public in 1975 and has formed the basis of a significant amount of research into the impact of psychological preferences on learning styles, communication, conflict resolution, problem solving, and decision-making (Rutledge & Kroeger, 2005).
The theory of psychological type posits that people have a hard-wired preference in such areas as how people gather data, what kinds of information they pay attention to, and how they react to the information. While people can and do access all functions, the theory asserts that people are predisposed to behave in certain ways, especially in times of stress (Rutledge & Kroeger, 2005). Each preference style is valid and useful. However, awareness of differences and honoring those differences in psychological type can diminish people’s frustrations and lead to greater organizational effectiveness.

Personality preferences, for example, impact people’s reaction to change initiatives. People who are often perceived as resistant to change may not have received the information that they needed in order to buy into the change initiative. People with a Sensing preference, for example, need to hear concrete, practical details. They will change if present realities dictate the need. They will focus on all parts and steps of the initiative and will look to either the past or the present for guidance (Kummerow, Barger, & Kirby, 1997). People with a preference for Sensing define long-range as three to six months long (Kise & Russell, 2008).

People with a preference for Intuition, on the other hand, look to the future for change. They may define long-term as three to five years long (Kise & Russell, 2008). Given their focus on big-picture possibilities, they may propose large-scale sweeping change. People who prefer Intuition are comfortable proceeding based on a model, general idea, or hunch (Demarest, 1997).

When thinking through the proposed changes, people with a preference for Thinking are likely to prize principles and objectively focus on the content of the change. They seek impersonal, objective data and appreciate “logical” approaches to problem solving (Kummerow, Barger, & Kirby, 1997). This contrasts to people with a preference for Feeling
who are apt to take the impact on individual people and relationships as a fundamental consideration. People with this preference are also likely to want to know how others feel about the proposed initiative (Demarest, 1997; Kummerow, Barger, & Kirby, 1997).

Psychological type also impacts people’s approach to problem-solving. Lawrence (1993) profiles how three personality types tackle problems. He states that a person with a preference for Extraversion, Intuition, Feeling, and Perceiving (ENFP) would naturally brainstorm and explore possibilities. This person would be excited by creating something new and dreams would govern with some consideration given to feelings. This type of person would be comfortable with transformational change. A person with a preference for Introversion, Sensing, Feeling, and Judging (ISFJ), on the other hand, would view problem solving more as an approach to adjusting day-to-day events to make life more satisfying. This type of person would be more comfortable with incremental change. Finally, a person with a preference for Extraversion, Intuition, Thinking, and Judging (ENTJ) would view problems as confusions that need logical analysis to clear them up. This type of person would have a natural bias towards applying a cause-and-effect formula to the situation. The desire for closure may block consideration of other appropriate facts and options.

Table II-2 summarizes how people with different preferences are likely to approach problems and make decisions:
Table II-2: Decision-Making Styles

<table>
<thead>
<tr>
<th></th>
<th>Sensing, Thinking</th>
<th>Sensing, Feeling</th>
<th>Intuition, Feeling</th>
<th>Intuition, Thinking</th>
</tr>
</thead>
</table>
| They pay most attention to | • Specific, realistic “hard” data  
• Past experience | • Specific information about people in their environment  
• The opinions and ideas of people who are important to them | • Relevant stories and anecdotes  
• Imagery, symbols and metaphors | • Patterns and meaning they see in the data  
• The long-range view |
| To make decisions they will typically | • Apply current standard operating procedures  
• Use tried and accepted measures | • Try to find an alternative agreeable to everyone  
• Emphasize the needs of the specific people they have identified as important | • Generate interesting new ways to see the problem—use their insights  
• Use analogies: “This is like…” | • Generate and test hypothetical alternatives  
• Judge solutions by their own conceptual framework |
| This results in decisions that may be slanted toward | • Established practices  
• The status quo | • Socially desirable solutions  
• Decisions that everyone feels at least reasonably good about | • New, novel solutions  
• Inspirational programs that will create enthusiasm | • Their own system of understanding  
• Overly rational, unifying models |
| May ignore | • Patterns | • Data | • Practical considerations | • Data that doesn’t fit their mental model |

Sources: Kummerow, Barger, & Kirby, 1997; Kise, 2008

The different approaches to change, problem solving, and decision making can lead to miscommunication and conflict among collaborative teams. The communication of people
with Sensing and Thinking preferences may be experienced as blunt, rude, noncollaborative, impersonal, cold and bossy by other team members. People with a preference for Sensing and Feeling may be perceived to be too personal, phony, and superficial, as well as overly compliant with authority. People with a preference for Intuition and Feeling may be heard as rambling, vague, overly optimistic, idealistic, and irrelevant. Similarly, people with a preference for Intuition and Judging may be experienced as arrogant and pompous, overly complex and theoretical, detached, and combative. Even the way conflict itself is handled varies by psychological type. Conflict is perceived to be normal and is tolerated by people with a Thinking preference whereas people with a preference for Feeling would rather avoid or bury conflict (Kise & Russell, 2008).

**Organizational Learning**

Organizational learning encompasses how individuals learn and how that learning is shared and transformed as it radiates throughout the organization. The step of creating collective knowledge distinguishes organizational learning from individual learning (Stoll, Bolam, McMahon, Wallace, & Thomas, 2006). When an organization creates a culture in which continuous, focused, shared learning is fostered, it is considered to be a “learning organization”.

Organizational culture is, “a learned pattern of unconscious (or semiconscious) thought, reflected and reinforced by behavior, that silently and powerfully shapes the experience of people (Deal, 1985, p. 301).” Hofstede, Hofstede, and Minkov (2010) define organizational culture as “the collective programming of the mind that distinguishes the members of one organization from others.” They posit that an organization’s culture is holistic, reflects the history of the organization, relates to rituals and symbols, and is created
and preserved by the group of people who together form the organization. While Deal and Kennedy (2000) also cite rites and rituals, heroes, and values as elements of an organizational culture, they also elaborate that the environment in which the organization operates determines what it must do to be successful. Deal and Kennedy (2000) further focus on the informal means of communication within an organization, the cultural network, which any initiative must utilize in order to be effectively implemented.

The concept of organizational culture is related to the concept of national culture; however, it is not identical. Children learn national values throughout the first ten years of their lives and they form the most basic values. Hofstede, Hofstede, and Minkov (2010) have identified four dimensions of national cultures: power distance, collectivism versus individualism, femininity versus masculinity, and uncertainty avoidance. Furthermore, Hofstede, Hofstede, and Minkov (2010) have also identified the influence of occupational cultures because entering an occupational field means the acquisition of both values and practices that are occupation-specific. These national and occupational values influence the practices demonstrated within an organizational culture because these values are part of the invisible software of employees’ minds. These values influence motives, emotions, and taboos.

For example, the United States ranked the highest of seventy-six countries surveyed on an Individualism Index (Hofstede, Hofstede, and Minkov, 2010). Employed persons in an individualistic culture expect to act according to their own interests. American teachers have also been assimilated into an occupation that has valued academic freedom. It is, therefore, not surprising that many American schools are characterized by organizational cultures where teachers view themselves as independent contractors.
Owens and Valesky (2007, p. 201) state that there is overwhelming agreement that, “organizational culture is powerful in determining the course of change in an organization. Not a few believe that it is often the most powerful determinant.” Furthermore, Fullan (1992) has argued that any attempt to improve a school without attending to the school culture is “doomed to tinkering” because the school culture influences readiness for change. Lasting organizational change, therefore, is most likely to occur when the initiative originates from within the organization. However, Hofstede, Hofstede, and Minkov (2010) also posit that there is widespread agreement that it is difficult to change an organization’s culture.

Kise (2006, p. 55) describes the type of collaboration that has been advocated for creating long-lasting changes in instructional practices:

Teaching teams engage in deep discussions about teaching and learning, serving as resources for each other in developing curriculum and lessons that meet the needs of all learners. Together, they unearth assumptions about teaching and learning, gain from each other’s natural strengths, share strategies and ideas, and learn more about what is possible in the classroom.

This type of collaboration, however, is rarely found in schools today. Eaker, DuFour, and Burnette (2002, pp. 10-11) have stated, “It has been said that the traditional school often functions as a collection of independent contractors united by a common parking lot.” Typically, teachers have teamed for administrative tasks, cooperated on the creation of the occasional cross-disciplinary student projects, shared students for certain subjects, or implemented uniform rules and consequences for behavior (Kise, 2006).

The concept of professional learning communities is somewhat similar to the concept of communities of practice. Communities of practice are informally bound, self-organizing groups of people who exchange and interpret information in an effort to deepen their own understanding of what their practice is about (Wenger, 1998). Membership in the group is
voluntary and is defined by knowledge, not by task. Members of the group discuss novel ideas, problem solve, and retain knowledge in “living” ways. Professional learning communities are similar in that teachers collaboratively examine emergent problems within their own instructional practice, seek new perspectives, and experiment with potential solutions. Unlike communities of practice, however, membership in professional learning communities is frequently formalized through one’s role in the school (e.g., second grade teacher) and school leaders have often mandated participation in an effort to develop a collaborative organizational culture.

The type of collaboration that is required within real instructional contexts is different than the type of collaboration that is frequently required in the context of training. For efficiency purposes, instructional designers frequently teach problem solving based on well-structured problems. With well-structured problems the solution strategy is usually predictable and all relevant information is available to learners (Kirkley, 2003).

Learners often have difficulties transferring the information processing strategies they learn in this context to their actual work situations because the problems they encounter at work are of a different nature. Ill-structured problems are frequently vague, unpredictable, and may have some information missing. There may be several plausible solutions, or there might not be a fully satisfactory solution (Kirkley, 2003). With ill-structured problems, learners often face uncertainty about which concepts, rules, and principles are necessary for the solution (Jonassen, 1997). Learners must make and defend judgments of the nature and scope of the problem, offer possible solutions, assess potential impacts of solutions, and evaluate criteria.
In the context of school-based professional learning communities, school practitioners bring forth “problem cases” from their own instructional practice. These cases may center on a few struggling students or on lower-than-desired performance by the entire class. Problem solving models for ill-structured problems emphasize perception and pattern recognition as well as divergent and creative thinking in order to generate as many alternative representations of the problem as possible (Polson & Jeffries, 1985). Furthermore, there is an emphasis on finding the interrelationships between problems and solutions and integrating insights (Marquardt, 2004). Consequently, the use of reflective questioning rather than past knowledge and assumptions as a reference point is crucial.

The purpose of the questioning is to deepen the team members’ understanding of the situation they are addressing. Marquardt (2004, p. 77) states, “the key to problem solving is to start with fresh questions, not constructs and assumptions from the past. Questions enable groups to unpeel the layers around the problem and uncover the core elements of knowledge necessary to discover the solution.”

With ill-structured problems, problem solvers need to reconcile conflicting conceptualizations of the problem (Churchman, 1971). Indeed, Jonassen states, “Having identified the problem and the goals that different people have, it is important for the problem solver to identify all of the various perspectives, views, and opinions on that problem because ill-structured problems usually have divergent or alternative solutions. Ill-structured problem solving is a process of reflective judgment in which learners reconcile the uncertainty of knowledge through the process of inquiry into their beliefs (1997, p. 80).”

While learners are identifying the various positions and inferring how people who hold that position would solve the problem, they also need to gather concrete data about the
problem situation. School practitioners may present problem cases based on anecdotal observations. They need to augment their impressions with archival records including test scores and portfolios and/or data collected through observations, surveys, and interviews. Some of the information needed for making good decisions may reside in the school’s knowledge management repository.

Once learners have identified the various perspectives of the problems, examined the data, and inferred potential solutions, they need to examine constraints on the potential solutions before selecting the best solution given the context. Constraints may include a lack of resources such as money or time. Constraints could also include the lack of authority or lack of support from leadership to make the systemic changes needed to support the chosen solution. While some constraints will be known from the beginning, others may emerge during the collaborative problem solving process.

Learners’ mental models are developed through the process of refining the problem representations and arguing either for a preferred solution or against an alternative solution. Learners construct a solution, reflect on how they came to that decision, and justify that solution. By iteratively restricting the alternative arguments, learners arrive at the most defensible solution to test through active experimentation.

Once the members of the professional learning community have selected the most viable solution to test, they need to create an action plan for implementation. After the team has prepared a joint action plan, they try out the solution in their classes. In order for these collaborative problem solving teams to be able to function effectively, to be able to challenge the status quo, they must be empowered to take risks. They must be allowed to experiment
and to occasionally fail. However, the teams also must gather data about the progress of their initial implementation in order to make a formative evaluation of the solution.

The most important phase of the collaborative problem solving process is the growth phase. Throughout this stage the dialogue process is interwoven through individual and public reflection. Individual reflection is important because when people “critically examine their tacit values, assumptions, beliefs, and mental frameworks within which they understand and solve problems” it can lead to radical generative learning known as double loop learning (Argyris, 1982). This type of learning contrasts with incremental, single loop, learning where people do not question their mental models and assumptions. Since mental models and attitudes influence to which data a person attends, changes to mental models and attitudes can greatly change the way a person behaves.

Dialogue and collaboration can promote greater levels of learning than individual reflection can. Through dialogue, discussion, and a critical analysis of competing ideas, teachers can clarify their thinking and discover areas of cognitive dissonance. Fullan (1999) states that dialogue is even more valuable when it incorporates intellectual conflict because contrasting ideas and alternate explanations stimulate engagement and challenge learners to refine their thinking. Furthermore, dialogue can model strategic thinking for peers. Based on the Vygotskian perspective that self-regulation through speech is a precursor to the internal regulation of thinking, Pugach and Johnson (1990) suggest that explicit dialogue in socially interactive settings between more and less skilled individuals can provide practice for the kinds of internal dialogue that characterizes strategic thinking.

The professional learning communities should continuously question their assumptions and routines, test new possibilities, reflect on the results, and make adjustments
as necessary. This gentle challenging can also make organizations aware of discrepancies between espoused theory and theory-in-use. Their tacit mental maps govern their actual behavior, their theory-in-use. Their behavior frequently contradicts their spoken, explicit beliefs, or their espoused theory (Argyris & Schön, 1974). Reflection and dialogue can make learners aware of the discrepancy and provide them with an opportunity to bring their behavior and their beliefs into greater alignment.

Public reflection and focused dialogue among collaborative team members are also essential to the process of sharing knowledge. Much of the knowledge created by educators is “hidden, implicit, informal, and created in a myriad of ways” (Martin de Holan, Phillips, & Lawrence, 2004). Furthermore, much of the knowledge has been created in a fragmented manner. Therefore, public reflection and focused dialogue will help educators to articulate and disseminate their tacit knowledge.

Nonaka (1994) posits that organizational knowledge is generated through the interaction of four knowledge creation processes and starts at the individual level and moves to a group level, and finally to an organizational level. First, individuals can acquire tacit knowledge from other individuals through a socialization process whereby individuals observe and imitate. Second, individuals can also generate new knowledge through the “sorting, adding, recategorizing, and recontextualizing of explicit knowledge” (Nonaka, 1994, p. 19). Third, individuals can convert explicit knowledge into tacit knowledge through an internalization process. Fourth, tacit knowledge can also become explicit through an externalization process. Nonaka postulates that knowledge is created through the continuous dynamic interaction of these four processes as groups dialogue and engage in an iterative inquiry process.
The professional learning communities create ways to institutionalize the most valuable of the shared knowledge through the knowledge management system. The knowledge management system uses technological tools such as discussion boards, document repositories, and links to resources to facilitate not only access to shared information but also to facilitate the creation of new group knowledge.

Uncovering assumptions is also important in order to address the emotional fears related to change. Kegan and Lahey (2001) argue that hidden assumptions bolster our fear of change and work to sustain our current behavior. Articulated fears may be reframed and teachers may be encouraged to risk the vulnerability that occurs in the change process. Collegial conversations need to be focused so that difficult issues such as fears are actually addressed. A study of action research groups over a 6-year time period has shown that uncontroversial topics were chosen rather than topics which might have “shaken up” the established instructional routines of individual classrooms (Allen & Calhoun, 1998).

Conversation protocols can guide teachers as they engage in the psychologically risky process of inquiry. Protocols generally have two main features. First, they provide a structure for the conversation. In other words, they provide a series of steps that a group follows in a fixed order. Second, they specify the roles that different people in the group will play (Allen & Blythe, 2004). There are a variety of protocol types. They vary along a continuum from question finding to problem solving and vary in terms of the role of context, the roles of observation, interpretation and evaluation, and the role of a focusing question (Allen & Blythe, 2004). McDonald, Mohr, Dichter, and McDonald (2003) argue that protocols force transparency in the process and make clear the differences between talking and listening, between describing and judging, and between proposing and giving feedback.
Structured dialogue is important in order to produce a dialogue rather than a discussion. York-Barr, Sommers, Ghere, and Montie (2006, p. 52) define dialogue as, “an inclusive and opening-up process in which participants discover new perspectives and connections…Dialogue holds and reveals many possibilities without constraint.” In contrast, the purpose of discussion is to narrow options and make decisions. Refraining from making rapid judgments is important in the inquiry process. Cochran-Smith and Lytle (1999) argue that teachers need to develop an “inquiry stance” which tolerates ambiguity and avoids the quick fix.

Growth also occurs as team members formatively evaluate the progress of the implemented solution and make appropriate refinements based on the analysis. The team members then develop and implement the modified solution, gather data about the second implementation, and again engage in the cycle of reflection and dialogue. This process repeats until the solution is found to be satisfactory to all stakeholders.

Once the team decides that the solution is satisfactory, the team should report their progress to all stakeholders. Stakeholders may provide feedback which would prompt further refinements. Reporting to stakeholders also provides the opportunity to develop more shared knowledge.

Finally, as part of the growth phase, professional learning community members should reflect on their own learning process. The final stage of the collaborative problem solving process is the integration of the solution into the standard of practice. The new knowledge and skills should become part of the instructional routine when appropriate. Thus, transfer of the professional learning into practice is an inherent part of this process.
The collaborative problem solving process is an iterative process. It is not an occasional, pull-out event. Therefore, as soon as the professional learning community has satisfactorily solved one problem, it should start the same process over by addressing another emergent, ill-structured problem.

**Culture of Evaluative Inquiry**

The use of data is central to the advocated reflective inquiry process. Data allows teams to determine whether their perceptions match reality. The use of data should become institutionalized in the organizational culture. In other words, many educational researchers advocate the creation of cultures of evaluative inquiry. Concurrent with the rise of reflective inquiry as an educational reform strategy, the field of program evaluation has been changing to also advocate the development of cultures of evaluative inquiry. King (2004, p. 338) states,

> the highest form of evaluation is one that lives independently in an organization…[It] lives in a natural setting and reproduces itself in its organizational context. It is longitudinal, it encompasses both process and product use, and it focuses on building individuals’ capacity to engage in evaluation.

She continues,

> My evaluation teaching expands its curriculum from the specifics of a given study to more general habits of thinking, what I call ‘bringing evaluation to life’. This is evaluation capacity building, the creation of evaluation community within an organization (p. 338).

In other words, evaluation becomes integral to the way staff members go about their work.

Preskill and Torres (1999) concur with King. They posit that the frequent pace of change in today’s organizations necessitates a reconceptualization of evaluation. Rather than viewing evaluation as an event, organizations should integrate evaluation inquiry processes into their daily work. They believe that “evaluative inquiry” provides a process for
collaborating on challenging issues. They state (1999, p. 55) “We propose that evaluative inquiry can not only be a means of accumulating information for decision-making and action (operational intelligence), but that it is equally concerned with questioning and debating the value of what we do in organizations.” Other researchers have also explicitly linked evaluative inquiry to organizational learning. Cousins, Goh, Clark, and Lee (2004, p. 101) state, “Evaluative inquiry, as part and parcel of organizational culture and operations, has the potential to serve as a potent means for organizations to develop their organizational learning capacity.”

Researchers posit that it is possible to intentionally institutionalize a shared commitment to the use of program evaluation and a culture that includes the purposeful socialization of newcomers through Evaluation Capacity Building (ECB) and process use. Patton (1998), for example, believes that the process of participating in the evaluation process can create a “culture of evaluation”. Some of the values within this culture are:

- clarifying, specifying, and focusing;
- being systematic and making assumptions explicit;
- operationalizing program concepts, ideas and goals;
- distinguishing inputs and processes from outcomes;
- valuing empirical evidence; and
- separating statements of fact from interpretations and judgments.

Other indicators of a culture of evaluative inquiry include a common, positive understanding of Evaluation Capacity Building work as “how things are done around here” and a cultural value of inclusiveness regarding the involvement of all stakeholders (Baizerman, Compton, & Stockdill, 2002). Cultures of evaluative inquiry also include transparent and participatory practice. Indeed, Earl and Katz (2002) note that a “culture of inquiry” requires not only the creation of time for the process and an internal sense of
urgency regarding the need to examine data and undertaking a course of action, but also requires the involvement of other people to serve as critical friends with the interpretation of the data. Preskill and Torres (1999) identified several other distinguishing characteristics of a culture of evaluative inquiry. In a culture of evaluative inquiry, members of an organization have integrated evaluative inquiry into their work processes. While trained evaluators may provide training to members of the organization, members of the organization take a primary role in the undertaking of the inquiry. Also, the evaluative inquiry is not episodic or event-driven. Rather the inquiry is iterative and self-renewing. Finally, the best indicator that a culture of evaluative inquiry has been established is if staff members “stop thinking about data as something that is done to the school and start thinking about data as something that is done by the school and for the school” (Sutherland, 2004, p. 289).

Interest in cultures of evaluative inquiry, Evaluation Capacity Building, and process use has been increasing over the past twelve years. Since Patton’s 1997 explication of process use, an increasing number of studies have focused on how and what stakeholders learn from their involvement in the evaluation process (Harnar & Preskill, 2007). Evaluation Capacity Building, mainstreaming evaluation, and process use have all been themes of recent evaluation conferences and journal issues. Furthermore, there has been an increasing commitment to involving stakeholders in evaluation processes (Harnar & Preskill, 2007). Since Evaluation Capacity Building is building the capacity of staff to engage in data-based decision making, it is currently a growth industry in local school districts across the U.S (King, 2002).

Advocates for developing cultures of inquiry focus on the use of evaluation processes to improve decision-making and organizational learning. For more than thirty years
questions about the utility of evaluation approaches have been raised (Caracelli, 2000). One frequent question has been, “With what kinds of decisions has evaluation proved helpful?” The terms used to describe the utility of evaluation have evolved over time. At first the term “utilization” was used. Weiss (1981) suggested a linguistic shift from “utilization” to “use” claiming that “utilization” embodies an inappropriate imagery of instrumental and episodic application. Kirkhart believes that the term “use” is also characteristically instrumental and episodic and further posits that the term implies purposeful and unidirectional influence (Kirkhart, 2000). Kirkhart prefers that the discussion center around the term “influence” which she describes as “broader than use, creating a framework with which to examine effects that are multidirectional, incremental, unintentional, and noninstrumental, alongside those that are unidirectional, episodic, intended, and instrumental (2000, p. 7).” The construct of “influence” is appropriate for examining cultures of inquiry where there may be unintentional consequences that result from evaluation processes and the consequences may accumulate incrementally from a variety of evaluative activities.

Collaborative Use-Focused Approaches to Evaluation

Collaborative approaches that focus on evaluation use tend to have one or more of the following justifications: pragmatic, epistemological, or political. According to Weaver and Cousins (2004, p. 21) to achieve the pragmatic goal, “Members of a community of practice engage with researchers or evaluators to produce knowledge that bears upon identifiable practical problems. To the extent that research is grounded in the context for use and thereby rendered meaningful to those responsible for problem solving, decision making, or policy making, the knowledge produced will be of greater use.” A second goal is to produce valid representations of underlying phenomena through social constructivist dialogue and
interaction. For example, in Stake’s responsive evaluation approach, various aspects of the program are explored from multiple perspectives and participants are encouraged to elaborate upon or amend the data and their interpretations (Stake, 2004). The final goal is normative and political in nature. Some approaches strive to provide a voice in decision-making for marginalized stakeholders through direct involvement and participation (Weaver & Cousins, 2004).

With the belief that participation in the evaluation process will lead to greater stakeholder buy-in and resulting use, many of the use-centered approaches are collaborative in nature. Cousins and Whitmore (2007) have identified a number of collaborative and participatory approaches to evaluation including practical-participatory evaluation, developmental evaluation, stakeholder-based evaluation, democratic evaluation, and transformative-participatory evaluation. Practical-participatory evaluation (Cousins & Earl, 1995), developmental evaluation (Patton, 1994) and stakeholder-based evaluation (Bryk, 1983; Mark & Shotland, 1985) are grounded in the pragmatic justification to improve decision-making and program quality. Conversely, democratic evaluation (McTaggart, 1991) and transformative-participatory evaluation (Gaventa, 1993) are rooted more in the political goals of empowerment and social justice in a pluralistic society. All of the aforementioned approaches are grounded in a social constructivist philosophy of knowledge construction.

While a number of approaches share the same goals, they differ from each other in terms of five process dimensions. The number of identified process dimensions has also evolved over time. Initially Cousins and Whitmore identified control of technical decision making, stakeholder selection, and depth of participation; however stakeholder selection was
eventually teased apart into three different dimensions of form: diversity of participating stakeholders, power relationship among participating stakeholders, and manageability of the evaluation implementation (Cousins & Whitmore, 2007; Weaver & Cousins, 2004).

Technical control of decision making indicates the degree to which the evaluator or non-evaluator stakeholders exert control over the technical decision-making process. The diversity of participating stakeholders indicates the degree to which primary, secondary, and tertiary stakeholders are included in the process. The power relationship among participating stakeholders refers to the differential in power and influence that different stakeholders normally exert. Within a practical participatory approach where stakeholder participation may be limited to primary stakeholders, the power differential may be slight. In a transformative participatory approach, the diversity of stakeholders may be great and marginalized stakeholders, who don’t normally have a voice in decision-making, may be included. The depth of participation refers to the degree to which stakeholders are extensively involved in a wide variety of evaluative tasks including data collection, data analysis and reporting. Finally, the greater the number of participating stakeholders and the greater their participation, the greater the likelihood that the inquiry process will be unwieldy and challenging to manage.

An outcome of the collaborative participatory approaches, various process dimensions, and evaluation capacity building activities is the use of evaluation findings and processes. Theorists have identified a number of evaluation uses. For many years, the literature focused on the direct, immediate instrumental use of findings in order to make decisions regarding the continuation, expansion, revision, or termination of programs (Caracelli, 2000). The concept of findings use has now been expanded to include conceptual
and symbolic uses (Mark & Henry, 2004). Conceptual use, or enlightenment, refers to a
general learning that takes place through the analysis of evaluation findings. Symbolic use
refers to the use of evaluation findings to justify a pre-existing opinion or previously made
decision.

The concept of process use formally entered the evaluation literature just over a
decade ago. Patton first defined process use as, “relating to and being indicated by individual
changes in thinking and behaving that occur among those involved in evaluation as a result
of the learning that occurs during the evaluation process (1998, p. 225).” He originally
suggested that four outcomes might occur as a result of involvement in the evaluation process
but has since included two more outcomes (Patton, 2007). These outcomes include the
reinforcement of general program and organizational development and specific program
interventions. The process of measuring focuses program resources on priorities
(instrumentation effects). The dialogue and inquiry processes lead to enhanced shared
understandings within the program. The evaluation process makes evaluation meaningful
and understandable to participants thereby increasing participant engagement and ownership.
Finally, the process can lead to the infusion of evaluative thinking into the organizational
culture.

The relationship between the related concepts of evaluation influence, evaluation use,
evaluation capacity building, process use, and the development of a culture of evaluative
inquiry is currently an issue of debate among leading theorists. Furthermore, few
comprehensive conceptual frameworks currently exist (Preskill & Boyle, 2008). The crux of
the debate centers on the role of intentionality. After a recent survey of American Evaluation
Association members in which less than one percent of respondents talked about process use
as something that is intentionally built into the evaluation process, Harnar and Preskill (2007, p. 40) concluded that “process use is more reflective of incidental or informal learning—a by-product that occurs from stakeholders’ engagement in the evaluation—while ECB represents the evaluator’s clear intentions to build learning into the evaluation process.” King (2007) believes that process use may be an outcome of evaluation participation although the degree to which that use is intended varies between contexts.

King further believes that intentional process use builds evaluation capacity. Patton, on the other hand, states, “process use is not itself capacity building; rather, it is capacity built (2007, p. 106).” In other words, the result of effective ECB may be process use. The relationship between ECB and process use is not always strong and direct. Patton (2007) points out that intended process use can include ECB, however not all intended process use involves ECB. He states, “much process use has a greater and more direct impact on program or organization processes and effectiveness than on evaluation capacity itself (Patton, 2007, p. 107).” Finally, not all ECB involves process use. For example, direct evaluation training may not be a part of an evaluation process. This training, however, may contribute to the establishment of a culture of evaluative inquiry.

**Evaluation Capacity Building**

In addition to choosing an approach and process dimensions appropriate to the context, evaluation use can be fostered through evaluation capacity building activities. Baizerman, Compton, and Stockdill (2002, p. 109) define evaluation capacity building (ECB) as, “the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine.” This definition has been frequently cited by other researchers. However, Preskill and Boyle (2008, p. 444) recently offered a new
definition. In part it states, “ECB involves the implementation of teaching and learning strategies to help individuals, groups, and organizations learn about what constitutes effective, useful, and professional evaluation practice. The ultimate goal of ECB is sustainable evaluation practice—where members continuously ask questions that matter; collect, analyze, and interpret data; and use evaluation findings for decision-making and action.”

The short-term goal of ECB is to increase evaluation related knowledge, skills, and attitudes (Patton, 2007). In order to achieve this, facilitators need to be cognizant of the motivations, the sets of assumptions about evaluation and capacity building in general, and the implicit or explicit expectations for what the ECB activities in particular will achieve (Preskill & Boyle, 2008).

ECB frequently is explicit and intentional. Some of the intentional teaching and learning processes that have been used to help people develop the knowledge, skills, and attitudes to think evaluatively and engage in the evaluation process are workshops, internships, written materials, meetings, appreciative inquiry, technical assistance, and coaching (Preskill & Boyle, 2008). However, not all ECB efforts are explicit. King (2002) has described a long-term ECB initiative in Minnesota in which ECB was developed through participation in practical-participatory evaluations and the development of an infrastructure for data collection, analysis, and presentation. Other less direct ways of developing ECB include participation in communities of practice, development of data-oriented technologies, and involvement in evaluation activities.

In order to intentionally foster increased knowledge and skills, the evaluator plays an expanded role as compared to the traditional postpositivism-based evaluator. The evaluator
works as a facilitator, educator, and critical friend in the hopes of fostering individual, group, and organizational learning. The evaluator also serves as a mediator as stakeholders discuss values and perspectives that sometimes conflict. Finally, the evaluator serves as a collaborator, coach and consultant, helping stakeholders to inquire about their practice in a real-world context (Caracelli, 2000; King, 2007; Morabito, 2002; Preskill & Torres, 2000). Baizerman, Compton, and Stockdill (2002, p. 116) sum up the chameleon-like nature of the evaluator who fosters ECB: “ECB practice is art-like, craft-like, and science-like, and the ECB practitioner has at hand all of these occupational orientations to use ‘as necessary’ in each emergent moment.”

**Elements of Sustainable Practice**

The long-term goal of ECB, however, is to develop a culture in which ongoing evaluative inquiry becomes a way of life. In order to sustain evaluation practice, the organization has to commit to internalizing evaluation processes. Preskill and Torres (1999, p. 52) state, “An organization’s infrastructure can strongly influence the extent to which organization members learn from evaluative inquiry and use their learning to support personal and organizational goals.” The structures and reward systems should support collaboration and cooperation (King, 2002; Preskill & Torres, 1999). Time and space for collaboration needs to be scheduled into the work day. Team members should have somewhat close physical proximity, space to meet, and interdependent roles (Kruse, Louis, & Bryk, 1995; Stoll, Bolam, McMahon, Wallace, & Thomas, 2006).

Communication is very important for a sustained change to the organizational culture. Conducive communication captures the diversity of voices of the participants and focuses on the sharing of learning. Communication sheds light on the ways that members talk about
evaluation, their inclination to ask evaluative types of questions, their interest in using data for decision making, and their overall commitment to conducting meaningful, timely, and useful evaluations. Furthermore, the communication should celebrate the organization members’ use of evaluation findings for decision making and action. As evaluation findings are used to improve programs, evaluation will more likely become embedded in the organization’s culture (Preskill & Boyle, 2008).

In order to sustain evaluative thinking, the organization must be able to create, capture, store, and disseminate evaluation-related data and documents. Knowledge management systems recognize and organize tangible and intangible assets such as human capital, intellectual property, best practices, lesson plans, and national resources. The human capital represents the experience and knowledge that each member of the learning community brings to the group. All teachers come with their own best practices and lesson plans from their past experiences that often provide a spark of new ideas or a possible solution to a problem. There are a growing number of national resources that can provide certain approaches or solutions to problems facing the professional learning community. These assets enable activities like accessing shared information and knowledge-building tools.

Some of the tools that can be used to facilitate knowledge-management activities include data warehouses, repository document managers, discussion boards, announcements, and links to outside resources. These tools are used to facilitate active learning conversations, the sharing of information, and both individual and group learning.

The active support of the organization’s leadership is critical for the development of a culture of evaluative inquiry. Preskill and Boyle (2008, p. 445) state, “We propose that the
extent to which and the ways in which the organization’s leadership values learning and evaluation, creates a culture of inquiry, has the necessary systems and structures for engaging in evaluation practice, and provides communication channels and opportunities to access and disseminate evaluation information will significantly affect not only if and how people learn about evaluation but also the extent to which evaluation practice becomes sustained.”

In order to have staff members diagnose their own organizational problems systematically and work out solutions for them, leaders must empower participative decision making. Participative decision making is important in contexts where emergent problems are ambiguous and ill-understood and where responses to rapid change need to be nimble (Lindelow et. al, 1989). While the leader retains the ultimate decision-making authority, all members should be encouraged to express their feelings, offer knowledge and information, and have their views considered.

The leader needs to foster an organizational culture that is conducive to encouraging the participation of all stakeholders. Owens and Valesky (2007, p. 259) cite the following characteristics of a growth-enhancing organizational culture:

- Is intellectually, politically, and aesthetically stimulating;
- Emphasizes individual and group achievement;
- Places high value on the personal dignity of individuals;
- Accepts divergent feelings and views in a nonjudgmental way;
- Is oriented to problem solving rather than to winning or losing in intraorganizational skirmishes.

In their study of organizational learning at an elementary school in the Canadian province of Saskatchewan, Mitchell and Sackney (1998) identified two cognitive processes, two affective processes, and three standards of conduct that were of primary importance. They assert that reflection and professional conversation, both cognitive processes, are crucial to teachers’
individual and collective learning. Furthermore, they discovered that certain affective processes greatly influenced the degree to which reflection and professional conversation could thrive. These affective processes were an invitation to participate and an affirmation of colleagues. These affective processes were complemented by standards of conduct that the teachers identified as essential to the development of effective organizational learning.

These standards of conduct were:

- All staff members are responsible for school success
- Diversity among individuals is honored and valued, and
- Psychological safety is maintained during group deliberations

These standards of conduct served to develop an atmosphere of trust in which teachers were willing to work with and learn from their colleagues.

Several studies have confirmed the importance of trust in the school improvement process. Trust is defined as a composite of vulnerability, benevolence, reliability, competence, honesty, and openness (Hoy & Tschannen-Moran, 2003). Bryk and Schneider (2002) developed a three-tiered theory of relational trust based on the concept that interdependence in relationships creates vulnerability. At the interpersonal level, “relational trust is rooted in a complex cognitive activity of discerning the intentions of others (Bryk & Schneider, 2002, p. 22).” These discernments are situated, however, within a set of interpersonal role relations that are influenced by the school culture. The interpersonal role sets include the relationship of teachers with the principal, the relationship of teachers with colleagues, and the relationship of teachers with parents. Finally, these trust relations culminate in important consequences at the organizational level.

Quite a few studies have demonstrated a myriad of positive associations between increased levels of relational trust and organizational conditions. Bryk and Schneider (2002),
Hoy and Tschannen-Moran (2003), Tschannen-Moran (2004), and Zand (1997) have demonstrated that increased levels of relational trust are associated with greater collaboration and group cohesiveness. Tschannen-Moran (2004) has shown a positive association with collective self-efficacy of teachers and more productive conflict resolution strategies. Bryk and Schneider (2002) and Kochanek (2005) have demonstrated a positive relationship with increased openness to innovation, improved problem solving, and increased collective responsibility for a shared vision. Moreover, Louis (2000) and Sergiovanni (2005) have shown that increased levels of relational trust are associated with an increased likelihood of transformational leadership. Schools with higher degrees of relational trust have also shown greater diffusion of innovative practices and increased sustainability of their reform initiatives (Bryk & Schneider, 2002; Louis, 2000; Mishra, 1996). Most importantly, these schools have also achieved higher levels of student achievement (Bryk & Schneider, 2002; Goddard, Tschannen-Moran, & Hoy, 2001). Therefore, developing trust is an essential part of the effort to develop effective collaborative work teams in the context of school systems.

School leaders need to recognize that trust evolves over time through repeated interactions and the current level of trust should determine the nature of the appropriate trust-building behavior. Kochanek (2005) has developed a process model of developing trust in schools. She posits that schools must begin with easing the vulnerabilities of participants in order to induce them to enter into exchanges. This is done by setting the stage with positive conditions and by facilitating successful low-risk interactions. Through repeated interactions in which participants gain knowledge of one another’s trustworthiness, they become more willing to engage in risky exchanges. Kochanek illustrates the connection of the trust-building behaviors to Bryk and Schneider’s theory of trust in Figure II-1 (2005, p. 19):
Figure II-1: Kochanek’s Process Model of Trust Building in Schools

School leaders, therefore, need to assess the current level of trust among members of a work team and purposefully build in trust-building interactions appropriate to the developmental level of the work team.

**Barriers to Developing Cultures of Evaluative Inquiry**

The predominant organizational culture of most school systems poses a significant barrier to the development of effective work teams which utilize reflective collaboration and evaluative inquiry to achieve higher levels of learning. Stein, Smith, Henningsen, and Silver (2000) describe a culture of silence and secrecy that discourages teachers from engaging in meaningful conversations about instruction for fear of being perceived as incompetent or as a troublemaker. Furthermore, teaching has frequently been viewed as a form of creative expression. There is a cultural acceptance of individual styles and preferences which leads to a bias towards noninterference (Little, 1990). This individualism has led to a lack of
common goals and meaning. Therefore, collaboration is perceived to have a high cost in terms of competency, status, and obligations (Kise, 2006). Finally, time for collaboration has rarely been incorporated into the work day schedule.

The combination of this organizational culture with increasing bureaucratic pressure to collaborate has led to a state which Hargreaves and Dawe term “contrived collegiality” in many schools. Hargreaves and Dawe (1990, p. 227) state that contrived collegiality “consists of administratively contrived interactions among teachers where they meet and work to implement the curricula and instructional strategies developed by others.” Contrived collegiality enhances administrative control rather than fosters teacher and curriculum development.

Furthermore, there are barriers to implementing evaluative practices. Shulha (2000, p. 39) states, “It is rare to find the structures and resources necessary to support teachers and administrators in intentional and continual inquiry into programs and practices.” Preskill and Torres (1999) cite additional barriers such as reactive organizational cultures and the fear of making mistakes. In addition, many organizational leaders believe that evaluative activities cost too much in terms of money, time, and personnel resources.

While it is difficult to develop cultures of evaluative inquiry within professional learning communities, it is even more difficult to sustain effective PLCs. Hargreaves (2004) has suggested that involving a broad range of people, nurturing people, sharing responsibility, and developing the capacity of people to adapt and learn from each other in complex environments are all factors in sustaining change. However, there is little longitudinal research that demonstrates how to create enduringly effective PLCs and there is
evidence that some PLCs that had been deemed effective later suffered a decline in effectiveness (Stoll, Bolam, McMahon, Wallace, & Thomas, 2006).

**Using ECB to Develop Cultures of Evaluative Inquiry**

While teachers may be used to outside evaluations, school improvement plans have typically focused more on lists of activities than outcomes “creating the possibility of going through yearly evaluation motions” (King, 2002). Mandated PDSA cycles have resulted frequently in hollow evaluation processes where goals were set, strategies were implemented, and data collected with little support or accountability for the results. Teachers need help and support to learn how to effectively frame evaluation questions, how to put together high-quality data collection instruments, how to make claims using data, and how to use data regularly as part of their instructional process.

Using a socio-constructive collaborative participatory approach, teachers can examine their assumptions about learning and use the evaluation results to improve decision-making and program quality. By giving teachers the authority to take risks, evaluate their practices and innovate, teachers are also provided a voice in the decision-making process and leadership is evenly distributed. Within PLCs, teachers would have considerable control over the technical decision-making process with the evaluator serving as a consultant or critical friend. The depth of teacher participation is great. The diversity of stakeholder roles is not as great since most participants will be members of the same grade-level or content area. The power relationship, in other words the ability to influence and control the process, is fairly equal among most participants. Hopefully, since the topics chosen are meaningful to the participants, they will find instrumental and conceptual uses for the findings.

Additionally, a socio-constructive collaborative participatory approach should also enhance
teachers’ engagement and ownership for the program intervention, and the organization as a whole, and increase their shared understandings.

There is empirical evidence that professional learning communities can improve teaching and learning. Several studies based on data from the National Longitudinal Study of 1988 revealed that teacher community had a positive statistical effect on student achievement gains (Lee & Smith, 1995, 1996; Lee, Smith, & Croninger, 1997). Newmann and Associates (1996) found strong correlations between professional community, authentic instruction and student achievement. Finally, the Center for Research on the Context of Teaching (2002) demonstrated that students did better in schools where teachers examined student achievement data together and collaborated to develop and assess interventions.

There are many descriptions in the research literature about the ways in which professional learning communities can work to improve student achievement through an evaluative inquiry process. Yet there is little evidence about which conditions, processes, strategies, and norms help schools develop productive evaluative inquiry teams (McLaughlin & Talbert, 2006). Indeed, Graham (2005, pp. 197-198) states,

While we know it is possible to create schools in which all or nearly all children prosper, we are not sure exactly what the constituent ingredients of such schools are. We know that they can exist for we have seen them, but we do not know how to construct or universalize them...Even if we are able to name the categories, we have not yet become able to organize the practices that follow from these categories...Our evaluation methodologies are much better at documenting policies that have been established than they are at discerning those practices which interact together to create conditions for a good school.

The scarcity of teaching teams which effectively implement cycles of evaluative inquiry attests to the difficulty in establishing and sustaining cultures which support the long-term implementation of these collaborative inquiry processes. There is a need to holistically
investigate the complexity of the process of developing a culture of evaluative inquiry within real-world contexts in order to provide school leaders with guidance how to accomplish this difficult task.
Chapter III: Methods

This study focused on the processes used by the staff members of North Mesa Elementary School, a public elementary school in the Suburban School District, to develop a culture of evaluative inquiry. The school leadership hoped that the development of an internally-driven, reflective culture of evaluative inquiry would lead to long-term gains in student achievement. The purpose of this chapter is to describe the research questions, the research site, the role of the researcher, the research design, and the procedures utilized in this case study.

Restatement of the Problem and Research Questions

Research shows that long-term student achievement trends have not been impacted by a number of top-down educational reform policies including a Nation at Risk in 1983, Goals 2000, the Improving America Schools Act in 1994, and NCLB in 2001 (Mintrop & Sunderman, 2009). In order to accelerate growth in student achievement, many researchers are now advocating the development of cultures of reflective inquiry centering on collaborative problem solving and the use of data. This study was a comprehensive investigation of how school staff members at a public elementary school in Suburban School District endeavored to develop a culture of evaluative inquiry. The study explored the following questions:

Central question:
How do staff members of a public elementary school develop a culture of evaluative inquiry?

Subquestions:
• How do members of the school leadership team increase their knowledge and skills related to evaluation and data analysis?

2 Pseudonyms have been used for the name of the school and the district in order to preserve the anonymity of the participants.
• How do members of the school leadership team apply their evaluative inquiry knowledge and skills to emergent issues within their school community?

• How do members of the school leadership team consciously build the evaluation capacity of teachers within their professional learning communities?
  
  o What training activities are held?
  o What roles do leadership team members play?
  o What learning processes are included in the evaluation capacity building activities?
  o How does the leadership team provide support in terms of communication, systems and structures and organizational culture?
  o How does the leadership team address the information needs of differing MBTI personality types (function theory)?
  o What challenges are encountered in the evaluation capacity building process?

• How does this initiative change the collaboration processes among teachers in the professional learning communities? Are there unintentional consequences in addition to the goals of the initiative?

• Do changes in demographics, student achievement, perceptions, and processes occur over the course of the case study?

Research Site

The unit of analysis for this case study was North Mesa Elementary School. North Mesa Elementary School is one of ten elementary schools in the Suburban School District located in a rapidly growing city of Suburbia\(^3\) in the U.S. Southwest. The population of Suburbia has grown exponentially over the past 40 years as Table III-1 shows, making Suburbia the third largest city in the state in terms of population:

---

\(^3\) A pseudonym was used to protect the anonymity of the participants.
In the ten year period from 2000 to 2010, the population of Suburbia grew by 69.1%. The high rate of growth is expected to continue over the long-term although the recent economic downturn has slowed the pace of growth over the short-term.

As the city of Suburbia has grown, the Suburban School District has also grown. In 1994, the Suburban School District was formed with the five elementary schools and one middle school which had been built within the city of Suburbia by two neighboring school districts. Constructed in 1995, North Mesa Elementary School was the first school built by the newly formed school district. The rapid growth in student enrollment, portrayed in Figure III-1, necessitated the hurried construction of additional schools and the frequent adjustment of school attendance boundaries.

<table>
<thead>
<tr>
<th>Year</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1964</td>
<td>100</td>
</tr>
<tr>
<td>1969</td>
<td>1,500</td>
</tr>
<tr>
<td>1974</td>
<td>5,377</td>
</tr>
<tr>
<td>1980</td>
<td>10,208</td>
</tr>
<tr>
<td>1990</td>
<td>32,505</td>
</tr>
<tr>
<td>2000</td>
<td>51,765</td>
</tr>
<tr>
<td>2005</td>
<td>66,599</td>
</tr>
<tr>
<td>2010</td>
<td>87,521</td>
</tr>
<tr>
<td>2020 projection</td>
<td>100,041</td>
</tr>
</tbody>
</table>

Sources: US Department of Commerce, Bureau of the Census; and U.S. Census Bureau Interim State Projections 2005; U.S. Census Bureau State & County Quick Facts.
As of August, 2009, there were ten elementary schools, four middle schools, two comprehensive high schools and two alternate high schools in Suburban School District. The student population at North Mesa Elementary School experienced similar growth. The student enrollment grew steadily each year with relief from overcrowding coming in 2005 and 2008 when the opening of new elementary schools in the district prompted the re-zoning of attendance boundaries and some students were reassigned to different schools within the district.
Figure III-2: Student enrollment at North Mesa Elementary School, 2002-2010.

Since only mature neighborhoods were included in the attendance boundaries drawn in 2008, the student enrollment remained relatively stable from 2008 to 2010, which was the time boundary of this study.

The rapid growth in population profoundly affected the demographic composition of the student body at North Mesa Elementary School during its thirteen year history. Over the past fifteen years the ethnic composition of the students across the district steadily moved from primarily Caucasian to a “minority-majority” composition as Figure III-3 shows:
Figure III-3: Distribution of Ethnicities within Suburban School District, 1994-2010.

While this trend was also reflected in the school demographics, the boundary changes created sudden shifts in the demographic composition of the school. Since re-zoning attendance boundaries adds and deletes entire neighborhoods from the school’s attendance zone, shifts in the socioeconomic and ethnic profile of the student body occurred approximately every three years. For example, one neighborhood consisting of smaller, older, and less expensive houses were rezoned to another elementary school southeast of North Mesa Elementary School in 2005. When a new elementary school was opened in 2008 one mile to the north, approximately 500 students from the newer and more expensive subdivisions were transferred to the new school and the students in the older neighborhood were transferred back into North Mesa Elementary School’s attendance zone.
The rezoning of the attendance district in 2008 changed the demographics of the school’s student body. From 2005-2007, Caucasians and Hispanics had approximately equal representation at the school with the two ethnicities comprising approximately 90% of the student body. In 2008, however, Caucasians represented just 38% of the student body and Hispanics represented 51%. While the percentage of Caucasians has slowly been declining across the district, the sudden gap at the school was due to re-zoning of the attendance boundaries.

Figure III-4: Ethnic composition of North Mesa Elementary School and Suburban School District.
The socioeconomic composition of the student body had also been shifting over the years. In 2000, 34.34% of the student body participated in the federal Free and Reduced Lunch program. By 2008, the start of this study, the proportion of participating students had increased to 62.43%, an 82% increase over eight years. Over the two years of this study, the percentage of students participating in the Free and Reduced Lunch program increased further to 72.73% of the total student population.

The Suburban School District has long enjoyed a reputation of high academic quality and the district’s scores on the statewide Standards Based Assessment (SBA) have consistently been higher than the statewide average. For the first time since the ‘No Child Left Behind’ legislation of 2001, North Mesa Elementary School failed to achieve Adequate Yearly Progress in 2007-2008. This event, combined with an increasing number of schools across Suburban School District which failed to make Adequate Yearly Progress, created a sense of urgency regarding the need to improve student achievement. The school leadership decided that developing a culture of evaluative inquiry was the best way to achieve the needed improvement in student achievement.

The focus of this inquiry was on the knowledge, beliefs, attitudes, and actions of the administrators and teaching staff of North Mesa Elementary School as they endeavored to develop a culture of evaluative inquiry. The school administration consisted of a principal and an assistant principal. Both the principal and the previous assistant principal started at North Mesa Elementary School in the summer of 2007. Both administrators described their goal for the first year of their tenure as “to survive” given the extremely overcrowded conditions at the school and a number of personal problems that beset each of the administrators during the 2007-08 school year. The previous assistant principal helped plan
this initiative but she left the school in November of 2008 for personal reasons. An assistant principal from a secondary school within the district was transferred to the site and decided to continue as the assistant principal at North Mesa Elementary School rather than to return to the secondary school.

The downsizing of the school in 2008 provided the school leadership with the opportunity for a “fresh start”. After attending a national Title I conference in February of 2008 the principal decided that sending a team from North Mesa Elementary School to a five-day data portfolio conference in the summer of 2008 would be a good way to kick-start an initiative to develop a culture of evaluative inquiry. In June of 2008, the principal stated that she hoped the conference would help the teaching teams organize their data and create an aligned school improvement plan. The administrative team’s goal for the year was to develop the functionality of the grade-level professional learning communities. Specifically, the administrators planned to focus on the structure of the PLCs and the development of norms, guiding questions, and protocols to help teams compare their perceptions to the data and engage in a reflective inquiry process.

**Role of the Researcher**

I served as a participant observer, coach, and researcher as the school worked to develop a culture of evaluative inquiry. I am currently employed as the Executive Director of Research, Assessment, Data, and Accountability in the district office of the Suburban School District. From March, 2008 through June, 2009 I served as a Research/Data Analyst and from July, 2009 to December, 2011 I served as the Director of Assessment and Data Analysis in the same office. Part of my professional responsibilities was to support the school with the collection, compilation, and interpretation of data from a variety of sources.
In that role, I was called upon to provide resources and coach users on the interpretation of assessment data. However, some of the activities of this inquiry such as observing meetings of the Instructional Leadership Team and grade-level professional learning communities fell beyond the scope of my normal professional duties.

**Research Design**

The focus of this study was on the process that school members used to develop a culture of evaluative inquiry. The most effective way to explore this process in the context of one school community was with qualitative research. Patton (1985, p. 1) describes qualitative research as:

> An effort to understand situations in their uniqueness as part of a particular context and the interactions there. This understanding is an end in itself, so that it is not attempting to predict what may happen in the future necessarily, but to understand the nature of that setting--what it means for participants to be in that setting, what their lives are like, what’s going on for them, what their meanings are, what the world looks like in that particular setting--and in the analysis to be able to communicate that faithfully to others who are interested in that setting...The analysis strives for depth of understanding.

Consistent with social constructivism, qualitative research assumes that reality is subjective and multiple. The goal is to obtain an insider, or emic, perspective through fieldwork and to report the different perspectives that may exist (Creswell, 2007).

A descriptive case study was the most appropriate method for investigating these research questions for several reasons. The first reason was that the questions dealt with links that needed to be traced over time (Yin, 2003). Second, the research questions focused on contemporary events. Finally, I had a limited ability to control the events and processes. As a participant observer who acted as a critical friend, I was able to occasionally informally
manipulate processes, but the locus of control resided with the school leadership and staff members.

Yin (2003, p. 13) defines a case study as:

An empirical inquiry that
- investigates a contemporary phenomenon within its real-life context, especially when
- the boundaries between phenomenon and context are not clearly evident.

Case study methodology was appropriate for this empirical context because the research questions focused on insight, discovery, and interpretation rather than hypothesis testing. Indeed, case study methodology is different from other research designs by what Cronbach calls “interpretation in context” (Cronbach, 1975). The knowledge gained is more contextualized and concrete than the knowledge gained from other research methods (Stake, 1995).

Merriam (1998) further describes the characteristics of case study methodology. She claims that case studies are particularistic, descriptive, and heuristic. Particularistic means that case studies focus on a particular situation, event, program, or phenomenon. While case studies examine specific instances, they illuminate general problems and suggest to readers what to do or not to do in a similar situation (Merriam, 1998).

Descriptive means that the end product of a case study is a rich, “thick” description of the phenomenon under study. Case studies obtain information from a wide variety of resources and illustrate the complexities of the situation. Case studies can show the influence of personalities and the passage of time on the issue (Merriam, 1998).

Finally, case studies are heuristic. Heuristic means that case studies illuminate the reader’s understanding of the phenomenon under study. A case study can “explain the
reasons for a problem, the background of a situation, what happened, and why” (Merriam, 1998, p. 31).

This empirical inquiry was a single case study with more than one unit of analysis. The rationale for a single case study was that the case represented the typical or representative case (Yin, 2003). “When typical site sampling strategy is used,” Patton (1990) states, “the site is specifically selected because it is not in any major way atypical, extreme, deviant, or intensely unusual.” The challenges that North Mesa Elementary School faced are representative of many other schools not only within Suburban School District but also across the United States. The lessons learned from this case are assumed to be informative about the experiences of many other schools. Within this single case, attention was given to several subunits: the school administration, the Instructional Leadership Team, and the grade-level professional learning communities (PLCs). Hence, this inquiry utilized an embedded case study design (Yin, 2003).

**Research Procedures**

**Participants.** The school administration was comprised of a principal and an assistant principal. The principal, originally from North Carolina, has bachelor’s and master’s degrees in early childhood education and an educational specialist certificate in educational leadership. She taught for five years at North Mesa Elementary School before serving as the assistant principal at a nearby school. She became principal of North Mesa Elementary School in July of 2007. The current assistant principal came to the district in July of 2008 in order to serve as an assistant principal at a district high school. He started as assistant principal of North Mesa Elementary School in November of 2008. A native of New
Mexico, he has a bachelor’s degree in mass communication and a master’s degree in sports administration.

The Instructional Leadership Team (ILT) was comprised of the school administration and representatives from each of the instructional areas. This included one representative from each of the grades from Kindergarten through grade five plus representatives from educational assistants and ancillary areas such as ESL, bilingual, Reading Recovery, physical education, music, and art. The educational technology specialist, instructional coach, site specialist, and math interventionist were also members. Of the fourteen members during the 2009-10 school year, thirteen members were female and one member was male. This proportion was reflective of the predominance of female staff members across the school. Four of the members had bachelor’s degrees and nine had master’s degrees. The members were heterogeneous in their ages: three were in their twenties, three were in their thirties, four were in their forties, and four were in their fifties or older. The group also represented a variety of experience levels. Two members had taught less than five years, two had taught between six and ten years, six had been educators between eleven and nineteen years, and four had more than twenty years of experience.

The instructional leadership team met for one hour twice per month. The stated purposes of the instructional leadership team were:

- To improve instruction and student achievement results schoolwide;
- To guide, enforce, and reinforce the school plan;
- To assist with the development of agendas for the grade level and cross-grade level meetings;
- To ensure the implementation of standards and the vision within and across the grade levels;
- To monitor progress and address concerns;
- To be a resource for the implementation of standards and district curriculum;
- To review data and plan for improvement;
• To disseminate content information from the district, state, and federal government;
• To troubleshoot the concerns of teams;
• To enable others to act and model the way; and
• To encourage the heart (North Mesa Elementary School Leadership Structure, Fall 2008)

All teachers were expected to participate in professional learning communities. The professional learning communities are communities of practice focused on instructional issues. In 2008 teachers were supposed to meet once per week with the focus of the meetings rotating based on a fixed schedule:

1st Wednesday of the month: Grade-level PLCs at school site / Reflective inquiry
Ancillary Teachers at district office

2nd Wednesday of the month: Vertical PLCs at school site / Subject area articulation

3rd Wednesday of the month: Staff meeting at school site
Grade-level PLCs at school site / Reflective inquiry

4th Wednesday of the month Grade-level PLC / Student Assistance Teams

The stated purposes of the grade-level teams were to “maintain unity of curriculum, instruction, [and] assessment, and to implement the standards at each grade level”.

Expectations for teachers included:

• Teachers will coach and support the implementation of the standards and the vision in each other’s classrooms.
• Teachers will study and support each other’s implementation of best practices. (North Mesa Elementary School Leadership Structure, Fall 2008)

The stated purposes of the vertical PLCs were very similar to the purposes of the Instructional Leadership Team: “to maintain unity of curriculum, instruction, and assessment in each subject area across grade levels. Specific goals of the vertical PLCs were:

• To improve instruction and student achievement results schoolwide;
• To ensure the implementation of standards within subject areas and across the grade levels;
• To advise the leadership team of progress and concerns of grade-level and cross-grade level team meetings;
• To coach and support the quality implementation of subject areas;
• To demonstrate the implementation of subject standards for teachers in each subject area;
• To review data and plan for improvement; and
• To disseminate subject information from the school, district, state, and federal government. *(North Mesa Elementary School Leadership Structure, Fall 2008)*

Some stated purposes apply equally to both grade-level and vertical PLCs:

• To review and clarify standards;
• To implement the state standards;
• To support the implementation of the standards in every classroom;
• To share best practices;
• To share examples;
• To review data (student learning, questionnaire, demographics, school process, student learning style preferences);
• To review student work;
• To develop rubrics for student work related to the standards;
• To implement the curriculum map;
• To help teachers format classroom tests to resemble the state assessment;
• To develop standards assessments and benchmarks. *(North Mesa Elementary School Leadership Structure, Fall 2008)*

During the 2007-2008 school year, the typical size of a grade-level PLC was approximately thirteen to fourteen members due to the large size of the student body. With the opening of a new school within the district and the redistricting of North Mesa Elementary School’s attendance zone, the typical size of a grade level PLC was reduced to approximately eight members during the time boundaries of this study.

**Data collection methods.** Consistent with case study methodology, data was collected through a variety of methodologies including interviews, observations, and document reviews. Over the two years of this case study, I interviewed eleven staff members of North Mesa Elementary School and one instructional support provider from Suburban School District who had worked directly with one of the PLC teams during the course of this
study. Four of the interviewees were classroom teachers who played leadership roles in their assigned grade levels. Due to reassignment during the second year of the study, some of these teachers were able to speak to the processes used by PLC grade level teams at the first, second, third, fourth, and fifth grade levels. Four additional interviewees were ancillary teachers who participated in PLC meetings at four of these grade levels. Two of the interviewees were school-based instructional support providers who assisted all of the PLC teams within the school. Four of the interviewees were interviewed twice, once at the beginning of the study and once at the end of the study. Finally, the principal was interviewed for approximately one hour on three separate occasions over the course of the study.

In addition to the interviews, I observed a number of collaborative meetings. The observations included one staff meeting, fourteen Instructional Leadership Team meetings and several PLC team meetings. Two of the Instructional Leadership Team meetings were day-long retreats. Furthermore, I participated in three interventions with some staff members of the school: a five-day data portfolio conference, a three-day data dialogue conference, and a one-day Instructional Leadership Team meeting with an independent consultant. Supplementing the observations, I obtained agendas or notes from four additional Instructional Leadership Team meetings, PLC team meeting agendas, notes, and artifacts from four different teams, and Plan-Do-Study-Act (PDSA) documents from every grade level.

In addition to artifacts from PLC team meetings, several dozen additional documents were also reviewed. These documents included the school’s Education Plan for Student Success (EPSS), the principal’s Professional Development Plan (PDP), the NM CLASS
school improvement plan, Quality of Education survey results, achievement data, demographic data, forms designed to facilitate collaborative discussions such as an Essential Learnings form, and informational documents describing team purposes and structures as well as calendars.

All sixteen interviews, twelve of the meeting observations, and two of the intervention observations were digitally audiorecorded. I transcribed all of the recordings and each recording was played back several times to ensure the accuracy of the transcription. Proper steps were taken to ensure the confidentiality of data as described in the consent form located in Appendix A.

**Trustworthiness and consistency of the data.** Trustworthiness was developed through both construct and external validity. Three strategies were taken in order to ensure construct validity (Yin, 2003). First, I used multiple sources of evidence including observations, interviews, and document reviews in order to develop converging lines of inquiry. The data were triangulated so that multiple measures of the same phenomenon were provided. Second, I maintained the chain of evidence by citing relevant portions of the case study database, recording the circumstances in which the evidence was collected in the database, and following procedures as specified in the proposal. Finally, a key informant reviewed the draft case study report.

External validity was developed through analytic generalization, not statistical generalization. In other words, I generalized results to the theories of organizational cultural change and evaluation capacity building in the context of school systems. Furthermore, similar findings were derived by two other researchers who independently reviewed the data. The researchers were graduates of the University of New Mexico’s Ph.D. in Organizational
Learning and Instructional Technologies program who also had experience teaching in the public education system. One of the researchers currently serves as an independent education consultant; the other is retired.

Consistency was developed through adherence to a case study protocol and through the development of a case study database.

**Data analysis.** An inductive, iterative process of reading and rereading the transcription was used to produce subcategories for information and analysis. Statements were partitioned into units, grouped in common category headings, analyzed, and summarized. The plausibility of subcategories was established by testing them with new information until all relevant information had been assigned to a category (Hancock & Algozzine, 2006).

This report follows a linear analytic format (Yin, 2003). The report commenced with a description of the problem and proceeded with a review of the relevant prior literature. The report then covers the methods used, the findings from the data collected, and the implications of the findings.

In order to present a rich, thick description which illustrates the influence of personalities, the passage of time on the issue, and the complexities of the case, each of the subunits of this embedded case is profiled in a holistic manner in Appendix B. The next chapter, however, organizes the data in a linear analytic manner in order to answer the research questions.
Chapter IV: Results

Research shows that long-term student achievement trends have not been impacted by a number of top-down educational reform policies. In order to accelerate growth in student achievement, many researchers are now advocating the development of cultures of reflective inquiry centering on collaborative problem solving and the use of data. This study was a case study of a typical school, North Mesa Elementary School\(^4\), as the school staff members endeavored to develop a culture of evaluative inquiry. The actions and beliefs of the school administration, the Instructional Leadership Team, and several professional learning community teams (PLCs) were highlighted in this case. The professional learning community teams profiled in this case were the first, third, fourth, fifth, and bilingual teams.

The purpose of this chapter is to summarize the results of the observations, interviews, and document reviews conducted to answer the central question of this study: “How do staff members at North Mesa Elementary School develop a culture of evaluative inquiry?” While a holistic profile of the case is presented in Appendix B, this chapter is organized by the specific subquestions that underpin the central question of this study.

1. How do members of the school leadership team increase their knowledge and skills related to evaluation and data analysis?

There was no systematic method for all school leadership team members to increase their knowledge and skills related to evaluation and data analysis. Rather, a few members of the leadership team attended professional development opportunities as they arose. These trainings included national conferences sponsored by Education for the Future, Solution Tree, MiraVia LLC, and the University of Arkansas. The focus of these trainings was on

\(^4\) Pseudonyms have been used for the name of the district, the name of the school, and the name of participants in order to protect the anonymity of the participants.
creating and analyzing data portfolios, utilizing data within professional learning communities, creating productive dialogues around data, and providing intervention to struggling students. Each of the trainings provided relevant material which participants cited as useful for increasing their knowledge and skills related to evaluation and data analysis. Several participants also mentioned that some trainings, like the data portfolio conference sponsored by Education for the Future and the Data-Driven Dialogue conference sponsored by MiraVia LLC, not only increased knowledge and skills, but also changed their attitudes about the nature of data.

While training participants frequently did share a brief synopsis of the training with other Instructional Leadership Team members, the synopsis was not as effective in developing the knowledge and skills of other team members as if they had been able to participate in the complete training themselves. Thus, team members participated in different trainings from each other and did not develop shared schema. Furthermore, not all team members participated in outside training and some team members who did participate like Leslie, an ancillary teacher, and Wendy, a primary grades teacher, were no longer members of the Instructional Leadership Team by the second year of the case study.

The entire Instructional Leadership Team did participate in one professional development opportunity. In February of 2009, an independent consultant associated with Education for the Future named Patricia facilitated a day-long consultation with the team. While most of the day was spent eliciting current practices from team members, it did help to set a shared vision of what best practices might look like. While the action plan from that

---

5 Pseudonyms have been used for all participants in order to preserve the anonymity of participants.
day was discussed at a leadership retreat later that semester, the action plan did not become a
guiding document for further action.

Throughout the course of the case study, there were discussions about having
leadership team members participate in a book study so that they could develop a shared
vision and common vocabulary. The book study did not occur during the two years of this
case study. A lack of time was cited as the reason for not following through with the idea of
a book study. Therefore, any increase in skills and knowledge related to evaluation and data
analysis which occurred through reading professional literature was a result of individual
initiative and was not consistent across all team members.

2. How do members of the school leadership team apply their evaluative inquiry knowledge
and skills to emergent issues within their school community?

The mandated Plan-Do-Study-Act (PDSA) process was the primary structure through
which the leadership team members applied their evaluative inquiry knowledge and skills to
emergent issues within their grade levels. The members of the leadership team led their
grade-level colleagues in the analysis of data to identify emergent issues with the students at
their grade level. The requirement established in the second year of the study to create
weekly grade-level wide interventions to help students progress with the learning targets in
the identified area prompted teams to take the PDSA process to a deeper level than they had
previously. As Diane, the third grade chair had noted,

It’s less compliance and more using it [evaluative inquiry] as a formative tool. It’s
more of, ‘What did you find out and how are you going to change what you are doing
based on what you found out?’ I think it’s coming naturally…because we have to do
stuff every single week. If you have to do stuff every single week then you have to
know where the children are and you have to have good solid grounds to stand on
versus just ‘I feel like’ or ‘I noticed that’. (2-22-10 p. 18)
Teams were forced to create common formative assessments to measure student progress towards the PDSA goals. The process of creating common formative assessments and then analyzing the resulting data impelled some of the teams to further define the learning targets and discuss how best to measure the learning targets in addition providing effective instruction in this area. The experience of the first grade team as described by Liz, the former chairperson, and Wendy, a teacher, illustrates this process. In 2009 during the first year of the study, Liz stated,

It was quite interesting to hear the different opinions about what people thought number sense was. As a group we didn’t all necessarily have the same view about what number sense is. (2-24-09b p. 22)

In 2010, Wendy shared the experience of the team after they were forced to create common formative assessments to measure progress on PDSA goals.

Our PDSA goal for math really revolved around the importance of number sense [and] the importance of vocabulary. It was hard because people wanted something fast and easy. When we really sat down, [we discovered] this test does not accomplish that…When we really developed a test that spoke specifically to the vocabulary, then people began to say, ‘I didn’t realize that my kids didn’t know this’…People were saying, ‘You know, I really saw a difference with this’…I think in some ways for our grade level it pulled people together more…People had to start using the same vocabulary. People had to start saying ‘What do you mean by that?’ Um, ‘I’m not sure how to accomplish that.’ ‘What is it that we are specifically looking at?’ So I think it had to get more detail-oriented. (5-17-10 p. 5)

Teachers at several other grade levels also indicated similar experiences within their teams.

Overall, however, the results of the PDSA and common intervention processes indicate that members of the school leadership team who served as Professional Learning Community (PLC) team facilitators had different levels of evaluative inquiry knowledge and skills and also had various levels of knowledge application. Led by Diane, the third grade team already had experience with evaluative inquiry and was knowledgeable about how to
use data to provide effective interventions for the students. In the second year of the study, Anne took over the facilitation of the historically contentious first grade team. While interested and capable, it took a while for Anne to develop effective facilitation skills and for the team to become productive with the inquiry process. At the fourth and fifth grade levels, knowledgeable ancillary teachers served as data coaches in order to provide support to the grade level chairs who had not yet developed evaluative inquiry knowledge and skills. At each of these three grade levels, the leadership team members were unsuccessful within the time boundaries of this study in developing a consistent vision of a culture of evaluative inquiry across all of members of their team. Indeed, some members still questioned the need to acquire the skills.

Under the guidance of Leslie, newly formed bilingual PLC also applied evaluative inquiry knowledge and skills to the emergent problem of declining test scores of ELL students. The team examined both schoolwide and classroom level data and utilized continuous improvement tools such as root cause analyses to develop action plans in five areas: program structure and staffing, student placement, curriculum, parent and community involvement, and communication. This study ended before it could be determined how effectively they could carry out the tasks in the newly developed plans.

The Instructional Leadership Team (ILT) members rarely applied their evaluative inquiry knowledge and skills collectively towards schoolwide emergent issues. There was some review of data in ILT meetings such as SBA data at the retreat or the intervention wall at a regular ILT meeting. However, the schoolwide goals and action plans were not collaboratively generated by the leadership team although the team members did have the opportunity to give feedback to already developed plans. The principal, for the most part,
created the school’s Educational Plan for Student Success and completed the state required
NM CLASS school improvement data review and priority setting process singlehandedly.
ILT members had the opportunity to provide feedback after the reviews had been created but
were not an integral part of the process. Toward the end of the case study, the ILT did take
more of a leadership role in the evaluation of the effectiveness of the common interventions
and the revision of the plan for the following school year.
3. How do members of the school leadership team consciously build the evaluation capacity
of teachers within their professional learning communities?

- What training activities are held?

At the beginning of the study in 2008, there were efforts to develop shared norms and
to collaboratively assess the perception of where the school fit along a number of school
improvement dimensions. However, challenges were encountered with both processes and
the efforts were truncated. The norms suggested by different grade levels were very diverse
and there was neither a process nor the time to continue with the norm setting process so the
process was simply discontinued. The initial self-assessment against the continuum
dimensions was unsuccessful due to logistical difficulties. The self-assessment was repeated
at the end of the first year of the study in the form of a survey, but the survey was not
repeated at the end of the second year to note any changes in perceptions of the schools’
evaluation capacity.

Few other training activities were held with all staff members at the same time. In
2009, the school did ask me as the Director of Assessment and Data Analysis to briefly
profile the progress of the district and school overall on the state accountability assessment
and Northwest Evaluation Association’s Measures of Academic Progress during the
orientation day. Then grade level teams were asked to review appropriate grade level data which was provided to them and to state the facts that they noticed in the data, to speculate on the meaning of that data, and to brainstorm next steps. Some teachers stated that it was the first time that they had reviewed a school-level and grade-level profile.

At the beginning-of-the-year orientation in 2009, Sharon, the principal, did mention to all staff members the new requirement to provide weekly grade-level wide interventions to students. However, the plan was only discussed in generalities and many ILT members commented that they felt that the need for the radical shift in instructional practices had not been adequately communicated to all staff members. No further whole staff meetings were held to discuss details and answer questions. ILT members felt that staff members received different messages since communication about details and the need for intervention had to be relayed through the grade level chairs who had different backgrounds and levels of understanding themselves.

Other trainings that were held for members of the professional learning communities included trainings to teach members how to use the Student Assessment Portal in the fall of 2008 and the Data Driven Classroom data warehouse and assessment management system in the spring of 2010. In 2008, Sharon and Wanda trained teachers to enter Developmental Reading Assessment and Math Quarterly Assessment data into the new Student Assessment Portal but did not show the teachers how they could access and use the reports to help with them analyze data. In 2010, in my role as the Director of Assessment and Data Analysis, I trained each grade level how to access data warehouse reports and create answer sheets to scan into the Data Driven Classroom assessment management system. This system enables teachers to analyze student progress by content standard. While staff members were trained
during this study, the actual use of the system did not take place until after this study had been completed.

Several PLC teams consciously incorporated activities into some of their meetings which were designed to increase the evaluation capacity of team members. The first grade team, for example, used me as an outside facilitator to model a process for reviewing data in a non-threatening manner. The fifth grade team also had a district instructional coach who helped them organize and analyze their NWEA data. The bilingual team asked an outside facilitator to manage a root cause analysis with them. Several teams, including the first and fourth grade teams, consciously included conversation protocols into some activities to help train members how to discuss data in a non-threatening, productive manner. However, these activities were team-specific and the same training was not provided to all teams.

- *What roles do leadership team members play?*

The role of the leadership team members has evolved over time. Under previous school administrations, the Instructional Leadership Team members would take notes about administrative matters at the ILT meeting and return to their grade levels to disseminate their notes. If decisions needed to be made at the ILT level, the issue was discussed both at the PLC team level and at the leadership team level. This process sometimes took months to complete. Sometimes communication from the grade levels to the administration was perceived even by team members “to be a litany of complaints” (10-23-08, p. 6). However, this process did make teachers feel like there was an avenue through which they could express their opinions.

Under the leadership of Sharon, the school principal, the Instructional Leadership Team members are still intermediaries between the administration and the staff as a whole.
However, the role has shifted from grade level representatives who disseminate information about “business” items to facilitators of cultural change. While the ILT members still had the opportunity to bring up issues related to their grade level in the ILT meetings, Sharon also expected the grade-level chairs to lead their teams through the process of evaluative inquiry. Furthermore, she expected them to lead their colleagues through a cultural shift whereby they would collaboratively plan grade-level wide interventions. This was a significant shift in the role of leadership team members.

There was no explicit schoolwide communication with all staff members about the changing role of leadership team members. This led to some confusion among some staff members about the role of the Instructional Leadership Team. Some staff members expected ILT members to continue behaving in the same manner as before, rather than focusing on data and interventions.

The confusion was further compounded by Sharon’s belief that grade level team representatives could make decisions on behalf of the team. Sharon stated,

Isn’t the purpose of [the grade level representatives on the ILT] to make those decisions for the grade level? (2-11-10 p. 14)

However, not all staff members concurred as one grade level representative said,

My grade level [is] a little concerned that decisions are coming out of ILT, are just being made and it comes across as a mandate and less like decision-making. (2-11-10, p. 13)

Contributing to the frustration of staff members was the fact that Sharon chose to appoint grade-level representatives in School Year 09-10 rather than have grade level PLC team members elect their representative. The ILT members were placed in the difficult role of trying to develop buy-in and acceptance of a cultural change when the need had not been
clearly established to all staff members, a shared vision had not been developed, and the legitimacy of the grade level chair as a change agent may not have been established.

There also was no systematic leadership training provided to the Instructional Leadership Team members. While some of the more long-standing team members had attended a two-hour district-sponsored workshop in 2007 or 2008 about how to facilitate PLC meetings in general, many of the newer grade level representatives had received no formal training. Thus, the grade level chairs had been put into new roles without consistent professional development to help them learn how to facilitate teams under the new paradigm of evaluative inquiry.

- *What learning processes are included in the evaluation capacity building activities?*

While many of the staff members did not receive formal training related to the development of a culture of evaluative inquiry, engaging in the process did lead to significant learning. Each of the grade levels had at least one person who took a leadership role in the analysis of data, even if that person was not the grade level chair. For example, Natasha, an ancillary teacher attached to the fifth grade PLC team, looked up examples of items in the area of measurement which led team members to better define the learning targets. She also helped the team to create a pretest for students and shared the results of an item analysis with team members which helped them to differentiate instruction and also examine the quality of the test questions which they had developed. Through this process, the fifth grade team came to realize that there was a mismatch between their instruction and the assessments that they were using.

Even the fourth grade team, which struggled the most to shift towards a culture of evaluative inquiry, had team members who had a significant learning experience and applied
that knowledge. One of the normally reluctant team members had participated in a district-wide team which revised the Math Quarterly Assessments. As part of that process, the district team analyzed a report which showed how many points the students at that grade level had earned on each benchmark in math on the state accountability assessment. The team also analyzed the number of points associated with each benchmark. This team member brought this knowledge back to her PLC and used that information to suggest focus areas for the team.

Thus, engaging in the process led to a significant increase in knowledge even if formal training was not provided.

- **How does the leadership team provide support in terms of communication, systems and structures, and organizational culture?**

  The leadership team did provide support in terms of systems and structures, communication, and organizational culture. The leadership team tried to provide tools that would help PLC teams be more productive and stay on task. This included the introduction of a standardized form for recording notes from PLC meetings and having PLCs meet in a common area so that resource personnel would be nearby if questions arose. Many teams, however, felt that these changes resulted in meetings which were too structured. Some team members felt that this structure was inconsistent with what their team was actually doing and Sharon, the principal, agreed to give teams more flexibility to meet where they would like and to structure their notes as they would like.

  The leadership team also supported the initiative of particular PLCs to change the way that they operate. For example, the school administration and bilingual coordinator used federal bilingual education funds to provide substitutes for the newly formed bilingual PLC
to meet for a half day once a month during the spring of 2010. The ability to meet regularly for an extended period of time allowed the team to discuss philosophy, develop a shared vision for the bilingual program, and create a plan to restructure the program staffing, curriculum, and instructional materials.

The leadership team also supported teams by inviting outside facilitators to come in for PLC teams that were struggling with interpersonal dynamics or struggling with particular groups of students. For example, district staff members were invited to facilitate several meetings at the first and fifth grade levels to help the teams set a direction and establish a process for examining data and discussing instruction. Furthermore, the leadership team supported the introduction of conversation protocols to several teams, specifically first and fourth, which were struggling with having safe, productive conversations.

Several structures set up by the district also supported the development of a culture of evaluative inquiry. The district has curriculum frameworks and has identified “Power Standards”, those enduring ideas to which teachers should pay the most attention. Furthermore, the district retooled its district-wide Math Quarterly Assessment in School Year 09-10 so that teachers could measure the growth of students on the Power Standards across the school year. The use of the Developmental Reading Assessment and the NWEA Measures of Academic Progress meant that there were standard growth measures for reading and math at each grade level. Furthermore, the district was in the process of introducing a data warehouse and assessment management system which would enable teams to more quickly and easily access students’ data and track the progress of students by standard.

The leadership team also structured time for PLC teams to provide feedback to the school administrators. Part of the agenda during each ILT meeting was time for each of the
representatives to bring up issues of importance to that team. However, there were not clear and consistent structures for schoolwide communication. In order to preserve time for PLC meetings, staff meetings were minimized. A few were held throughout the year, primarily for mandated topics such as test security training and for training in new initiatives such as how to use the new electronic gradebook. Other than orientation days at the beginning of the year, no time was set aside to discuss the direction of the school with all staff members present.

The desire to protect the time of the teachers also had other unintended consequences. In order to minimize the time that Instructional Leadership Team members would miss from their classroom instruction or from their time to prepare for instruction afterschool, Sharon did most of the schoolwide planning (e.g., the Educational Plan for Student Success and the NM CLASS school improvement plan). She then informed the ILT and asked for feedback rather than facilitating a truly collaborative process with input. While this minimized the time that the leadership team members needed to spend away from their classrooms, it may have contributed to some staff members feeling like they did not have a voice and that plans were mandates. Furthermore, since the need for change had not been clearly established with all staff members, it put the instructional leadership team members in the difficult position of trying to move initiatives forward which they themselves did not fully understand while also conveying the questions and concerns of their teammates to the school administrators.

The lack of schoolwide activities, the siloed approach toward communication, and the increased structure to PLC meetings may have contributed to a growing sense of social and professional isolation. One teacher had noted,
[We] don’t feel as close to people as in other years…Even when we were a big campus, [we] at least communicated a little bit more than we seem to now. We rarely see each other as a group whether it is for social [or] creating vision and policy or anything…There’s not that camaraderie. (5-19-10 pp. 20-21)

The administrators had also noted that the school atmosphere had changed. Sharon, the principal, stated,

Something that has bothered me greatly is the lack of social connections that we have had here and I’m not sure why…I know things have changed…I think if you go back and look, the expectations have ramped up. That is systematic and also deliberate. The school under [a former principal was a] whole different school than it is under me. The expectations have changed. In some ways I think it is good because the needs [of children] are clearer. When you are really looking at stuff and you are examining data, the paperwork is tremendous and exhausting. I can’t be that happy person anymore that says, ‘It’s okay. Don’t worry about it’…I would like to do it sometimes. I look at their faces and I know it’s hard. It’s hard on me, but we’re doing what’s best for children. (6-15-10 p. 14)

While increased focus and productivity is hopefully beneficial for the students, the relational trust developed through social interaction and camaraderie can be used as social capital in the change process.

While a decreased sense of camaraderie worried many of the school leaders, the organizational culture was still marked by a remarkable sense of honesty. Members of the instructional leadership team felt safe enough to share that some of their colleagues planned to sabotage the intervention process. Furthermore, they felt safe enough to directly question the principal, Sharon, about her leadership style in an Instructional Leadership Team meeting. One teacher asked,

When we get our grade level’s feedback, is that going to be taken into consideration? Because I think that’s where it comes from with the grade levels with the whole ‘We’re not included. Our voice doesn’t matter.’ Because it’s like if you have an idea in your head, that’s going to be it whether we give feedback or not. (3-3-10 p. 5)
This honesty indicates that there is still a significant level of trust within the leadership team members.

- *How does the leadership team address the information needs of differing MBTI personality types (function theory)?*

  While staff members did not have the opportunity to take an MBTI personality types inventory assessment, issues related to function theory do surface throughout the course of this study.

  The principal’s actions and statements reveal that she has a preference for “Feeling”. Her natural tendency is to first consider the impact on individual people and relationships. For example, William, the Assistant Principal, stated that he and Sharon had debated the impact of the brand of donuts to be served to staff members,

  [Sharon] and I were talking about Krispy Kreme versus Wal-Mart donuts. I’m a guy. You bring donuts, I’m happy. But apparently Wal-Mart donuts are not quite good enough. You know, Krispy Kreme says we love you and Wal-Mart says we kind of like you. (2-11-10 p. 23)

  Sharon is aware of this preference and has consciously tried to move out of her comfort zone if she believes that it is in the best interest of the students. She stated,

  The important thing is what is best for the children here at this school. If I can live with that at night, even if I have hurt somebody’s feelings, even if things have not gone the way somebody wanted them to go, then I am okay…In the past I would have thought, ‘It’s best for children but I might hurt somebody’s feelings’. So [with] the things I’ve done this year, I kind of look at myself and go, ‘Holy Cow’! (6-15-10 p. 9)

  The differences in preferences amongst staff members may have contributed to issues with communication. Sharon noted that communication is an ongoing issue. In 2009 she stated,
Every year I have seen on evaluations on me [that] communication could be better and I think unless I stand on my head and run around naked in the courtyard, I can’t do any more. (5-28-09 p. 14)

In 2010, she also stated something similar,

I feel secure [that] I have tried every single way I can to communicate and validate what people need to hear. (6-15-10 p. 13)

While Sharon has consciously tried to communicate in a variety of manners in order to reach all staff members, it is evident from staff members’ comments that not all of them perceived that they had the information that they needed, be it hard data, symbols, stories, or patterns, in order to willingly want to change and problem solve. Different preferences may have also contributed to the significant conflict and miscommunication that was evident in a number of teams, especially the first and fourth grade teams.

- **What challenges are encountered in the evaluation capacity building process?**

Numerous challenges were encountered in the evaluation capacity building process. These challenges include changes in the external economic and political backdrops, deeply held cultural beliefs inconsistent with developing a culture of evaluative inquiry, a lack of a shared vision, process norms, and assessment literacy, misalignment within the system, and a lack of resources.

Unfortunately, a global economic crisis started around the same time as the beginning of this study. For the first time in its history, Suburbia School District needed to cut a significant amount from its annual budget. An atmosphere of fear and uncertainty permeated the school climate as teachers feared that their positions might be eliminated or that their compensation would be reduced. Barbara, an instructional support provider who works with
all of the teachers in the school, explained how the economic crisis affected the change initiative,

I think that the economy is pulling people down. I think that if everything were rosy we could push them a little harder. People are so low right now because they are thinking, ‘Am I going to lose my job? Am I going to lose my job?...Are we running out of toner in the printers? You know, all that stuff even if they are not voicing it out loud I know that it is at the back of their minds. Because some people that normally would be pleasant, aren’t. And then the people that you know are not going to be pleasant are worse. (11-24-08 p. 5)

Although no personnel were ultimately laid off or furloughed during the time of this study, extremely tight financial resources became a new reality. Furthermore, the ongoing cuts kept staff members from feeling centered and secure. The ongoing economic crisis also diverted staff members’ time and attention away from school reform as the students’ behavioral and social needs also increased significantly.

While the school has always lacked adequate resources, the exacerbated lack of resources had an impact on the initiative to develop a culture of evaluative inquiry. As a Title I school, the school was fortunate to have some federal funds available for some professional development and for substitutes to allow certain teams like the bilingual team additional time to meet to create a shared vision, analyze data, and design an action plan. However, the school could no longer afford to play “fairy godmother” to staff members and provide them with whatever materials they requested. Sharon, the principal, described this as “the popcorn effect”, the allocation of resources in a manner that is not aligned to a common vision and aligned plan. Furthermore, there were no state funds available to support the initiative and teachers had less support in other areas of their work. Teachers felt overwhelmed by all of their duties as a whole and felt as if they did not have control over
their work life. Wendy, a primary grades teacher, described just a portion of the stress that teachers were feeling even at the beginning of the study in 2008,

They [the school administrators] are very concerned about the reading scores this year. On our last inservice day…they very specifically focused on how guided reading looks at the different grade levels. [Some questions we have are] what about the resources to do this? [When do we have] the time to plan these lessons, get into the leveled library, pull all these books, and do the lessons? What they are doing with the SMART goals takes a lot of time already and then you’re adding what they see as this additional [requirement]. It’s always that issue of when are they going to find the time? This year even more is coming, is being put on them. Every year it’s more and more and more and nothing is ever taken off. (10-14-08 p. 3)

Given the already high stress levels of the teachers, any small changes in expectations for teachers provoked very strong reactions. It is into this setting that Sharon endeavored to create transformational shifts in the manner in which teachers collaborated and used data.

In addition to the changes in the economic context, there were also changes in the political landscape. While fewer federal and state politicians touted the school accountability measures from the ‘No Child Left Behind’ Act as the mechanism which would produce significant increases in student achievement, more and more federal and state politicians began to advocate tying teachers’ evaluations and compensation to increases in student achievement. Many teachers feel that the politicians are looking for simplistic solutions to complex problems and that the methods that might be used to measure “their effectiveness” may be inherently unfair and not represent the true value of the instructional and emotional support that they provide to their students. Thus, many teachers may have associated a move towards developing a culture of evaluative inquiry as a step towards proposed accountability measures which could affect the teachers’ livelihoods.

Related to this concern is the belief among many teachers that “data” represents only a small fraction of the educational experience. Several instructional leadership team
members including Leslie, Cheryl, and Wendy expressed relief after attending training that data actually included perceptions, processes, demographics, and qualitative student achievement data as well as student achievement data that could be validly represented by a number. They had, like many teachers, viewed an emphasis on data as an actual narrowing of the curriculum and a lack of well-rounded support for students. Thus, some teachers believe that data-driven instruction could actually harm the students in the long-run.

Traditionally, teachers at the school have enjoyed an extremely high level of autonomy. Sharon, the principal, stated,

I remember coming here in 1998 and it was like a free-for-all here. Honest to goodness, you did whatever you wanted to do pretty much. No one had the same kind of materials. No one had the same standards. (6-15-10 p. 15)

While the district had emphasized greater alignment and had raised expectations since that time, teachers still operated, by and large, as independent operators. Wendy, a primary grades teacher, described the reaction of some teachers when the district set the expectation that all teachers would collaborate together in professional learning communities,

This group has been together for a long time and there was this really big change in the district of push, push, push…You will get [the students] to this [achievement level]…So when this began years ago to [raise] the bar and the expectations, there was this massive push, on primary to get them from A to 24. ‘Get your butt together and get it going’ and so it was this massive, ‘Oh my god, what am I doing?’ And there was this frenzy. And I think there’s still some of that in there and now suddenly it is like, ‘Oh, will you guys please work together now’? (2-29-09b p. 17)

Collaboration is perceived by many to be an infringement of academic freedom and a time drain. Rather than viewing collaboration as an opportunity to socially construct new knowledge, it is often viewed either fearfully or as a waste of time. Furthermore, some teachers interpret the new expectation that all teachers engage in collaborative inquiry as evidence that others don’t believe that teachers know what’s best for their students.

104
Part of the hesitancy to collaborate also centers around the fear of deprivatization of data. Liz, the former first grade team representative, explained,

Teaching is very personal (2-24-09b p. 13)…Instead of looking at the PLC as a group that’s supposed to help everybody, they don’t see it that way. They are afraid to talk about what they do or they are afraid to ask for help because they think that that reflects badly on them…They don’t ask ‘What are you doing? Your scores look really good’. They don’t want to know. They just say, ‘Well, this is what I am going to do’. (2-24-09b pp. 11-12)

She added,

Unfortunately there are a lot of the teachers who feel like when they put forth their data that it looks punitive to them. So when a lot of times if somebody is not going to be at a meeting, they don’t want to leave their data without being there because they are afraid that data is going to make them look bad. (2-24-09b p. 10)

A focus on student learning is actually a transformational shift for many teachers. In the traditional paradigm, the focus has been on how well a teacher taught a particular standard. If a student failed to master the standard, it was assumed that student-related factors were to blame. Under a student-learning paradigm, teachers are held to a higher standard as they have to differentiate the instruction to meet the differing needs of various students in order to ensure that all students master the standard. The focus on learning also is daunting for teachers in that they do not control all of the factors in the learning process, yet are held accountable for the outcome.

One obstacle to creating a culture of evaluative inquiry is that many teachers don’t see the need to make this type of transformational change. The teachers frequently have the perception that the quality of instruction has been high and that the students have been learning to the best of their innate ability. Furthermore, the achievement scores of students in the district have traditionally been significantly higher than the average scores of other students from around the state. Wendy, a primary grades teacher, elaborates,
The entire framework of the school was not messed up. It was this one group we were working with [students with disabilities]. It feels like we’re going in with hatchets and chopping everything out instead of saying, ‘Let’s look at this area and see where we can go ahead with that’. I don’t understand that. (10-23-08 p. 11)

The perception of some teachers was that everyone was being punished because two subgroups did not achieve AYP targets instead of believing that creating a culture of collaborative evaluative inquiry was a good thing for all of the students in the school. The lack of time to meet as an entire staff and to establish the need for change hindered the process of creating a new shared vision. Without establishing the need for change, teachers who are already overloaded are unlikely to willingly change long-established routines and behavior patterns.

One of the biggest changes that the intervention process required of the teachers was to truly shift to standards-based teaching instead of following the sequence within the textbook. The intervention process demanded that the teachers focus on PDSA goals until the students mastered those goals. This initially caused some confusion and discomfort with some teachers and demonstrated how incomplete the implementation of the paradigm shift to standards-based teaching, which started over a decade ago, has been.

Another obstacle to creating a culture of evaluative inquiry was a pattern of short attention spans when it came to previous change initiatives. Like other educators across the country, the teachers at North Mesa Elementary School had seen many education reform efforts come and go. Many teachers had adopted a “This too shall pass” philosophy about externally introduced initiatives. Even within the school the administrators sometimes had not followed up to insure that previous initiatives were implemented with fidelity. For example, while the requirement to complete a PDSA process has been around for a number
of years, the process had been more one of filing reports rather than an actual evaluative inquiry process. Another example is that teachers were supposed to keep Running Records to track student progress in reading and to complete reading continuums. However, the difficulty which teams have had in this study to create and measure reading skills indicated that many teachers are not as experienced with this expectation as had been thought.

The pattern of short attention spans with change initiatives also led teachers to demand immediate results. Rather than expecting an implementation dip as they engaged in new processes, some of the teachers expressed concern that while teachers may have benefitted from the collaborative planning, the benefits were not yet transferring to the students. For example, Wendy stated,

Did it help the kids? No. Did it help people begin to speak a little bit more? Sure. (5-17-10 p. 5)

Expecting but not realizing immediate results for students may cause teachers to not persevere through the process long enough to see the desired results. As Diane, the third grade chair, pointed out to other teams, the third grade team had had to start the process and make adjustments as they went along but they ultimately felt that the students benefitted. Many of the teachers, however, were uncomfortable with the transition process as new systems and behaviors were developed.

Another challenge to creating a culture of evaluative inquiry was the pervasive lack of shared vision and lack of shared definitions of proficiency and acceptable growth rates. For example, the district instructional coach found that some teachers were satisfied with struggling students making the average growth of students in a national norm reference group. These teachers felt that there was no need to alter instructional practices since typical growth was achieved. However, some of other staff members felt that there was a need to
change instructional practices in order to accelerate the growth of these struggling students and help them reach proficiency much faster.

The lack of shared vision also extended to differing views regarding process norms. At the beginning of the study, the leadership team tried to work with all of the staff members to collaboratively create process norms. However, the norms proposed by various teams were so significantly different, the process was not completed.

Interpersonal conflicts within teams were a significant challenge at many grade levels. Two grade level teams, the first and fourth grade teams, had to work through the interpersonal conflicts before they could become productive. Anne, the new first grade chair, described the dynamics at the beginning of School Year 09-10.

> It seemed like anything presented from anyone would be shot down. It seemed to get really personal, like personal attacks…People were just really scared to talk and people from the grade level would come to me afterward and tell me how they felt. (3-22-10 p. 14)

The dynamics within the first grade team shifted due to a combination of factors. First, one of the more domineering members was absent for a while and new patterns were able to be established. Second, the new grade level chair learned to become more assertive in facilitating the team discussions. Finally, the team learned how to utilize some tools such as conversation protocols and depersonalized presentations of data.

The fourth grade team also had interpersonal issues which thwarted productivity for a while. Marion, the grade level chair, explained her view as to how the situation was resolved,

> They [were] just not a cooperative group until we had a big blowout and I think that is what helped all of us change and develop into better teachers. We finally got to the point where instead of whining and complaining…we try to find solutions to concerns and issues… (3-11-10 pp. 2-3)
In addition to the blowout, however, the team also utilized conversation protocols and depersonalized data in order to increase the psychological safety and change the team dynamics.

Sometimes it was not interpersonal conflicts that challenged teams, but rather the different backgrounds and philosophies of the participants. The bilingual team, for example, had members who decided that they could not teach in a maintenance Spanish bilingual program due to their own personal backgrounds. Philosophically, they felt that maintenance programs would hurt the students in the long run. In this case, the team chair and the school administration discussed what role these teachers could take that would support the program without putting them in a role that would be incongruent with their personal beliefs.

Another challenge to the collaboration process was the ownership that some teachers felt over their students. One member of the Instructional Leadership Team had stated,

I know there are a lot of teachers that are hesitant to send their kids to these groups because they think that they are not going to be teaching them the way that they would teach them. (9-16-09 p. 8)

The reluctance to share their students with their colleagues through the grade-level wide interventions appears to indicate that there was some level of distrust that other teachers would be able to teach their students as effectively as they do. Since the team members were collaboratively planning for the interventions and were working towards the same PDSA goal, the concern should not have been that students would not be striving to reach the same targets.

One issue that was pervasive throughout the study was the lack of knowledge regarding assessment literacy. While staff members indicated that various assessments took
a significant portion of class time, they were unable to explain to Patricia, the independent consultant, what the purpose of the various assessments was and how the assessments related to each other. Many of the teachers viewed district-mandated assessments like the NWEA Measures of Academic Progress and the Math Quarterly Assessments and reporting tools like the Reading and Writing continuums as burdensome impositions without utilizing that information to inform instruction as had been intended by the district. For example, Eleanor, an ancillary teacher attached to the fourth grade PLC team, stated,

A lot of people will...take their kids to the NWEA test and then not do anything with the data because they don’t know what to do with the data. (3-2-10, p. 18)

Unfortunately, the lack of assessment literacy meant that the teachers did not understand how the assessment data could be integrated into the instructional process. They viewed the assessments and their instruction as totally separate activities. Eleanor explained her colleagues’ reaction when she showed them how the NWEA scores could be utilized to create flexible groups for targeted instruction,

So at that meeting I was able to show them DesCartes [a tool within NWEA that links scores to specific skills] and it was like ‘Whoa!’ They pulled specific things from DesCartes for Intervention Wednesdays…It just made the job so much easier and that is part of the efficiency piece. (3-2-10, p. 18)

Furthermore, the difficulty that some teams have had creating reading goals and measuring student progress towards these goals was evidence that teachers have not integrated Running Records and the Reading and Writing continuums into their routine instruction as had been the expectation. The lack of assessment literacy contributed to the time burden that teachers felt because they did not understand how existing assessments could be integrated into their daily instruction and the intervention process.
Moreover, teachers had not been provided with the necessary training to help them create their own high quality assessments which accurately measured specific goals. The diversity of roles within the teams, including teachers, student teachers, ancillary teachers, educational assistants, etc., meant that team members also came with a wide range of knowledge and skills and it was hard to provide meaningful, targeted professional development to meet each of the team members at their current knowledge level.

Systems-level issues also contributed challenges. Some teachers felt that the district-mandated Math Quarterly Assessments did not provide useful information because it had been developed to show only student progress on standards to be taught during that quarter, not on growth across the school year. The Math Quarterly Assessments were aligned to the district-wide curriculum maps. However, the school adopted new math instructional materials in SY08-09 and the sequence of the new basal series did not match the order in which standards had been presented in the curriculum map. Rather than pick and choose from the book to support the instruction of the standards, some teachers moved through the skills in the order presented in the book. Therefore, there was a systematic mismatch between the Math Quarterly Assessment and the instruction in some classrooms. During the second year of this study, a district-wide team developed a new math assessment designed to show student growth on key standards across the school year so that the mismatch would be eliminated.

Some teachers also felt that the district-wide interim assessments were not helpful for the measurement of student growth towards PDSA goals because the information provided by the interim assessments was more general than the PDSA goals unless the teachers dug into the data. During the first year of the study, teachers had to manually calculate scores by
standard for the Math Quarterly Assessment. During the second year of this study, the district introduced a new data warehouse and assessment management system which automatically reports how students performed by standard for the Math Quarterly Assessments or other assessments which teachers might create.

Finally, there was also a mismatch between district- and state-mandated improvement plans. The state requires that the schools create an Educational Plan for Student Success (EPSS) based on global data disaggregated by subgroup (e.g., ethnicity, participation in specific programs). The district-mandated PDSA plan, on the other hand, had asked teams to analyze student performance by specific skills. Once the team identified specific skills such as measurement that need improvement, they then further identified specific subskills and provided differentiated instruction to all students in that skill area. This mismatch has meant that two separate processes need to take place, contributing to the lack of ownership in the EPSS.

4. How does this initiative change the collaboration processes among teachers in the professional learning communities? Are there unintentional consequences in addition to the goals of the initiative?

The degree to which this initiative changed the collaboration processes among teachers in the professional learning communities varied across the teams. Within the third grade team there already had been a high degree of trust and psychological safety and the team had already initiated similar actions the prior year. However, the schoolwide initiative did nudge the third grade team to slightly more sophisticated levels of data analysis. Unfortunately, in order to conform with the schoolwide model of incorporating ancillary teachers and others into the intervention process, the third grade teachers felt that more time
was spent teaching others how to teach the particular skills and less time was spent on defining learning targets and differentiating instruction for students.

With several of the other teams, the requirement to create common formative assessments and collaboratively plan intervention blocks brought interpersonal differences to the forefront which had to be dealt with before the teams could make a significant transformation. While direct confrontation (a “blow out”) and third-party interventions were cited as catalysts for behavior change in at least several cases, several teams also utilized some tools which helped to create an atmosphere of increased safety. The first and fourth grade teams and the bilingual PLC all utilized conversation protocols learned at the Data Driven Dialogue training at several meetings to shift the dynamics in how members conversed and to make members more aware of norms of professional dialogue. The facilitators also learned to utilize depersonalized data, such as grade-level wide data, to decrease the risk to individual teachers.

These confrontations and tools did have some unintended consequences as well. First, Anne, the new leader of the first grade PLC, learned that it was okay for her to be more assertive and to stand up to people in a productive manner. Other leaders such as Marion, the fourth grade chair, also had to reflect on how to best lead the group without acting as “a mother”. Second, the use of protocols helped ancillary team members become more accepted. Eleanor, a literacy ancillary teacher who had been attached to the fourth grade PLC for two years, stated,

The other thing for me is that [the use of protocols] allowed me to be in that circle rather than an outsider. (3-2-10 p. 9)
Indeed, the inclusion of ancillary personnel in the intervention process also helped to broaden the view of student performance. Diane, the third grade chair, stated,

It has kind of forced us to look more at the whole child and get a more total picture. There’s been more of discussion [about how the child is doing in related arts] than I have ever seen before…So I know it was painful for the poor ancillary people to come in and do it but I think it was really good at the same time because it brought us together in a different way. (2-22-10 p. 7)

While one goal of the collaborative planning process was to promote greater ownership of the success of all students across homeroom teachers at the same grade level, it was perhaps less expected that the ancillary teachers would broaden the viewpoint of the homeroom teachers.

In addition to the conversation protocols and depersonalized data, concrete, specific, and precise tasks helped some teams change their collaboration processes. Wendy, a primary grades teacher, explained how the first grade team was able to move from contentious discussions to productive discussions,

Because we had something so focused to talk about, which vocabulary is necessary to make this connection, it was safe ground for people…Vocabulary was very specific, very precise, and people could go ahead and open up on that…rather than just how do we teach reading. (5-17-10 p. 8)

By focusing the discussion on specific skills and the students rather than teaching, the team members were able to shift their collaboration dynamics.

The initiative also highlighted the need for PLC team leaders who have good facilitation skills. For example, the productivity of the third grade team was credited primarily to the facilitation skill of its chair. When asked what had made the third-grade PLC effective, Karen, an ancillary teacher attached to the third grade team, stated,

I would say the leadership of the PLC is first and foremost the most important thing. We have an excellent leader…She’s very focused, very structured, very
knowledgeable. She has control of the group. If we start getting off on tangents, she brings us all back. (5-19-10 p. 12)

Team leaders were also instrumental in changing the collaboration processes of the bilingual and the first grade teams. After years of discussing the need to change the bilingual program, Leslie was successfully able to pull together a disparate group of staff members to create a comprehensive action plan for restructuring the bilingual program. Under the guidance of Anne, the first grade team was also able to stay focused on its task and change their collaboration dynamics. She stated,

We’ve looked at the data [from the common formative assessments] every time that we said we would. Everyone stayed on track with reporting their data to me. And then I put it into a graph the way that you showed me and that has been really effective. Because we can see where the kids are and we’ve had good discussions about is this a good goal. It’s really more of us looking at our teaching and if that is a good goal. (3-22-10 p. 3)

While the principal, Sharon, had tried to select teachers who she felt would be able to serve as good facilitators for this initiative, the team leaders were not systematically provided training to help them understand and manage the collaboration processes. Therefore, the quality of the team facilitation varied depending upon the talents that individuals possessed.

This initiative also highlighted the need to understand and address staff members’ personal histories and underlying belief systems. For example, several teachers’ own experiences as bilingual speakers in elementary school greatly influenced the type of program that they were able to philosophically support. Their own experiences led them to believe that a maintenance bilingual program could damage the ability of the students to excel in an English-language environment in the long run.

Another unintended side of effect of the focus on data and evaluative inquiry was that the degree of camaraderie at the school decreased over the two years of this study. By trying
to protect the time for PLC team planning and discussions, there were fewer opportunities to get to know each other as whole people. Sharon reflected on the decreased camaraderie,

       Honestly I don’t know the root cause. I mean, I think we’ve offered things. I think if you go back and look, the expectations have ramped up. That is systematic and also deliberate. The school under [a former principal] is whole different school that it is under me. The expectations have changed. (6-15-10 p. 14)

The decreased level of camaraderie was concerning because trust can act as social capital as the school engages in new and uncomfortable processes. Given the more focused work of PLCs under this initiative, it became more important, as Leslie, the bilingual team chair, explained, to incorporate relationship development processes into the work of PLCs,

       In the ‘real world’ we don’t have time for sharing…[like] telling about our pets…[It’s] just get to work, but there is a way to do the work in a very inter-relational kind of way. (5-19-10 p. 9)

Through that work, a sense of interdependency can be fostered. Diane, the third grade chair, explained the impact that the work and the common interventions had on the third grade team,

       One thing I noticed is when we all started working together and having to share the kids, everybody became everybody else’s problem. Or everybody else’s success. And so it was a lot less like ‘My kids do this’ or there was a lot less of that going on. It was ‘We need to get everybody to…’ (9-2-09 p. 16)

That sense of interdependency likely had the largest impact on the team collaboration processes within the initiative to develop a culture of evaluative inquiry.

5. Do changes in demographics, student achievement, perceptions, and processes occur over the course of the case study?

       During the two years of this case study, small improvements in student achievement were noted on the state accountability test.
<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th></th>
<th></th>
<th>Math</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2006</td>
<td>2007</td>
<td>2008</td>
<td>2009</td>
<td>2010</td>
<td>2006</td>
<td>2007</td>
<td>2008</td>
<td>2009</td>
<td>2010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Students</td>
<td>67.9</td>
<td>65.3</td>
<td>65.1</td>
<td>67.9</td>
<td>68.6</td>
<td>46.6</td>
<td>53.3</td>
<td>55.6</td>
<td>59.4</td>
<td>60.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>75.5</td>
<td>71.6</td>
<td>71.3</td>
<td>76.0</td>
<td>77.3</td>
<td>53.6</td>
<td>63.3</td>
<td>61.5</td>
<td>64.6</td>
<td>63.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>59.1</td>
<td>61.0</td>
<td>57.7</td>
<td>60.3</td>
<td>63.1</td>
<td>41.7</td>
<td>47.4</td>
<td>50.8</td>
<td>53.4</td>
<td>59.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English Language</td>
<td>61.3</td>
<td>43.9</td>
<td>47.2</td>
<td>21.4</td>
<td>31.6</td>
<td>41.3</td>
<td>36.6</td>
<td>39.6</td>
<td>39.3</td>
<td>57.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learners</td>
<td>Students with Disabilities</td>
<td>37.0</td>
<td>32.1</td>
<td>20.8</td>
<td>25.0</td>
<td>18.0</td>
<td>21.7</td>
<td>28.3</td>
<td>26.4</td>
<td>32.1</td>
<td>18.0</td>
<td></td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>62.5</td>
<td>59.6</td>
<td>53.7</td>
<td>66.0</td>
<td>66.0</td>
<td>38.2</td>
<td>46.0</td>
<td>49.5</td>
<td>56.0</td>
<td>59.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Given the large changes in the school’s demographics since 2008 like the increased percentage of students participating in the Free and Reduced Lunch Program, achieving even modest student achievement gains is significant. However, the modest gains have not been enough for the school to meet the Adequate Yearly Progress requirements of the ‘No Child Left Behind’ Act.
Table IV-2: Adequate Yearly Progress Trends at North Mesa Elementary School, 2008-2010

<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2008</td>
<td>2009</td>
</tr>
<tr>
<td>All Students</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Caucasian</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>African-American</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hispanic</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English Language Learners</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Students with Disabilities</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The school is currently designated as a School in Need of Improvement at the School Improvement II level and the staff members continue to fear that they will lose more autonomy if they are unable to raise student achievement enough to meet the mandated targets.

Changes in demographics, perceptions, and processes are described in different sections of this chapter and Appendix B.

Summary

Thus, the process of how staff members at North Mesa Elementary School strived to develop a culture of evaluative inquiry is quite complex. The school administrators purposefully held PLC teams more accountable for using the existing Plan-Do-Study-Act (PDSA) continuous improvement process to monitor student progress towards certain learning targets. Furthermore, they required that the PLC teams use the data gathered through the PDSA process to guide collaborative intervention blocks for the students.

The grade level representatives on the Instructional Leadership Team (ILT) served dual roles. First, they facilitated the collaborative inquiry process with their colleagues.
Second, they represented their colleagues at the ILT meetings where they expressed concerns not only about the initiative to create a culture of evaluative inquiry but also about other school matters.

The team members faced many challenges when implementing the initiative including external factors such as changing political and economic conditions, a deeply ingrained culture of instructional autonomy, the lack of agreement that change was even necessary, and the lack of a shared vision and common goals. Ultimately, significant progress was made in developing a culture of evaluative inquiry. However, the changes made were more idiosyncratic due to the efforts of individuals rather than a systematic and sustainable change throughout the school culture. While some teams did see improvement in student achievement on certain skills and the school was able to maintain overall student achievement levels despite changing demographics, appreciable increases in student achievement were not observed during the course of this study.

This chapter presented the results of this study in a linear analytic manner and Appendix B provides a richer, holistic description of the case. In the next chapter, I will discuss the implications of these findings for the field of educational reform.
Chapter V: Discussion and Implications

Over the course of this two-year case study, the instructional staff members of North Mesa Elementary School endeavored to improve student achievement through changing the manner in which teachers collaborated. The theory was that if teachers used data to identify instructional areas where their students struggled, collaboratively reflected on their instructional practice in those areas, and tried new practices, then student learning would improve. It was hoped that through the creation of a culture of evaluative inquiry, teachers’ mental models would be transformed and significant, long-lasting improvement to student achievement could occur. Since few educational reform efforts have achieved significant, long-lasting change in instructional practices, there is a need to document successful practices and understand barriers to achieving success in reform efforts.

Based on the assumption that the development of a culture of evaluative inquiry would lead to improved student achievement, the central question of this case study is: “How do staff members of a public elementary school develop a culture of evaluative inquiry?” This case study documents what actions teachers did and did not take to develop a culture of evaluative inquiry from 2008 through 2010 and profiles the multiple perceptions of staff members towards this initiative. If the staff members were successful in transforming the organizational culture to be evaluative inquiry focused, then it would be possible to see whether the transformation had an effect on student achievement. If the staff members were not fully successful in transforming the organizational culture, then it would benefit the professional literature to understand the barriers that were encountered so that other sites might be able to plan for and address those barriers in the future.
In this chapter, the major discussion points are presented with implications for other school districts that would like to create a culture of evaluative inquiry, followed by a model for the development of a culture of evaluative inquiry, and concluded with questions for future research.

**Discussion**

1. *Instances of significant learning and changes of practice were evident but the staff was not successful in fully developing and sustaining a culture of evaluative inquiry.*

   Every grade-level PLC and the Instructional Leadership Team demonstrated instances of significant learning. For example, teams reported that through the process they had learned to better define what students should know and be able to do, had developed more common vocabulary, and had improved their assessment literacy. In turn, this led to some teams using data in a formative manner rather than just collecting data for compliance purposes and greater alignment between instruction and grade-level goals. Furthermore, the process led to some shifts in attitude including more common ownership of students and a more holistic manner of looking at students as ancillary teachers were included in the conversations. Finally, the process resulted in increased teamwork and greater acceptance that differentiated instruction and interventions are needed to better meet the needs of the students.

   While significant learning and changes of practice were evident, the shift to a sustainable culture of evaluative inquiry was not completed during the timeframe of this study. King (2004, p. 338) stated that a sustainable culture of evaluative inquiry “is one that lives independently in an organization”. In other words, evaluation processes are naturally integrated into the everyday work of the organization. Deal and Kennedy (2000) explain that
means that virtually all of the people in the organization identify with new role-model heroes, tell different stories to one another to explain what is occurring around them, spend their time differently on a day-to-day basis, ask different questions, and carry out different work rituals.

Even at the end of the study, the principal stated that she was still receiving feedback that “PLCs were doing too much data” and the degree to which evaluation processes had been incorporated into instructional routines varied widely across staff members. Thus, it cannot be said that the changes that had been made would be sustained independently of continued conscious efforts from the Instructional Leadership Team.

It is not surprising that the school was unable to fully achieve a cultural change within two years. Deal and Kennedy (2000, p. 163) state, “It literally takes years to achieve fundamental change in an organization’s culture.” The force of the old culture can emasculate proposed changes when those changes threaten important, long-held values. Even when staff members understand and accept the change, they may not have the skills and knowledge necessary to engage in new practices (Deal & Kennedy, 2000).

Since the staff members were not fully successful in creating a culture of evaluative inquiry despite conscious efforts to do so, it is appropriate to explore the barriers that were encountered so that other sites might be able to plan for and address those barriers in the future.

2. Evaluation Capacity Building was hampered by a lack of resources and a lack of system alignment and communication, but process use was evident.

The shift to a culture of evaluative inquiry was to be facilitated through Evaluation Capacity Building (ECB) activities. Preskill & Boyle (2008, p. 444) state,
ECB involves the implementation of teaching and learning strategies to help individuals, groups, and organizations learn about what constitutes effective, useful, and professional evaluation practice. The ultimate goal of ECB is sustainable evaluation practice—where members continuously ask questions that matter, collect, analyze, and interpret data, and use evaluation findings for decision-making and action.

Evaluation Capacity Building is often explicit. Some explicit training did take place. For example, seven members of the Instructional Leadership Team did attend the data portfolio conference sponsored by Education for the Future and four staff members participated in a data-driven dialogue conference. Furthermore, an outside consultant did work with members of the Instructional Leadership Team for one day to assess their situation and develop an action plan. However, a lack of funds and time hampered the ability to have all staff members participate in explicit Evaluation Capacity Building activities.

The lack of investment in the development of culture change is not surprising. Graham states (2005, p. 198),

Deeply understanding the culture, the constraints, and the opportunities of schooling was not high on the intellectual agenda of educational researchers for much of the twentieth century...At least one explanation for this oversight was the lack of money for the task. Money for educational research has always been scarce in comparison with what we spend for other worthwhile activities, such as research on health, agriculture, or defense...In a nation that has traditionally believed in research and development in other fields, there has been precious little “R” on which to develop the “D” in education.

Graham (2005, p. 199) compares the percentage of funds spent on Research and Development in 1989 among various federal agencies as evidence that less is invested in education than in other fields. For example, the Office of Educational Research and Improvement invested 5% of its budget on research and development activities in 1989 whereas the Agriculture Research Service spent 46.6% of its budget, the National Institute of
Health spent nearly 60% and National Science Foundation spent 93.5% of its budget on research and development.

Moreover, even if funds had been available to sponsor additional trainings or a book study and to pay for substitutes for the teachers, the teachers felt overwhelmed and did not feel like they had the time to participate in explicit ECB activities. Furthermore, some teachers would have resented any time taken away from their immediate classroom needs, even if they did see the long-term relevance.

Fortunately, not all ECB efforts need to be explicit. Process use is one mechanism through which evaluators can indirectly but actively help organizations build evaluation capacity. Cousins, Goh, Clark, and Lee (2004, p. 107) state, “if evaluation becomes integrated into the ongoing activities within an organization, it may become a learning system that fosters the development of shared values and understanding among organization members.” By requiring teams to collaboratively set and evaluate PDSA goals and to collaboratively plan for grade-level wide interventions, the Instructional Leadership Team created a structure in which teachers could learn evaluative skills while engaged in an evaluation process.

Process use occurs when “individual changes in thinking and behavior, and program or organizational changes in procedures and culture...occur among those involved in evaluation as a result of the learning that occurs during the evaluation process” (Patton, 2007, p. 90). The results of this study indicate significant process use throughout the inquiry as many teachers cited examples of how they had learned about the process for setting goals, the process for defining proficiency, and how to assess in a meaningful manner. Furthermore,
there were numerous examples of teams learning how to use the data to sort students into flexible groups and provide differentiated instruction.

Through the support of the school administrators, the teachers at the school were able to reach a crucial goal as described by King (2007, p. 53),

The key question is how to engage multiple people in continuing discussions related to evaluative thinking so that many learn from ongoing activities. It means systematically and purposefully applying the experiential learning cycle...If possible, every member of the organization needs to understand that evaluation is a part of his or her job description and there is a structure in place to support people in learning about what this means.

By the end of this study, there was an acknowledgment by teachers that the expectations of them as professionals had changed for the long-term. One teacher stated that although the process had been difficult to implement,

The entire campus has now come together and seen that there is a real need to meet these kids and provide them with interventions. I mean I don’t think it was universally recognized…Everybody knows that it is going to affect and impact their life now. (2-22-10 p. 7)

While some evaluative inquiry skills undeniably were learned through this structure, even the most evaluation-oriented teams felt that they “have a long way to go” before the use of evaluative inquiry skills could be considered a normal part of their daily routine.

Part of the struggle is that not all parts of the system are aligned to facilitate the learning of evaluative skills. Many of the assessments that teachers are required to administer are mandated by the district. There has been insufficient communication regarding the purpose of each of the assessments and the alignment among the various assessments. Teachers have had available resources that would help them monitor student progress over the course of the school year such as the NWEA Measures of Academic Progress yet many teachers either did not know how to access or did not have time to access
the data in meaningful formats or they had not yet learned how to integrate that information into their instruction. The district needs to ensure that there is a balanced assessment system which meets the needs of all the information users and that all users understand and can access the data that they need.

Other impediments were also noted in this study. The district-wide Math Quarterly Assessment, for example, was not aligned to the sequence of skills that was being taught within the school. The absence of a data warehouse hampered the retrieval of information regarding student achievement. Since the end of the study, the district has created a Math Quarterly Assessment which measures student growth on the same standards across the year so that the alignment between instruction and the assessment has improved. Furthermore, the district has implemented a data warehouse and assessment management system. Thus, impediments to sustained learning may continue to exist but the nature of those impediments change over time.

3. Neither top-down mandates nor complete teacher empowerment would have led to a significant change in organizational culture and teachers’ instructional practices. However, it is challenging to create a balanced system of defined autonomy where teachers are focused on “the right work”.

Although the school had not made Adequate Yearly Progress for several years, many teachers in the school did not perceive that significant changes in their instructional practices were necessary. One teacher stated,

The entire framework of the school was not messed up. It was this one group we were working with [students with disabilities]. It feels like we’re going in with hatchets and chopping everything out instead of saying, ‘Let’s look at this area and see where we can go ahead with that’. (10-23-08 p. 11)
This viewpoint is typical not only of many teachers within North Mesa Elementary School, but also across the United States. Changing the organizational culture and the way teachers interact requires significant changes in instructional practices. Teachers’ instructional practices are intricately linked with their personal identities and self-efficacy as teachers. Van Clay, Soldwedel, & Many (2011) point out that teachers may ask why they have to change what they have perceived to be successful for a number of years when change initiatives like shifting to a culture of evaluative inquiry are suggested. In order to embrace the change, they need to see how the changes align with rather than conflict with that which they believe that they do well. Given the personal nature of instructional practices, instruction is often the last area that teachers will change. Indeed, there is ample evidence that giving educators site-based autonomy does not ensure that they will use that authority to focus on matters essential to teaching and learning (DuFour & Marzano, 2011; DuFour, DuFour, & Eaker, 2008).

Furthermore, granting teams full autonomy with all instructional matters will potentially lead to inconsistency and misalignment across the school and across the district. Fullan (2009) noted that “Bottom-up change—so-called let a thousand flowers bloom—does not produce success on any scale. A thousand flowers do not bloom and those that do are not perennial.” Prior history at North Mesa Elementary School demonstrates this inconsistency as some grade levels had collaboratively engaged in grade-level wide interventions and other teams rarely even discussed instructional topics. Furthermore, different teachers and different teams had idiosyncratic goals and methodologies leading to widespread misalignment that Sharon, the school principal, had termed “the Popcorn Effect”.

127
In addition to being reluctant to reflect on and change instructional practices, many teachers were hesitant to become interdependent with other teachers and to subsequently give up some of the autonomy within their classroom. Collaboration time is perceived by some teachers as the loss of valuable time that could be spent within their own classroom. As one teacher explained,

"I think part of it too is the shift to the PLC in general. It’s the shift to having conversations all the time. I don’t know that everybody wants to have a conversation all the time about what to do." (2-24-09d p. 2)

She explains further,

"Teaching is very personal (2-24-09b p. 13)…Instead of looking at the PLC as a group that’s supposed to help everybody, they don’t see it that way. They are afraid to talk about what they do or they are afraid to ask for help because they think that that reflects badly on them…They don’t ask ‘What are you doing? Your scores look really good’. They don’t want to know. They just say, ‘Well, this is what I am going to do’. (2-24-09b pp. 11-12)

Hesitation to shift from a culture of independent contractors to interdependence and collaboration is also found in many schools across the United States. DuFour, DuFour, and Eaker (2008, p. 354) have asserted that “collaboration by invitation won’t work.” However, there is also significant research that demonstrates that attitudes follow behavior (Pfeiffer & Sutton, 2000). Therefore, it would appear that mandating that teachers collaborate regarding instructional issues would help at least some teachers overcome their initial hesitations and lead to lasting change.

Paradoxically, top-down change also doesn’t work “because it fails to garner ownership, commitment, or even clarity about the nature of reform” (Fullan, 2009). While two grade level teams had already voluntarily implemented collaborative grade-level wide interventions and had started shifting towards using common formative assessment data, the
other four grade levels perceived the initiative to develop a culture of collaborative evaluative inquiry as a top-down mandate. Some teachers viewed the initiative as an infringement on their autonomy as this comment illustrates,

> It’s time to quit ramming everything down everybody’s throat. Let us do the things that we want to do and I think that will make a huge difference…We don’t need the shotgun approach of ‘Some of them aren’t doing it, so let’s blast it’. (10-23-08 pp. 9 & 10)

Some teachers expressed concern about a lack of clarity regarding the vision and the parameters of the perceived mandate and they expressed concern that collaborative grade-level interventions and common planning would not have the desired effect. In fact, there was not only a lack of ownership and commitment among some teachers at the beginning of the initiative, but some teachers also openly vowed to sabotage the process.

Thus, there are serious questions about whether either a top-down or a bottom-up approach to school reform can work. A balance between a top-down approach and a bottom-up approach, like a balance between a post-positivist approach and a socio-constructivist approach, may be necessary. The question then becomes how a school like North Mesa Elementary School can develop a culture of evaluative inquiry. Researchers have recently begun to advocate a system of defined autonomy whereby educators enjoy some latitude within specific non-negotiable parameters (DuFour & Marzano, 2011; Marzano & Waters, 2009; Von Clay, Soldwedel, & Many, 2011). The theory is that leaders should encourage autonomy and creativity (“be loose”) within well-defined parameters and priorities that must be honored (“held tight”) (Van Clay, Soldwedel, & Many, 2011, p. 24). The art of leadership is to accurately determine an effective balance between the two.
Sharon appears to have tried a defined autonomy approach to the implementation of this cultural shift. She consistently told Instructional Leadership Team members that collaboration to provide differentiated interventions for their students was compulsory. Furthermore, teams needed to set common goals, analyze common formative assessment data, and collaboratively plan for the interventions. Within those parameters, teams could choose their goals, how they were going to measure progress towards the goals, and how the differentiated interventions would be provided. According to Van Clay, Soldwedel, and Many (2011), Sharon focused on the right areas: ensuring a focus on learning, building a collaborative culture, and establishing a results orientation. Yet the school still was not successful in fully creating and sustaining a cultural shift within the two year timeframe of this study.

One issue may have been the lack of common vocabulary. DuFour, DuFour, & Eaker (2008) state, “many organizations settle for superficiality in language, using terms so ambiguously and loosely that they can mean very different things to different people” (p. 29). Given the lack of time to discuss and develop a shared vision and the lack of consistent professional development, it is likely that there was not a common understanding of what the non-negotiable expectations actually were. As Sparks (2006, p. 93) said, “it is difficult to establish something when there is no clear picture as to what it should be.”

Although the school was not successful in fully changing the organizational culture during the course of this study, the teachers were engaging in “the right kind of work”. Through the collaborative dialogue process, the teachers were building shared knowledge. They were collaboratively defining what students should know and be able to do at each grade level and they were sharing ideas about how best to teach certain skills to reach the
differentiated needs of students. Thus, the time spent was valuable although it was insufficient to fully shift all teachers’ attitudes and beliefs within the timeframe of this study.

4. Challenges encountered in the development of a culture of evaluative inquiry indicate an incomplete shift to standards-based teaching.

Observations from this case study show that a previous paradigm shift to standards-based teaching had not been fully achieved at North Mesa Elementary School. The requirement for standards-based teaching was incorporated into the 1994 Elementary and Secondary Education Act (National Academy of Education White Paper). The goal of the policy was to clarify what students should know and be able to do so that a variety of textbooks would not be the de facto curriculum. However, many of the teachers in this study were challenged with identifying particular standards on which to focus in their PDSA inquiry process. Together they struggled to specify definitions of proficiency on those skills. Many of the teachers also were uncomfortable with pulling together a variety of resources to support instruction during the intervention block instead of following the sequence of activities in a textbook.

Research has shown that many schools across the country lack an understanding of the changes that were needed to shift to true standards-based instruction and lack the capacity to make those changes happen (National Academy of Education White Paper, p. 2). Few states invested in the sustained professional development that would be necessary to develop the capacity. Furthermore, in many states the alignment between standards, instructional resources, and assessments tend to be superficial which makes focused instruction even more difficult (National Academy of Education White Paper, p. 4).
If the teachers had fully implemented standards-based instruction, the shift to a culture of evaluative inquiry would have been easier for them. Once they collaboratively determined the standards on which they would focus their PDSA inquiry process, they could collaboratively define the performance standards through which proficiency would be demonstrated and could create the common assessment. Then each teacher could plan how best to prepare students to reach that performance standard. Thus, the failure of previous initiatives to be fully implemented diminished the probability of success of new initiatives that had been based upon the prior initiatives. It may be necessary to include professional development regarding standards-based instruction into professional development for developing a culture of evaluative inquiry.

5. *Data is personal and threatening to teachers’ self-efficacy.*

Traditionally, teachers have been able to set the proficiency targets for students and determine how to help their students reach those targets on their own. Furthermore, teachers created their own assessments to measure student progress. Any differences in the achievement of different classes on standardized tests, if noted at all, were often attributed to differences between the groups of students.

Teachers have also protected themselves by focusing on teaching, rather than learning. It is less threatening for a teacher to take a teaching-focused approach and say ‘I covered the curriculum and the students chose not to learn it’ than to take a learning-focused approach and say ‘I did not adequately facilitate their learning on this standard’. A culture of evaluative inquiry requires the acknowledgment that no matter how good a teacher is, there are specific areas of instructional practice that can be improved. Even if most students in the class meet expectations, the teacher must believe that a tweaking of practices could either
lead to *all* students meeting expectations or that some students could perform at a higher level.

In a culture of collaborative evaluative inquiry, the focus must be on the students’ divergent skills and teachers must be willing to remediate those skills where students are not making adequate progress. Some of teachers’ resistance to truly analyzing students’ progress and using it to drive their instruction may be based on insecurity. The shift towards common goals and common assessments enables the scores from different classes to be compared to each other. Teachers are afraid that the scores of their students will be compared and will be found lacking which is partially a reflection of their ability as a teacher. As one participant in the study noted, “[The teachers] feel that if they make their scores public and they really look at that data in a reflective way that it is a judgment on them and themselves as a teacher.” (6-3-10 p. 6) Adding to that fear is the fact that teachers do not control all of the factors that influence student achievement. Some teachers may be afraid that they could be unfairly held accountable, or at least will be perceived in a negative light, for reasons that are partially beyond their control. Recent legislative initiatives including basing at least fifty percent of a teacher’s annual evaluation on the scores of their students on one standardized test have only served to increase the anxiety around the use of data and have raised additional questions about the validity and reliability of some decisions that could be made based on the data.

Furthermore, the proposals within the recent legislative initiatives are inconsistent with the creation of collaborative cultures. While the goal of creating a culture of evaluative inquiry is to work collaboratively to improve the achievement of all students, the legislative proposals include the awarding of financial incentives to individual teachers. Prior to the introduction of these legislative initiatives one teacher had noted, “Teaching is not a
competition…But not everybody buys into that.” (2-22-10mi p. 3) Evaluation structures that reward individual achievement rather than collective achievement will only serve to reinforce the culture of competition.

Based on the findings of this case study and the current research base, I propose the following model for school districts who would like to create a culture of evaluative inquiry in order to create sustained improvements in student achievement.

**The Miller Evaluative Inquiry Model**

This case study has shown the very complex nature of trying to institute a shift in organizational culture and illustrates how contextual factors influence the implementation of the change initiative. Yet the issues that the teachers at North Mesa Elementary School faced are typical of the issues that teachers across the country face. Thus, it is appropriate to offer a generalized model which combines the research literature with the experience of the staff members at North Mesa Elementary School in order to help other schools in their quest to develop a culture of evaluative inquiry.

**Purpose and goals of the Miller Evaluative Inquiry Model.** The Miller Evaluative Inquiry Model details a collaborative process for small teams to solve ill-defined problems which emerge during the course of their instructional practice through an iterative evaluative inquiry process. While explicit Evaluation Capacity Building activities can be utilized to supplement the learning of team members, the majority of learning takes place through process use. The model also touches on some knowledge management processes and tools which facilitate learning and communication with colleagues which in turn promotes organizational learning.
The Miller Evaluative Inquiry Model has three goals based on a socio-constructivist perspective of individual and organizational learning:

1. School practitioners will change their understanding or will create new personal constructions of knowledge in order to solve ill-structured problems that emerge in their instructional practice.

2. School practitioners will develop collaborative group process skills in order to socially construct new knowledge.

3. Knowledge that is created in one professional learning community will be transferred to other professional learning communities via the knowledge management system.

**The big picture.** The Miller Evaluative Inquiry Model addresses many of the concerns raised about traditional professional development practices such as an emphasis on stand-alone workshops isolated from the instructional context and the differentiated needs of the students and teachers. In this model, teachers are clustered around common interests. For example, in an elementary school, the same grade level teachers could form a PLC team. In a high school, members of the same department could form a PLC team. Within a school community, there would be multiple professional learning community teams comprised of three to eight members each which concurrently solve problems related to their own instructional practice through evaluative inquiry. In order to solve the problems, both individual and group learning must take place.

Figure V-1 depicts “the big picture” of the Miller Evaluative Inquiry Model.
Figure V-1: The Big Picture of the Miller Evaluative Inquiry Model.

The teams work within a system of defined autonomy that is aligned between the school and the district. Within this system there are some non-negotiable expectations: a shared overarching vision; a focus on standards, learning, and results; and the need to collaborate. Both the district leadership and the school instructional leadership team need to work together to create a set of systems and structures that are aligned and are supportive to the creation of an effective culture of evaluative inquiry. Information needs to be communicated between the district leadership, school leadership, and PLC teams through a
knowledge management system. Underpinning the work of both the district staff members and school staff members is interaction with national best practices.

The first element of the system of defined autonomy is an overarching vision that is shared at all levels: district, school, and team. A study conducted by researchers at Midcontinent Research for Education and Learning (McREL) found that 71% more instructional time than is currently available would be needed to adequately cover the content in state and national standards (Marzano & Waters, 2009). It is necessary to prioritize and set a non-negotiable vision towards which all staff members work. The creation of a shared vision allows teams and schools to create a guaranteed viable common curriculum across the schools. Furthermore, it enables the creation of supportive structures and systems.

While the overarching vision and goals should be non-negotiable, teams should be empowered to have some autonomy regarding how best to meet the overarching vision and goals. There is a tension between the need for standardization to guarantee a common viable curriculum and educational experience, and the need to allow for differentiation to best meet student needs and to allow teachers to tap into their creativity and develop ownership.

Furthermore, teams should have some latitude to start with the goals that are the most concrete or interesting for the teachers within the overarching vision so that the PLC teams can establish new routines and processes before tackling more challenging goals. Some skills are easier to address than others. For example, the first grade, third grade, and fifth grade teams all explicitly stated that it was easier for them to set goals and collaboratively plan for mathematics instruction as compared to reading instruction. They felt that math was more concrete and progress was easier to measure. Furthermore, there was less disagreement within teams about the best way to teach mathematical skills.
Therefore, it makes sense for teams that are new to the process to select goals that are concrete and of interest to them. Establishing new routines and processes is challenging even when concrete, non-controversial goals are selected. However, the additional burden of struggling to measure complex or ill-defined skills or navigating through the interpersonal dynamics with a controversial topic can overwhelm teams and cause them to give up in frustration. Establishing new patterns of behavior and developing ownership of the process are more important initially than focusing on the area of greatest need.

The second element in the system of non-negotiable expectations and defined autonomy is a focus on standards, learning, and results. A shift is currently taking place in education from a focus on teaching, or the activities of the teachers, to a focus on what and how much students have learned. In order to focus on student achievement across schools, it is important that teachers focus on standards—a description of what students should know and be able to do—instead of following the content that different publishers have selected. A system-wide focus on standards, learning, and results is essential to the successful establishment of a culture of evaluative inquiry. School leaders should ensure that essential skills and concepts have been defined and that all teachers are focusing on the acquisition of those skills and concepts. Moreover, school leaders need to ensure that teachers formatively and summatively assess student progress towards the common learning targets and systematically provide assistance to those students who need additional support. Within these broad parameters, however, teams should be empowered to determine how best to help their students reach the common targets that the PLC team members collaboratively set.

Another non-negotiable expectation in the system of defined autonomy is the creation of a collaborative culture. The school instructional leadership team can help foster the
development of a collaborative culture by creating interdependency and establishing the need
to actually use the data. In this study, the school leadership created interdependency by
requiring that PLC teams collaboratively plan for a weekly grade-level wide intervention
block. Teams needed to create common formative assessments in order to place students into
flexible groups which increased the awareness of and use of formative assessment data.
Without this structure, many of the teams may not have moved in this direction. However,
interdependency and the need to use the data alone may be insufficient. School leaders must
support teams by providing additional training and time so that teachers learn how to use data
in a more effective manner.

Furthermore, school leaders and team facilitators need to pay attention to
interpersonal dynamics and teachers’ fears and then scaffold the activities in order to develop
both interpersonal trust and trust in the evaluative inquiry process. Teachers’ instructional
practices are intricately linked with their personal history and belief system. Personal beliefs
and fears need to be articulated and addressed. For example, through discussions staff
members in the bilingual program realized that their instructional philosophies were directly
related to their own personal experiences. The articulation of these beliefs allowed school
leaders to address some staff members’ fears and also led to at least one staff member
realizing that she needed to take a different role than had initially been discussed. It is,
therefore, important to give everyone a voice in the change process so that their personal
needs can be addressed.

The discussion of these beliefs and fears, however, can be very threatening and are
best conducted in a climate of interpersonal trust. Therefore, school leaders may need to
scaffold activities to develop trust. For example, school leaders can start by having teams
analyze depersonalized data such as grade-level wide data instead of classroom-specific data. In this manner, the focus is on student learning and not on outcomes within individual classes. Over time, as trust is developed, the growth data can be organized on the classroom level to facilitate conversations about successful strategies that some teachers have used.

This case study confirmed that interpersonal dynamics play a large role in the collaboration process. Several teams could not productively work towards their goals until interpersonal conflicts were addressed through outside mediation, outside facilitation or confrontations. The dynamics observed at North Mesa Elementary School are typical of the dynamics within many schools and organizations. Therefore, school leaders should consider training team members to use conversation protocols in order to create the psychologically safe conditions that would allow for the development of interpersonal trust. As teachers learn to productively discuss instructional issues in a safe atmosphere, they will begin to also develop more trust in the process and the scaffolding supports can slowly be removed.

In order to achieve these non-negotiable expectations, the district and school leadership need to create structures to facilitate the work. For example, the district and school leadership need to create time within the workday for team members to collaborate. They need to have well-articulated curricular and assessment expectations and resources. Furthermore, they need to provide a robust knowledge management system that not only allows team members to quickly and easily access data and manipulate it, but also serves as a knowledge repository so that the learning of one PLC team can be disseminated to other teams. The knowledge management system should also give teachers access to relevant information about educational research and practices that are currently regarded as best instructional practice. Between the knowledge management system and multiple venues for
communication such as success sharing meetings, group learning can be transformed into organizational learning.

Finally, resources and actions need to be aligned to the shared vision. School leaders need to evaluate whether requests for resources or actions follow the plan for achieving the school’s vision and must have the strength to make some staff members unhappy in order to avoid “the Popcorn Effect”. While the vision must be held constant, the tactics for reaching that vision may need to be adapted over time based on lessons learned.

**The evaluative inquiry process.** The evaluative inquiry process embedded within the aforementioned big picture helps school practitioners learn from the challenges that they face on a daily basis. The process has a dual purpose: to improve the professional knowledge and practice of the school practitioners and to improve the learning outcomes of their students. The process has five main stages: Direction Setting, Problem Exploration, Incubation, Growth, and Integration as shown in Figure V-2. Once a satisfactory solution has been integrated into daily practice, then the professional learning communities start the process over with a different challenge. Not only is the entire process iterative in nature, but some of the sub-stages, Reflection and Refinement, reflect the iterative nature of reflection, modification, and testing which occur within the process.
The first stage of the evaluative inquiry process undertaken by PLC teams is Direction Setting. Although time is a precious commodity, the facilitator must take a significant period of time at the beginning of the cycle to establish the need for change. Teachers generally are expected to accomplish more than they reasonably can within their contracted hours. Any time a person engages in a new process, there is a learning curve and the person takes longer to accomplish the process than when they have experience and have become proficient at the process. There also is an emotional toll when moving away from
something that is familiar and comfortable to the unknown. Therefore, teachers are unlikely to embrace change of any sort unless the need for the change has been clearly established.

Establishing the need for change may not be easy given that teachers have different belief systems. For many years throughout the history of education in the United States, the purpose of schooling was to sort students (DuFour & Eaker, 1998). Not all students were expected to achieve at very rigorous levels. Graham (2005, p. 160) states,

Achieving academically had always been the goal for a few, including those living in poverty and those of color, as well as those who expected or whose parents expected them to attend an academically demanding college and for some who simply found academic study congenial to their talent and temperament. But it was not an expectation that was applied to all students.

When not all students achieve at high standards, some teachers cite that as evidence that students lack personal responsibility or that students are limited by their innate aptitude. Therefore, great care and time needs to be taken to have teachers reflect on their personal beliefs and the role that teachers can play in helping struggling students achieve at high levels.

These philosophical discussions lead into the creation of a shared vision. Catch phrases that may arise like “all students can learn” or “world class” need to be translated by staff members into specific actions and behaviors. Frequently, phrases and terms including “collaboration” and “professional learning communities” are not adequately discussed or defined. Therefore, unless there are specific discussions, different staff members may actualize the terms in different ways which would create a misalignment even though each staff member may believe that they are striving to reach a shared vision. The facilitator also needs to ensure that the shared vision and collective behavior commitments of the team are aligned with the overarching vision of the district leadership and the school leadership.
The second stage of the evaluative inquiry process is Problem Exploration. PLC team members conduct an in-depth analysis of several “big picture” summative assessment reports such as reports from the Standards-Based Assessment (SBA) or Northwest Evaluation Association’s Measures of Academic Progress in order to identify potential areas for exploration. The data should directly relate to the performance of the students that they teach. With the support of the facilitator, the PLC team members assess the viability of the identified instructional problem areas for collaborative problem solving action. The following questions (based on questions found in Marquardt, 2004 and Rothwell, 1999) serve as guidelines when selecting the most appropriate problem case for further inquiry by the team:

- What is the description of the instructional problem as perceived by the team members?
- What will happen if no action is taken to solve or manage the problem?
- Is the problem sufficiently important to the students and teachers to warrant devoting time and effort to improvement efforts?
- Does the problem lend itself to improvement based on concerted action by the professional learning community?
- Is there a time frame for taking action?
- Will the effort build important knowledge and competencies that teachers will need in order for students to be successful in the future?

Based on the answers to these questions, the members of the professional learning community select the problem most suited for further evaluative inquiry.

After an instructional problem has been selected for further inquiry, the PLC team members will analyze more specific achievement data, if available, and undergo a problem exploration process. More specific achievement data could potentially provide more nuanced information about the particular instructional strengths and weaknesses of the students. However, it may be helpful for the team members to predict what that data will show prior to
actually examining this data. As they create their predictions, they should also list the hunches and assumptions that underlie those predictions. In this manner, assumptions can be brought into consciousness. Furthermore, reflective questioning can also open up the possibility of new perspectives and interpretations of the data to team members. However, facilitators may need to closely manage this process, possibly with the help of conversation protocols, to ensure that all team members have the opportunity to participate in a psychologically safe atmosphere.

The purpose of the questioning is to deepen the team members’ understanding of the situation they are addressing. Marquardt (2004, p. 77) states, “the key to problem solving is to start with fresh questions, not constructs and assumptions from the past. Questions enable groups to unpeel the layers around the problem and uncover the core elements of knowledge necessary to discover the solution.” Table V-1 shows ten types of questions which build the group’s capacity to reframe the problem, build common goals, develop potential strategies, and take effective actions.
### Table V-1: Types of Questions to use in the Evaluative Inquiry Process

<table>
<thead>
<tr>
<th>Type of question</th>
<th>Definition</th>
<th>Example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open question</td>
<td>Gives the person or group a high degree of freedom in deciding how to respond</td>
<td>What would be the best results if we took that action?</td>
</tr>
<tr>
<td>Affective question</td>
<td>Invites members to share feelings about an issue</td>
<td>How do you feel about sharing students for intervention purposes?</td>
</tr>
<tr>
<td>Reflective question</td>
<td>Encourages more elaboration</td>
<td>You said these students struggle with measurement; what do you think causes these difficulties?</td>
</tr>
<tr>
<td>Probing question</td>
<td>Causes the person or group to go into more depth or breadth on a topic</td>
<td>Why is this happening?</td>
</tr>
<tr>
<td>Fresh question</td>
<td>Challenges basic assumptions</td>
<td>Why must it be that way? Has this ever been tried?</td>
</tr>
<tr>
<td>Question that creates connections</td>
<td>Helps to create a systems perspective</td>
<td>What are the consequences of these actions?</td>
</tr>
<tr>
<td>Clarifying question</td>
<td>Results in further descriptions and explanations</td>
<td>Are you saying that? Could you explain more about this situation?</td>
</tr>
<tr>
<td>Explorative question</td>
<td>Opens up new avenues and insights and leads to new explorations</td>
<td>Have you explored/thought of…? Would such a source help?</td>
</tr>
<tr>
<td>Analytical question</td>
<td>Examines causes and not just symptoms</td>
<td>Why has this happened?</td>
</tr>
<tr>
<td>Closed question</td>
<td>Can be answered by “yes” or “no”, or a quantitative response; can be useful to clarify or seek further understanding and quickly move the group forward</td>
<td>Did you agree with this decision?</td>
</tr>
</tbody>
</table>

Source: Based on Marquardt, 2004, p. 77-78.

Questions are used throughout the evaluative inquiry process, especially in the problem exploration phase. Since most instructional problems are ill-structured, in other words they are frequently vague, unpredictable, and may have some information missing, there may be several different plausible solutions or there might not even be a fully satisfactory solution (Kirkley, 2003). It is important that PLC team members understand the perspectives of all the stakeholders and be receptive to different suggestions for solving instructional problems.
Once PLC team members have identified the various perspectives of the selected instructional problem, examined the data, and inferred potential solutions, they need to examine constraints on the potential solutions before selecting the best solution given the context. Constraints may include a lack of resources such as money or time. Constraints could also include the lack of authority or lack of support from leadership to make the systemic changes needed to support the chosen solution. While some constraints will be known from the beginning, others may emerge during the evaluative inquiry process.

Through the process of iteratively exploring alternative arguments and possible solutions, team members not only arrive at a defensible solution to test through active experimentation, but they also develop their own mental models.

The third stage of the evaluative inquiry process is called Incubation. Once the members of the professional learning community have selected the most viable solution to test, they need to create an action plan for implementation. This action plan not only details what action each team member must take in a given timeframe, but it also details the needed resources and mini-trainings. PLC team members likely will need access to additional information. If they already possessed the necessary knowledge and skills, they would have already utilized them in their instructional practice unless contextual barriers were insurmountable. Therefore, the team members need access to case-based examples, on-demand mini-trainings arranged by the facilitator and school administration, and instructional resources. The needs of each professional learning community within a school will vary depending upon the topic of the selected problem and the background of the team members.

Once the team members have prepared a joint action plan, they will try out the solution in their classes. In order for these evaluative inquiry teams to be able to function
effectively and to be able to challenge the status quo, they must be empowered to take risks. They must be allowed to experiment and to occasionally fail. However, the teams also must gather data about the progress of their initial implementation in order to make a formative evaluation of the solution.

The most important phase of the evaluative inquiry process is the Growth phase. Individual reflection, public reflection, and the dialogue process are interwoven throughout this stage. Individual reflection is important because when people “critically examine their tacit values, assumptions, beliefs, and mental frameworks within which they understand and solve problems” it can lead to radical generative learning known as double loop learning (Argyris, 1982). This type of learning contrasts with incremental, single loop, learning where people do not question their mental models and assumptions. Since mental models and attitudes influence to which data a person attends, changes to mental models and attitudes can greatly change the way a person behaves.

The PLC team members should continuously question their assumptions and routines, test new possibilities, reflect on the results, and make adjustments as necessary. This gentle challenging can also make organizations aware of discrepancies between espoused theory and theory-in-use. Team members’ tacit mental maps govern their actual behavior, their theory-in-use. Their behavior frequently contradicts their spoken, explicit beliefs, or their espoused theory (Argyris & Schön, 1974). Reflection and dialogue can make team members aware of the discrepancy and provide them with an opportunity to bring their behavior and their beliefs into greater alignment.

Facilitators can guide team members in productive reflection and dialogue by using some of the following questions:
• Are we using open, reflective, and probing questions?
• Is everyone involved in the questioning?
• Is the listening attentive and open, or is it evaluative and inattentive?
• Are we filtering out what the person is saying?
• Are new insights arising, and are people making connections with the diversity of questions and opinions being offered?

(Based on questions found in Marquardt, 2004)

Public reflection and focused dialogue among collaborative team members are also essential to the process of sharing knowledge. Much of the knowledge created by teachers is “hidden, implicit, informal, and created in a myriad of ways” (Martin de Holan, Phillips, & Lawrence, 2004). Furthermore, much of the knowledge has been created in a fragmented manner. Therefore, public reflection and focused dialogue will help educators to articulate and disseminate their tacit knowledge. The professional learning communities create ways to institutionalize the most valuable of the shared knowledge through the knowledge management system.

Growth also occurs as team members formatively evaluate the progress of the implemented solution and make appropriate refinements based on the analysis. The team members then develop and implement the modified solution, gather data about the second implementation, and again engage in the cycle of reflection and dialogue. This process repeats until the solution is found to be satisfactory to all stakeholders.

Once the team decides that the solution is satisfactory, the team should report their progress to all stakeholders. Stakeholders may provide feedback which would prompt further refinements. Reporting to stakeholders also provides the opportunity to develop more shared knowledge.

The final stage of the evaluative inquiry process is the integration of the solution into the standard of practice. The new knowledge and skills should become part of the
instructional routine when appropriate. Thus, transfer of the professional learning into practice is an inherent part of this model.

The evaluative inquiry process is an iterative process. It is not an occasional, pull-out event. Therefore, as soon as the professional learning community has satisfactorily solved one problem, it should start the same process over by addressing another emergent, ill-structured problem.

Questions for Further Research

This study resulted in a number of findings about the process for developing a culture of evaluative inquiry. As part of the case study methodology, a number of interrelated factors were described which, over a two year time period, influenced the shift in organizational culture. The impact of particular activities was not able to be isolated and causality was not able to be determined. Further research is necessary to explore the effect of certain processes on the shift to a culture of evaluative inquiry. The following questions might provide valuable information:

- Given a context where there are few funds and little time, which Evaluation Capacity Building activities are most effective?
- How can Evaluation Capacity Building activities best be structured to foster the development of interpersonal trust?
- To what extent does process use contribute to building the capacity for long-term organizational learning and evaluative inquiry capacity? How can this capacity building be sustained over the long-term so that a culture of evaluative inquiry becomes freestanding?
• In a system of defined autonomy, which elements need to be non-negotiable in order to ensure the attainment of essential skills and concepts and which elements can be determined by PLC teams in order to foster ownership?

• Which district, state, or federal policies positively impact the development of a culture of evaluative inquiry? Which policies act as impediments to the development of a culture of evaluative inquiry?

Conclusion

This study was designed to explore the processes used by staff members at North Mesa Elementary School to develop a culture of evaluative inquiry. While instances of significant learning and changes of practice were evident, significant challenges were encountered and the staff was not successful in fully developing and sustaining a culture of evaluative inquiry. A lack of resources including funding and time hampered explicit Evaluation Capacity Building activities; however, the requirement to monitor student progress towards grade-level goals and collaboratively plan grade-level wide weekly interventions led to significant process use. The study raised questions about how school leaders can create a system of defined autonomy that balances top-down direction with bottom-up empowerment and how best to meet staff members’ psychological safety needs as they start shifting towards a culture of evaluative inquiry.

While this study provided a rich description of the challenges that a typical school faced while endeavoring to create a culture of evaluative inquiry, further research is needed to isolate specific activities that effectively lead to Evaluation Capacity Building and process use. Further research is also needed to determine how Evaluation Capacity Building and process use can lead to a sustained shift in organizational culture. As these questions are
explored, the relationships between program evaluation, Evaluation Capacity Building, process use and organizational learning will be examined and will inform future evaluation practice in the context of schools.
References


Demarest, L. (1997). *Looking at Type in the workplace*. Gainesville, FL: Center for Applications of Psychological Type, Inc.


Appendices

Appendix A: Consent Form ........................................................................................................ 168

Appendix B: The Case .............................................................................................................. 171
Appendix A: Consent Form

University of New Mexico

Consent to Participate in Research

Introduction
You are invited to participate in a dissertation research study conducted by Happy Miller, doctoral student, from the Educational Leadership and Organizational Learning department at the University of New Mexico. The data from this study could be combined with data from previous studies in a future publication. You have been selected to be a part of this study because of your participation in or support of the Instructional Leadership Team and/or Professional Learning Committees at North Mesa Elementary School.

Purpose of the study
The purpose of this study is to understand the factors that affect the process of developing a culture of evaluative inquiry at a public elementary school. Particular focus will be placed on the role of the leadership team in developing evaluation capacity and fostering ongoing, contextualized learning.

Procedures and activities
I will observe you during evaluation capacity building activities, leadership team meetings, and professional learning community meetings. I will take notes either in a notebook or on a laptop in order to describe the context and I may digitally record the conversation. I will then transcribe the conversations to look for emergent themes. I may ask you to periodically volunteer to participate in a 20-60 minute interview to share your perspective of the evaluation capacity building organizational culture building processes. These interviews will take place throughout the year. I may ask to review data that teams have put together into a school portfolio. You will not receive payment for your participation.

Potential risks and discomforts
The observations are naturalistic in nature; there is no effort to change your behavior. The risks are considered to be minimal. That means that “the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests”.

Potential benefits to participants and/or to society
Over the past decade, the concept of teacher collaboration in professional learning communities has gained great support as a means for creating organizational change. Through collaboration, teachers are expected to both improve practice and generate knowledge. Some researchers suggest that the most appropriate approach to evaluation for this context would be one in which evaluation is ongoing and iterative, integrated into regular work routines, and is performed primarily by organization members. The approach should develop a culture of evaluative inquiry in order to promote organizational learning. Interest
in Evaluation Capacity Building and the development of cultures of inquiry has grown in recent years. This case study would provide data about the process of developing a culture of evaluative inquiry and some of the challenges encountered. The results would be of interest not only to the larger educational community but also to the Suburban School District since the goal is for North Mesa Elementary School to serve as a model for other schools to undergo a similar process in future years.

Confidentiality
Any information obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. Digital recordings will be immediately uploaded onto a separate USB thumb drive and will be kept in a locked safebox at my house. Your information will be protected by a firewall on the computer during the transcription process. Field notes will also be kept in the locked safebox.

You will be referred to only by a code in the transcription and by a pseudonym in the write-up of the study. Anyone to whom you refer in a conversation will also be given a code and pseudonym. The school also will be given a pseudonym and the district will not be identified.

The transcriptions will only be used for educational purposes. The data will inform future studies and may be included in a future dissertation publication. All data will be destroyed upon the completion of my dissertation and related publications.

Participation and withdrawal
You can choose whether to participate in this study or not. If you volunteer to participate, you may withdraw at any time. You may also refuse to answer any questions you do not want to answer and still remain in the study.

Identification of investigators and review board
If you have any questions or concerns about the research, please feel free to contact: Happy Miller, researcher, at hmiller@unm.edu or Dr. Patricia Boverie, chair of the Department of Educational Leadership and Organizational Learning, at pboverie@unm.edu. If you have other concerns or complaints, contact the Human Research Protections Office at the University of New Mexico, phone (505) 272-1129.

Signature of Research Participant
I understand the procedures described above. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been provided a copy of this form.

Name of participant (please print)

Signature of participant Date
<table>
<thead>
<tr>
<th><strong>Signature of Investigator</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>In my judgment the participant is voluntarily and knowingly providing informed consent and possesses the legal capacity to give informed consent to participate in this research study.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of investigator</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Signature of investigator</th>
<th>Date</th>
</tr>
</thead>
</table>
Appendix B: The Case

As discussed in Chapter III, the methodology selected for this empirical study is a case study. In order to provide “interpretation in context” (Cronbach, 1975), this appendix presents the case primarily in a chronological order from the beginning of the study in July of 2008 to the end of the study in May of 2010. Studies of several units of analysis are presented within this chapter: the school as a whole, the Instructional Leadership Team, four grade-level PLC teams, and the bilingual program PLC team. This data is summarized in a linear analytic manner in Chapter IV.

The Case

The context

In 2007, Sharon⁶ was named Principal of North Mesa Elementary School where she had been a classroom teacher before becoming an Assistant Principal at a nearby school. She described her first year as principal when there were 1,250 students enrolled at the elementary school,

When I became principal it was just hold on for dear life. I didn’t want to push things too much because everyone was in such flux. It was basically just making sure that everybody had a space, materials in place, everybody had a chair. It was that kind of modality. (6-15-10 p. 1)

The following year, the Suburban School District opened two new elementary schools and attendance boundaries were redrawn. Enrollment at the school dropped from 1250 to 750 and many teachers transferred to the new schools. While there was relief to escape from the

---

⁶ Pseudonyms have been used in order to protect the anonymity of all of the participants.
overcrowded conditions, the changes also had an emotional impact. One teacher made this comparison a few months after the attendance boundaries changed,

Our school has literally gone through a divorce. It may have been an amicable divorce but we have gone from a 13 member grade level to 7. (10-23-08 p. 2)

Given the structural changes and the resulting changes to the status quo, Sharon felt that the time was right to promote a cultural change within the school.

The school had long had a tradition of teacher autonomy. Sharon described what the culture was like when she first started as a teacher at the school,

I remember coming here in 1998 and it was like a free-for-all here. Honest to goodness, you did whatever you wanted to do pretty much. No one had the same kind of materials. No one had the same standards. (6-15-10 p. 15)

As the district grew in size, greater efforts were made to align curriculum, instruction, and assessment. However, the size of North Mesa Elementary School made alignment efforts challenging. As one teacher pointed out,

We were all having to deal with such a large number just within the grade level we weren’t thinking [about] what was going on across grade levels. (10-23-08 p. 4)

With thirteen teachers at a grade level, there were few efforts to create alignment within grade levels as well. Most grade-level meetings focused on business items, not on instructional practices.

Another factor that prompted Sharon to push for a cultural change was that the school did not meet Adequate Yearly Progress under the ‘No Child Left Behind’ Act for the first time in 2008.
Not only had the target that the school needed to meet, the Annual Measurable Objective, increased significantly from 2006 to 2008 but the actual percentage of students with disabilities achieving proficiency in reading had declined from 37% in 2006 to 20.8% in 2008. The percentage of ELL students who were proficient in reading had also declined significantly but the overall percentage was still high enough to meet the target for that year.

### Table B-1: Adequate Yearly Progress Trends at North Mesa Elementary School, 2006-2008

<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2006</td>
<td>2007</td>
</tr>
<tr>
<td>All Students</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Caucasian</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hispanic</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Asian</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>American Indian</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>English Language Learners</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Students with Disabilities</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### Table B-2: Percentage of Students Proficient on the SBA, 2006-2008

<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2006</td>
<td>2007</td>
</tr>
<tr>
<td>All Students</td>
<td>67.9</td>
<td>65.3</td>
</tr>
<tr>
<td>Caucasian</td>
<td>75.5</td>
<td>71.6</td>
</tr>
<tr>
<td>African-American</td>
<td>59.1</td>
<td>61.0</td>
</tr>
<tr>
<td>Hispanic</td>
<td>59.1</td>
<td>61.0</td>
</tr>
<tr>
<td>Asian</td>
<td>61.3</td>
<td>43.9</td>
</tr>
<tr>
<td>American Indian</td>
<td>61.3</td>
<td>43.9</td>
</tr>
<tr>
<td>English Language Learners</td>
<td>37.0</td>
<td>32.1</td>
</tr>
<tr>
<td>Students with Disabilities</td>
<td>37.0</td>
<td>32.1</td>
</tr>
<tr>
<td>Economically Disadvantaged</td>
<td>62.5</td>
<td>59.6</td>
</tr>
</tbody>
</table>
Sharon explained part of the pressure in a meeting with the Instructional Leadership Team,

> We have to provide systematic interventions for every child at this school on their level. We have to. If we do not do this, the state will step in and do it for us… This is the thing that I want you to realize. This is not me sitting there and going, ‘Oh, this would be lots of fun, you know, just to really mess up everybody and their autonomy.’ So that’s what you need to realize number one. Because, in fact, I have thought about… inviting [another principal] to come over and talk to you guys because [he] can share with you what happened in [his previous district] when [the state] stepped in. (3-3-10 p. 6)

While there was tremendous pressure from the community for schools to meet the requirements of Adequate Yearly Progress, the requirements and focus were set by external stakeholders, not the staff members themselves. Many teachers believed that the exclusive focus on reading and math as measured through standardized tests unnecessarily narrowed the curriculum and dismissed the needs of the whole child. One teacher, Leslie, expressed her fears regarding the potential negative impact of the ‘No Child Left Behind’ Act of 2001,

> There’s a lot of pressure to meet AYP…We’re too focused on AYP versus meeting students’ needs…I don’t know what they are actually doing with those kids that are not meeting the SMART goals and so forth. Are they just pounding it down their throat or are they also using games and fun activities to reach that learner that didn’t get it the first time through the traditional way? To me that’s when you have to get really creative with those students and those are the ones who really need differentiated, not drill-and-kill, activities that might actually reach them through a different modality, through music, art, or dramatic play. (10-14-08 pp. 4 & 8)

Many teachers also believed that the AYP designation did not accurately reflect the quality of their school. Since students in every subgroup except Students with Disabilities and ELL had met the targets each year, some teachers objected to efforts to reform instructional practices across the school. Wendy, a primary grades teacher, explained her view,

> The entire framework of the school was not messed up. It was this one group we were working with [students with disabilities]. It feels like we’re going in with hatchets and chopping everything out instead of saying, ‘Let’s look at this area and see where we can go ahead with that’. I don’t understand that. (10-23-08 p. 11)
Even at the beginning of this case study, Wendy expressed frustration with efforts from the
district to raise expectations and develop a common, guaranteed viable curriculum,

It’s time to quit ramming everything down everybody’s throat. Let us do the things
that we want to do and I think that will make a huge difference…We don’t need the
shotgun approach of ‘Some of them aren’t doing it, so let’s blast it’. (10-23-08 pp. 9
& 10)

Later that year, Wendy described some of the history when discussing the culture
surrounding collaboration at a meeting,

This group has been together for a long time and there was this really big change in
the district of push, push, push…You will get them to this…So when this began years
ago to [raise] the bar and the expectations, there was this massive push, on primary to
get them from A to 24. ‘Get your butt together and get it going’ and so it was this
massive, ‘Oh my god, what am I doing?’ And there was this frenzy. And I think
there’s still some of that in there and now suddenly it is like, ‘Oh, will you guys
please work together now’? (2-29-09b p. 17)

Thus, there was already some frustration among teachers regarding perceived loss of teacher
autonomy from federal, state, and district mandates before Sharon decided to try to develop a
culture of evaluative inquiry.

*Initial steps towards implementing the initiative*

The first step in this initiative was providing professional development to a guiding
coalition of staff members. In July of 2008, a team of seven school staff members and three
district staff members attended a data portfolio conference sponsored by the Education for
the Future Initiative. The team from the school included

- Sharon, the principal;
- Wanda, the assistant principal at that time;
- Wendy, a primary grades teacher;
- Leslie, the bilingual teacher;
- Karen, the ELL teacher;
- Cheryl, the instructional coach; and
- Barbara, the educational technology specialist.
The team from the district included

- Myself in the capacity as the district’s data analyst;
- Maria, the Director of Federal and Bilingual Programs; and
- Elena, the Director of Professional Development.

The conference was led by Dr. Victoria Bernhardt, a prolific writer about translating data into information to improve teaching and learning. The objectives from the training included learning how to utilize data to improve continuous improvement planning and how to analyze data to identify which factors contributed to the improvement of or the stagnation of learning. They also included learning how to create a vision and action plan, how to close the achievement gap, and how to institutionalize data analysis.

The purpose of the training was to provide the school with a methodology for organizing and analyzing demographic, achievement, process, and perception data. However, the training also influenced the attitudes of some of the team members towards data. Leslie explained how the conference addressed some of her fears about the narrowing of focus through the AYP requirements,

I kind of was like, ‘I don’t like data’, and what I learned is that philosophically I am not anti-data; it just has to be used appropriately and not [be] single-focused [like on the state accountability assessment]…One of the things that I felt positive about coming out of that conference was that data is not only test scores. (10-14-08)

Cheryl concurred that the training had a similar effect on her:

I think that I learned to appreciate data a little bit more. I’ve never been much of a data person. I mean we collect it but we really don’t intensely look at it…I think the thing I liked most about that conference was the fact that they didn’t lose sight of the child. (11-24-08 p. 1)

The training also made some of the participants more reflective regarding the change process. Wendy, for example, said the following when asked what she learned from the training,
I think the predominant thing that I learned is that there needs to be time to have that common idea of where do we want to go and where do we want to be…Nothing else is going to happen unless you have that commonality. (10-23-08 p. 1)

Thus, in addition to providing practical tools, the training also influenced beliefs and gave participants an overview of how to manage a change process.

The conference was meant to be a “kick-off” for the initiative to develop a culture of evaluative inquiry. In preparation for the conference, Wanda, the assistant principal, did complete a draft data profile for the school. Upon the opening of the school year, the guiding coalition attempted to incorporate some of the processes into their school improvement efforts. During the staff orientation, staff members were asked to conduct a self-assessment by rating the school on a variety of continuum indicators.

There were seven main continuum indicators:

- Information and analysis;
- Student achievement;
- Quality planning;
- Professional learning;
- Leadership;
- Partnership development; and
- Continuous improvement and evaluation.

Each of these indicators was further subdivided into three subsections: approach, implementation, and outcome, and was described at five levels of implementation. Here is an example of the continuum description for Information and Analysis Implementation:
No information is gathered with which to make district or school changes. Student dissatisfaction with the learning process is seen as an irritation, not a need for improvement.

Some data are tracked, such as attendance, enrollment, and drop-out rates. Only a few individuals are asked for feedback about areas of schooling and district operations.

The district collects information on current and former students (e.g., student achievement and perceptions), analyzes and uses it in conjunction with future trends for planning. Identified areas for improvement are tracked over time.

Data are used to provide feedback to improve the effectiveness of teaching strategies on all student learning. Schools’ historical performances are graphed and utilized for diagnosis by the district.

Innovative teaching processes that meet the needs of students are implemented across the district. Information is analyzed and used to prevent student failure. Root causes are known through analyses. Problems are prevented through the use of data.

Posters of each of the continuums were posted on the walls of the cafeteria and in the hallway. Staff members were encouraged to read the continuums and place a dot on the rating they perceived to be most appropriate for each indicator. Unfortunately, due to limited resources, staff members were not provided with individual copies of the continuums. The dots covered part of the written descriptions for each indicator so that not all staff members were able to read the descriptions before placing a dot. Given these limitations, the data was considered to be invalid and was not analyzed.

Grade level teams also started the process of developing collaboration norms during orientation day. Teams were given a handout with guiding questions in six categories: time, listening, confidentiality, decision making, participation, and expectations. The principal, Sharon, collected the notes created by each of the grade level teams on that day and posted them on the wall for several weeks. She asked staff members to review them and place dots next to the ones that they considered to be most important in order to determine the schoolwide collaboration norms. Sharon described what happened next,
That didn’t work. They said they were too busy. So I wrote down a list of common things and gave them to the grade levels and said, ‘Just tally. Tell me your top five’. Well, the top five weren’t even close to each other. Wanda [the assistant principal at the time] said, ‘Oh, just tell them which ones’ and I said ‘No, we’re not going to do that. We can’t do that. I mean because then it’s going to become a piece of paper that’s just a piece of paper’. (9-30-08 p. 14)

While there was not schoolwide agreement on norms, one of the teachers did feel optimistic with the collaboration norms that were developed at her grade level,

While we have had [developed norms] in the past, it is still ‘Here’s our list and let’s go on from there’. I think that it’s a smaller group now and that people are feeling more comfortable talking and listening. I think that people are more open because it’s a smaller group, a smaller school. (10-23-08 p. 3)

Hence, teacher interactions may have changed somewhat from the structural changes and somewhat from the norm developing activities, but a great deal of diversity still existed in expectations and norms for teacher interactions.

Another initial step in the implementation of the initiative was the administration of staff and student surveys in September of 2008. The surveys were created by Education for the Future staff members in 1991 to assess perceptions. Items were drafted based on the literature about effective schools and revised after years of input (Bernhardt, 2008). The reliability quotient was .86 for the staff questionnaire and .93 for the elementary school student survey. The staff questionnaire responses “show the degree to which staff work together to create a continuum of learning for all students, if there is a clear and shared vision, and if the administrators are adding value to school processes from the perspective of the staff” (Bernhardt, 2008). Each of the statements in the teacher survey was framed in a positive manner. A total of ninety-one staff members took the survey. All of the mean responses ranged between neutral and strongly agree. The statements with which respondents most strongly agreed were:
• I feel that learning can be fun;
• I believe student achievement can increase through differentiating instruction;
• I believe student achievement can increase through providing a threat-free environment;
• I believe student achievement can increase through close personal relationships between students and teachers;
• I believe student achievement can increase through addressing student learning styles;
• I love seeing the results of my work with students;
• I believe every student can learn;
• I believe quality work is expected of me;
• I believe it is important to communicate often with parents;
• Learning is fun in my classroom; and
• I love to teach.

The statements with which the respondents agreed the least were:

• I feel recognized for good work;
• I feel that others are clear about what my job is at this school;
• My administrators facilitate communication effectively;
• My administrators support shared decision making;
• I believe the vision for this school is shared;
• I believe we have an action plan in place which can get us to our vision;
• Morale is high on the part of teachers; and
• Teachers in this school communicate with each other to make student learning consistent across grades.

In general, it appeared that those areas which were most under the control of the teachers were rated most positively by those teachers but the areas which required collaboration and support from others were rated the lowest.

Interviews reinforced some of the themes that emerged from the staff surveys. When asked what the school did well, one teacher mentioned “traditions like…our sock hop and library dress up week, those little things that kids look forward to…and student connections to their teachers.” (10-14-08 p. 14)

The results of the staff survey were disaggregated in several different ways. In general, staff members with one to three (N=17) or four to six (N=15) years of experience rated the statements the most positively. Interestingly, staff members with eleven or more years of experience (N=37) rated the statements more positively than staff members with seven to ten years of experience (N=12). There did not appear to be significant differences in the responses of the staff members based on ethnicity (Caucasian N=49; Hispanic/Latino N=30).

The effect of the economy on the implementation of the initiative

One external factor that affected staff morale was a downturn in the economy at the beginning of this study. The school district was forced to implement a variety of budget saving measures including restricting supply and substitute teacher budgets. For the first time in the district’s history, the district needed to contemplate the possibility of furloughs or
a Reduction In Force (RIF). Salaries were frozen while benefits were reduced. Cheryl explained part of the impact,

People are worried about jobs. People are worried about the economy. As good as these teachers are, morale is bad...There are a lot of [people] who are mad...I think that’s a big problem because as much as [we] can try to put on that happy face and go in there, those kids will know. And it’s going to transfer over. (11-24-08 pp. 13 & 15)

Barbara also described the impact of the economic downturn and budget cuts on staff morale,

I think that the economy is pulling people down. I think that if everything were rosy we could push them a little harder. People are so low right now because they are thinking, ‘Am I going to lose my job? Am I going to lose my job?’...Are we running out of toner in the printers? You know, all that stuff even if they are not voicing it out loud I know that it is at the back of their minds. Because some people that normally would be pleasant aren’t. And then the people that you know are not going to be pleasant are worse. (11-24-08 p. 5)

While the school district was able to stabilize the budget without laying off teachers during the time frame of this study, significant budget cuts had to be instituted and the ongoing economic downturn led to more budget cuts and the possibility that furloughs or layoffs would become inevitable. Given all the uncertainty of the larger economic context, teachers were hesitant to enter into the necessary ambiguity of a change initiative as well.

The economy also impacted student behavior and readiness to learn. The community of North Mesa was one of the least affluent in the Suburban School district with some of the lowest house prices. Sharon described the community,

When I came here people said [the community] was kind of like the Stepford Wives. The community looks really kind of cute on the outside but when you go in, it isn’t.” [School] is not on [the students’] radar. Surviving is. And I think that’s kind of the population that we have now to a pretty large extent. (6-15-10 p. 25)

When the new schools were opened and the attendance boundaries were redrawn, the newer subdivisions with more expensive houses were assigned to the new school. The percentage
of students participating in the Free and Reduced Lunch Program increased by about 10% at that time. However, the percentage of students participating increased by an additional 14% since the re-zoning process which reflected the impact of the economy on the financial security of the families.

In 2010, North Mesa Elementary School had the highest percentage of students participating in the Free and Reduced Lunch Program of any of the schools in the district. The impact of the shifting demographics wasn’t as apparent in the first year of the study as in the second year. Sharon noted,

[The nature of students] has shifted. We expected it last year…but last year I didn’t notice it as much. This year it is very apparent. I don’t know whether it is indicative of financial crises that are going on with people. There seems to be a whole lot more of that. Our free and reduced population did shoot up again. They’re falling off the edge…We saw a whole lot more CYFD [Children, Youth, & Family Division] referrals this year. Families just need more support. (6-15-10 pp. 23 & 24)

Wendy, a primary grades teacher, described the impact that the home environment had on the students’ readiness to learn,

This is the rental area. People can’t necessarily afford them [but] that’s what they have available to them. You have a lot of two people working to try to afford housing and so they are testy and snapping and don’t have time for kids…The parents aren’t talking to them a lot. Even things like basic colors and stuff. (5-17-10 pp. 1-2)

The school tried to meet the increased need of the students by starting two different programs. The first program offered free breakfast to all students. The students ate in the classroom at the beginning of the school day. Wendy described how the program impacted her classroom,

A lot of parents don’t feed the kids. I had six kids regularly come to school without breakfast. [Now] they are showing up and I can feed them breakfast without it having to come from my own pocket. (5-17-10 p. 2)
Prior to the free breakfast program, Wendy provided cereal or toast to students at her own expense. Sharon also described the impact of the program on the school as a whole,

The needs we would see – children hiding in bushes, children taking 2 or 3 or 4 apples so they could take them home…Teachers have liked the free breakfast program. Attendance has picked up. They are doing better in school. (6-15-10 p. 24)

In addition to the free breakfast program, several teachers at the school organized a food backpack program. The teachers solicited donations of canned foods and other non-perishables and filled backpacks with food for certain students to take home for the weekend. Sharon explained the need,

One of the best things that we have done this year was that backpack program…There are families that we have got that didn’t eat from Friday until Monday…At the end of the year there were forty families on that backpack program. We got a grant through the Roadrunner [Food Bank] for 100 families in the beginning of the year. (6-15-10 p. 24)

Thus, personal security needs dominated the lives of many teachers and students over the course of this study.

*The school principal’s vision for the initiative*

It is into this context that Sharon tried to develop greater accountability for the use of data. For Sharon, closely following the progress of students was a moral imperative. Sharon described her vision of a very student-centered school:

I just know in my head that we’ve got to do what we need to do for kids. And if the scores don’t show it, then we need to think about what we need to do next…Every day, every child, no excuses. We want these children supported in every single way possible they can be supported. We’re past the days where you come to school, you learn with papers, and then you go home. These children here deserve the support not only academically but emotionally and socially. (6-5-09 p. 9 & 6-15-10 p. 1)
Sharon gave an example of how she expected teachers to know students’ needs and adapt
their instructional practices to meet the students’ needs,

They were over at [the preschool] and they were talking about autism boards and the
things that they use. [The teacher] said, ‘I don’t do that’. So I said, ‘if something
works for a child, then that’s what we do’…It doesn’t have anything to do [with the
teacher]…[The teachers] may have to shift [their] thinking…It’s the ‘I’m in control of
this. This is what I do and it works for me’. (6-5-09)

In order to avoid the idiosyncratic responses of individual teachers, Sharon created the
expectation that grade-level Professional Learning Community (PLC) teams would
collaboratively set learning goals, create biweekly common assessments to measure student
progress towards those goals, analyze the data, and, in the second year of the study, provide
systematic interventions to help students. The structure of PLC teams had been established
in prior years by the district and the process of setting goals and measuring them was part of
the long-standing district-mandated Plan-Do-Study-Act continuous improvement cycle.
Sharon described her expectations for the PDSA continuous improvement process at an
Instructional Leadership Team meeting,

You are going to have to design…a very quick and easy and simple assessment.
What does it look like that’s not going to take you years to do?…That’s how you are
going to progress monitor. You are going to say, ‘We are working on number sense.
This is what we are working on. This is how we are tracking it’…The study comes at
the end of the first quarter. That’s when you look at your short cycle [assessments]
and go, ‘Okay, how did that go? How many are proficient? What are we going to do
for next quarter?’ (9-2-09 pp. 12 & 20)

A handout from the fall of 2008 entitled “North Mesa Elementary School Leadership
Structure” further elaborated on the desired role of the professional learning communities,
The purpose of grade-level PLC’s is to use data to drive instruction. The PLC’s will
promote curriculum coherence as well as develop instructional strategies and
common assessment to support student achievement at each grade level. Grade-level
PLC’s will meet in the cafeteria, library, and/or computer lab.
• Every teacher will participate in meetings with her or his grade level.

• Teachers will coach and support the implementation of the standards and the vision in each other’s classrooms.

• Grade-level teams will seek support from the subject-area teams.

• Teachers will study and support each other’s implementation of best practices.

• Support staff will be assigned to appropriate grade level teams.

• Collect and disseminate classroom data to improve classroom instruction and overall student achievement.

While both the structure of professional learning communities and the structure of the PDSA process had been in place, the organizational culture did not support the most effective use of these. Sharon’s goal was to create a cultural shift whereby the use of these existing structures would move beyond pro forma compliance to create a significant difference in instructional practices.

*Use of data prior to the initiative*

The school had long used some data to identify students who needed additional support. For example, the school had created “an intervention wall” in the Instructional Coach’s office. The wall consisted of pocket charts with color-coded cards. On the front of the card were a student’s Developmental Reading Assessment (DRA) history and a list of the interventions that the student had participated in previously like Reading Recovery or Literacy Group. Identifying information about the student was hidden on the back side of the card. In that manner, the Instructional Coach, Cheryl, could quickly and easily see which
students were receiving interventions, how many students needed services but were not yet receiving the needed assistance, and each student’s growth across the school year. Sharon described the purpose of the wall,

We’ve always used the intervention wall. [The instructional coach] would update it [quarterly] and we’d look at it and the team would look at it to see who had been served…It is a good, visual, consistent way to make sure children are being served and are being supported. (6-5-09 p. 20)

The data from the wall had also been used to establish the need for new or different types of pull-out interventions or programs for certain students. During an Instructional Leadership Team (ILT) meeting, Cheryl, the Instructional Coach, used the wall to illustrate the lack of progress of some ELL students to ILT members,

These kids are not in intervention right now but I want to look at them and see if I can get them in. Because according to these cards, at the end of 4th quarter in first grade they were a twenty. Baseline coming in they were a twenty. Now after nine weeks in second grade they are still a twenty. They are sticking. (11-4-09 p. 4)

Thus, a precedent for using data for making placement decisions and for monitoring growth had been established.

However, many assessments had frequently been given just to comply with district mandates and many staff members did not have the time, knowledge, or motivation to use that data to track student progress and alter their instructional practice. Cheryl, the Instructional Coach who managed the intervention wall, described schoolwide use of data prior to the case study,

We really just collected [data]. We’re not like some of the other schools where it drives everything and every decision is based on that data. Data was kind of like a second thought. And maybe we weren’t really good at collecting it either. (11-24-08 p. 2)

Diane, the third grade chair, mentioned,
We have a lot of assessments but it becomes a compliance issue. ‘Yes, I gave it. Here’s my scores. Now can I go back in and teach’? (2-24-09c p. 5)

Natasha, the math interventionist, concurred that,

[Some teachers] are just like ‘Well, I have been doing it this way for years. You know, ‘You want my numbers? I will give them to you but then I am going to move on to the next chapter’. It’s not ‘Well, these kids are really solid in this area so I can move them on or give them more advanced type of problems in that area. (2-22-10 p. 6)

A lack of consistent training in how to interpret the assessment results was also evident.

Charmaine, the Special Education Site Specialist, stated,

[The teachers] bring their NWEA scores...and can’t tell me anything other than this is the number. They don’t know what the number means. They have no idea…The kids don’t know what [the NWEA scores] mean either. Shouldn’t they at some point understand what that means for them? (2-24-09c pp. 4 & 5)

Hence, while data had been collected, it appeared that there had not been widespread systematic analysis of the data to make it useful in improving instructional practices.

**The use of data in PLCs during the Fall Semester of 2008**

In the fall of 2008 Sharon and Wanda raised their expectations regarding the use of data in teachers’ instructional practice and the PLC teams’ PDSA continuous improvement plans. They told teams, “We want you using common language and common assessments because when you look at data then it’s valid.” (9-30-08 p. 12) Sharon described the reaction she received,

I think we kind of hit a discomfort level with [the teachers] because in the past they would do their little PDSA and they would hand it in. They would give you tidy little data at the end of the quarter and it would look really good. It would sit in a drawer and that was basically it. So now we’re going, ‘Rank your areas of growth’. And we want semi-monthly assessments. They haven’t been very fond of that because they don’t understand. I said to one of the teachers, ‘It would be like if your baby was sick and you took it to the doctor in October but didn’t check again until December.
Would you do that with your own child?’ She kind of looked at me and I said, ‘That’s the same with our data. If we’re going to make changes and we’re going to see if this area that pinpointed needs modifications, we’ve got to look at it constantly’. (9-30-08 p. 5)

Many teachers, including Leslie and Karen, described a shift in the way that PLCs operated that fall,

   PLCs are definitely more data driven. They seem to be more productive. They are not just chitchatting about things. Everything seems to be ‘What are our numbers telling us? What do we need to do? Where do we need to improve? (10-14-08 p. 1)

Cheryl also noted some changes in the conversations held within PLCs,

   I think that there are some good conversations going on. I think some teams are really pulling together and it has caused them to have conversations that really do affect student achievement and their teaching and their instruction. (11-24-08 p. 4)

Thus, there was some hope that a shift towards a culture with greater embedded evaluative inquiry was actually taking place in the fall of 2008.

*Schoollwide challenge: Structures that didn’t match readiness levels of teams*

Many challenges existed, however, in institutionalizing the shift to a culture of evaluative inquiry. The first challenge was that the principal’s vision of a culture of evaluative inquiry was not yet a shared vision among the staff. Barbara, the educational technology specialist, stated in November,

   The whole PLC concept is so different. Instead of just getting together and discussing maybe random issues that come up, they are very focused. [The] agenda is really set…It just comes down from administration. ‘This is how we’re doing things and here’s the form that we’re going to record everything on’…The administration is the driving force…There is a handful of resistant people…[who] feel like they are being browbeaten. (11-24-08 p. 1)
The form that had been distributed for PLC notes actually had been distributed by the district’s Executive Director of Elementary Curriculum & Instruction to schools district-wide to help focus the work of professional learning communities on four main questions:

- *What is it we want all students to learn?* (Identify standards, benchmarks, power standards, or goals and objectives)

- *How will we know when each student has mastered the essential knowledge or skills?* (What does proficient look like? What does the data tell us? Who is not proficient? Who is advanced?)

- *How will we respond when a student experiences difficulty in learning?* (Differentiation, Instructional Strategies, Tier I & II Interventions, resources, tools – What are the main challenges to better results? How do we hold the students accountable for their learning? How do we hold each other accountable?)

- *How will we deepen the learning for students who have already mastered the essential knowledge and skills?* (Enrichment, differentiation, instructional strategies, resources, tools – What are the main challenges to better results? How do we hold students accountable for their learning?)

The form also had boxes for discussion notes and steps with which to follow up as well as space to communicate questions to the school administrators or site specialist. While comprehensive and research-based, many PLC teams found that the form did not work well for them. One of the teachers explained,

> It didn’t match the conversations that we were having and so we had to try to take the conversations that we were having and fit it into the form. I don’t think that was really reflective of what was necessarily happening. (5-28-09 p. 11)
Sharon elaborated on the feedback that she received regarding the form,

The feedback that I got was that it needed to be a little bit more simplistic…It was just too much for them to handle at the implementation stage that we are at now. (5-28-09 p. 1)

The form was revised at the end of the first year of this study. During the second year, at least one grade level utilized a format of “Check in (announcements), Review agenda and norms, Discussion, and Next Steps”. The “four essential questions” were listed in the footer of every page as a reminder to teams.

Schoolwide challenge: Lack of shared vision

Wendy, a teacher who had attended the training sponsored by Education for the Future, also discussed the progress in developing a shared vision in the fall of 2008,

We still don’t have a vision. We did some of those beginning activities which were good but then it kind of felt like it dropped and we never created a vision…As an IL [Instructional Leadership] Team we haven’t even created a vision. That’s the part that I’m a little disappointed about, I guess, because to me I thought that that was one of our first goals. If you don’t know where you are going, how are you going to get there whether you are using data or not? So for me, right now, all this data is all about is AYP…I think it needs to be a lot more than AYP. (10-14-08 pp. 5 & 8)

At the end of the first year of this study, Sharon acknowledged that it had been difficult to create a shared vision,

Different people have visions but not everyone has the same vision. I think that’s something we have got to work on. (6-5-09 p. 24)

Sharon explained some of challenges in developing a shared vision,

I think that Wanda and I have a road map of where we want to go but we can’t pull people into it too fast because then I think we’ve lost the buy-in. We’ve lost the empowerment and it just become rote and I don’t want it to become rote with them…You know how you do things because you know it’s the right thing to do, but you don’t want to do it? In some instances, I think that’s where we are right now. They know that this is the right thing to do but sometimes they just don’t want to do
it. I don’t want to give them a piece of paper that says ‘Here’s our common vision’ when we’re not there yet. (9-30-08 pp. 11 & 14)

She further elaborated later that year,

I think we’ve tried to [develop a shared vision]. We did it some at the beginning of the year…It’s hard to do at the beginning of the year because teachers are sitting there going, ‘I’ve got to do this in my room. I’ve got to do that in the room’. You know, ‘They’re coming in two days’. Their heads aren’t there so I don’t know what kind of buy-in you have but you’ve also got to do it at the beginning of the year because it sets the tone. (6-5-09 p. 25)

The process of developing a vision that was truly shared by all staff members was complex because it involved teachers’ personal beliefs. Insufficient paid preparation and professional development time for teachers exacerbated an already challenging process because it was difficult to motivate teachers to engage in the vision setting process over the more immediate need of preparing the learning environment and instruction for the students.

Schoolwide challenge: Perceived lack of time, control, and resources

Sharon and Wanda tried to model the use of data-driven decision making by using student achievement scores to determine schoolwide professional development needs. Sharon described how the administrators tried to guide teachers to pay attention to areas that were relatively weak,

This year on the teachers’ PDPs [Professional Development Plans] we had them do a reading continuum because reading has dropped…and research shows that small group guided reading with Running Records, you know, continuous assessment, helps. So I got a continuum from the instructional coach and when [the teachers] come in to me, I have them highlight where they are on the continuum and we talk about the next step…They’ve had training [on how to use Running Records] for probably three years…We have kind of said they should [use Running Records] but no one has ever really required it. We have to hold their feet to the fire and that’s where those two-week assessments are going to come in. At those meetings we can say ‘You’ve got a struggling reader, can we see your running record’? (9-30-08 pp. 23 & 24)
To provide teachers support with their Professional Development Plan goals, Sharon and Wanda provided professional development using internal resources during an in-service day in October of 2008 on the use of guided reading and Running Records to guide instruction.

Wendy described some of the teachers’ reaction to the training,

They [the school administrators] are very concerned about the reading scores this year. On our last in-service day…they very specifically focused on how guided reading looks at the different grade levels. [Some questions we have are] what about the resources to do this? [When do we have] the time to plan these lessons, get into the leveled library, pull all these books, and do the lessons? What they are doing with the SMART goals takes a lot of time already and then you’re adding what they see as this additional [requirement]. It’s always that issue of when are they going to find the time? This year even more is coming, is being put on them. Every year it’s more and more and more and nothing is ever taken off. (10-14-08 p. 3)

Lack of time and control were recurring themes in the teachers’ comments. One teacher described competing demands on their time at a meeting,

We’ve suddenly got new things happening this year. We’ve got ProgressBook [an electronic grade book] suddenly coming in. We’ve got SAP [a student achievement data warehouse] coming in. And so we’ve had to utilize some of that time too. (2-24-09b p. 21)

While most changes were small and were designed to improve educational outcomes and save teachers’ time in the long-run, each change in the teachers’ routines meant a new learning curve and was perceived as a mandated distraction by many teachers from what they considered to be important.

*Schoolwide challenge: Lack of integration of assessment into instructional practices*

Many of the teachers commented that they also felt that they were mandated to test too much. For example, in October of 2008 Wendy said in an interview,
All we’re doing is testing so I can’t really teach because I’m spending all of this time testing. (10-23-08 p. 7)

The instructional coach also stated in November of 2008,

There is too much assessing going on. The assessments take too long. (11-24-08, p. 2)

In February of 2009, a different teacher stated in a meeting,

One of our major sources of frustration is that we have all these things that we are giving, yet we can’t control what we can take off or add to our plate. And we don’t have the time or the resources to appropriately use that information. (2-24-09c p. 5)

At a retreat in May of 2009, Diane, the third grade chair, stated,

I think the difficulty for teachers comes at that short cycle level [which is] every 2-3 weeks...We just barely get the test results, don’t have any time to teach and then we turn around and test them again. So that’s where we fall down. (5-28-09 p. 7)

Diane further explained how the level of assessment led to a perceived narrowing of the curriculum,

There is a lot of stress because you feel like you are getting ready for this test and then you are getting ready for that test but we’re never really getting to the full curriculum. (9-2-09 p. 4)

Sharon acknowledged the teachers’ comments regarding the level of testing expected,

It’s been the biggest complaint we’ve had this year. They’re testing so much they don’t get to teach. (5-28-09 p. 10)

She also acknowledged that some teachers would rather not focus on assessment results,

“I’ve heard people go, ‘Why are we doing this data crap?’” (9-30-08 p. 21) Thus, it was challenging to teachers to find the right balance of assessment and to seamlessly integrate the assessment into the instructional process.
Schoolwide challenge: Culture of autonomy

The perception of lack of control extended to other areas of teacher practice as well. For example, some teachers felt that collaboration meetings impeded their planning time and autonomy. One teacher stated,

People are feeling that we meet too much. Part of that is because there are so many other things that we want to do and we have lost some of our planning time. (10-23-08 p. 5)

The teachers also resisted the effort to bring all of the teachers into the same general area for PLC meetings to facilitate better communication with resource providers. Sharon explained,

We made the parameters where they are all up here [in the cafeteria] together. Some of them weren’t fond about that. They felt like we didn’t trust them. I said, ‘No, what we’re trying to do is make sure that we’re all in a common area. If you have questions of people, ancillary people, myself, the assistant principal, we’re available. (9-30-08 p. 6)

Some teachers perceived efforts at greater alignment and consistency across grade levels as mistrust in the ability of the individual teachers to know best what their students needed. For some people, the desire to align efforts became intertwined with beliefs about self-efficacy.

Administrative response to teacher pushback

The teachers’ frustration and feeling of lack of control led to pushback from the teachers. Sharon was put in the position of trying to “stay the course” and hold teachers accountable while still allowing for teacher input. She described this as a process of defining what is ‘tight’ (non-negotiable) and what is ‘loose’ (negotiable). By November, she did allow the teachers to meet in their classrooms scattered throughout the school again. Another expectation that she changed was that teachers would need to formatively assess the progress
of students every three weeks instead of every two weeks. One teacher described the impact of this change,

If you just automatically say that you will do [common formative assessments] every two weeks, then you have almost invalidated everything that we’re doing because it is still coming top down…It was a psychological win for the people to say, ‘We do have some control over what we’re doing’. I do appreciate [that the principal] listened to that. (10-23-08 p. 8)

Sharon had to engage in a constant dance to drive teachers towards a common goal while still enabling them to feel empowered.

Professional development with an independent consultant, February of 2009

In February of 2009, Sharon used Title I funds to bring an independent consultant that the team had met at the Data Portfolio conference sponsored by Education for the Future in July of 2008 to the school. The purpose of the visit was to discuss how the school could improve its continuous improvement process in general and the outcomes for Special Education and ELL students specifically. The consultant, Patricia, had been contracted by Education for the Future to present a breakout session entitled “Closing the Achievement Gap” at the training and frequently provided consulting services to Education for the Future clients.

Patricia spent one day with members of the school’s Instructional Leadership Team. Her stated goal was,

To really discover a lot of things in the data…and by the end of the day have an action plan as well as probably a list of assessment questions. (2-24-09a p. 1)
Patricia facilitated a conversation regarding the use of data at each grade level and for specific populations such as Students with Disabilities and English Language Learners. One of her conclusions was,

You are assessment and data rich but analysis poor or data driven decision making poor. You have a lot of things available but it doesn’t funnel through in meaningful ways to your classroom teachers or even at the building level, your decision-making team here, for useful decisions. (2-24-09d p. 1)

In the action plan, Patricia wrote,

There are several assessments given (intermediary assessments) for various reasons, but results are not analyzed systemically and used strategically for improvement:
- Lack of consistent populations in the building has made year-to-year performance difficult to analyze meaningfully for decision making;
- Lack of consistent planning based on performance results for curriculum and instruction at grade levels and in vertical teams similarly limits meaningful decision making.

Teachers are unsure how to use data available from assessment results in instructional planning. (Action plan, p. 1)

She recommended that the school determine which assessments were most beneficial for predictive validity on the state accountability assessment and the most useful for directly informing instructional decisions for teachers. She further recommended that teachers be provided additional training on how to use the assessment results.

During the conversation with the teachers, Patricia also noted that,

Your discussions quickly spiral out to that systems level. You have a lot of things that impact that individual kid very quickly in terms of what his instructional day looks like that have been decided at this systems level. (2-24-09c p. 6)

She elaborated in the action plan,

Systemic issues are overwhelming – creating frustration and effect of teachers feeling like [they are] ‘spinning their wheels’ by investing energy in activities that don’t connect to instructional planning or data driven decision making. Some [is] due to disconnect within [the school] between and among departments and grades. Some [is] due to requirements at district or state level without follow up for how to use [the] information at the building level, especially the classroom level. (Action plan, p. 4)
Given the lack of in-depth understanding of the teachers of the various mandates, it was not surprising that Patricia also observed that the teachers, “Have a lot of data on a lot of different points but don’t know how to correlate”. (2-24-09b p. 9)

*Assessment literacy and assessment quality*

Assessment literacy and assessment quality issues reoccurred throughout conversations. Teachers did comment in meetings about a perceived disconnect between the district-mandated teacher-team created Math Quarterly Assessments which were based on the Power Standards [the most essential concepts] specified to be taught each quarter in the district’s curriculum map. The disconnect may have been exacerbated by several factors. First, the Math Quarterly Assessments were not designed to be part of a growth model. In other words, different standards were measured each quarter. Second, the district allowed schools to use two different textbooks which presented concepts in a different sequence from the other. When teachers followed the sequence of their chosen textbook and not the curriculum map, the sequence of instruction did not match the concepts measured in the Math Quarterly Assessments.

Teachers also struggled to align their PLC team-selected goals with the district-mandated assessments as Anne, the first grade chair, described,

What is being tested on the [Math Quarterly Assessment] is not necessarily our children’s greatest need. So we have been testing every three weeks on the greatest need and then all of a sudden they are thrown into money combinations when what we have been working on is subtraction or regrouping or whatever. And it is not necessarily reflective of what they are necessarily learning in class. (5-28-09 p. 8)
Anne further elaborated on the lack of alignment several months later,

> Our baseline math, all of our assessments are so like kind of random. They are not aligned at all and they are not really what we think would be a good foundation for the kids. Right now our conversation is do we want to give them a strong foundation or have them do well on the Math Quarterly Assessments because they are not aligned at all. (9-2-09 p. 2)

Diane, the third grade chair, added,

> It’s not that they’re not helpful be we can’t use them to track our PDSAs because they don’t track the same thing. (9-2-09 p. 4)

Thus, there was a need to really examine and align curriculum, instruction, and assessment across the district and not just within a grade level. This alignment, however, was often perceived as another external mandate which limited the teachers’ autonomy.

Teachers also struggled to create assessments that accurately measured the concepts taught without overwhelming both the teachers and the students. Sharon, the principal, noted,

> We found out through the year the assessments were sometimes the big factor. The assessment might not give you the information that you needed or give you too much information. (6-15-10 p. 6)

Natasha, a member of the 5th grade PLC, elaborated how her team came to realize that there was a mismatch between their instruction and their own common formative assessments in reading,

> Our reading was not good. We had such a hard time...Each time we didn’t like that assessment so we wanted to give a different kind. And then we didn’t like that one...Now it’s like, ‘You know what? What we’re teaching is not what we’re testing. We need to match the two’...It was a huge A’ha. We’re teaching, teaching, teaching [and] give them a test and nobody scored higher than a sixty. And then somebody said, ‘But look at the test and what we’re doing.’ And it’s like ‘Exactly!’...We need to work on drawing conclusions...and make sure that that’s what we’re trying to test. (2-11-10 p. 5 & 2-22-10 p. 24)
Both training and time were necessary to help teachers create more productive assessments that truly matched their instructional goal.

While quality assessment was frequently cited as one of the biggest challenges and frustrations for teachers, some of the team members came to realize that the process of assessment was important because it communicated what was important. Sharon noted in a meeting of the Instructional Leadership Team in February of 2010,

I thought it was interesting that first grade and fifth grade said they felt a lot of what happened was not so much what you were doing but the assessment of what you were doing...Your assessment drives what you do. (2-11-10 pp. 3 & 5)

By needing to collaborate to create a common assessment, some teachers, perhaps for the first time, began to change the order of instructional planning to goal setting – assessment development – activity development rather than to put together an assessment as an afterthought.

_Tension between autonomy and alignment_

The tension between teacher and team autonomy and schoolwide and districtwide consistency played itself out in the continuous improvement planning process as well. The PDSA (Plan-Do-Study-Act) plans were generated by the grade-level teams. The teams were told to analyze the data to determine strengths and opportunities for improvement, choose a specific area on which to focus, and create grade-level goals. The goals were generated by each grade level independently of the other and independently of a schoolwide plan. While the content levels all created reading and math goals, some grade levels like third and fourth grades also chose to add writing goals while others did not. The process used to generate the
goals also varied by grade level. Sharon described the process used by one grade level she observed to generate their goal,

I remember being in a meeting and they were trying to decide what their goal was going to be for math. They were throwing all kinds of ideas out there. It’s like that 20 idea brainstorming thing…It was the way they had normally thought: ‘Well, this looks good. We kind of have enough data. Let’s go’…The Assistant Principal and I didn’t drive the meetings. We just sat in the meetings because they had to be the ones owning it, not us. If I was in there going ‘This is what you should do’ then it wouldn’t be meaningful to them. It was hard to watch them make those kind of changes and sometimes not. (6-15-10 p. 5)

Thus, there was not a systematic process utilized across all grade levels.

The PDSA planning process was quite dynamic and fluid. When grade level teams reached their goal, then they selected a new goal. However, some of the grade levels evolved through their engagement in the process and decided to change their direction before reaching the first goal. Sharon noted in one Instructional Leadership Team meeting,

Some grade levels have new goals that they made. Some grade levels said, ‘Well, this one wasn’t appropriate.’ But that’s what this process is. It’s not about going, ‘Okay, we’ve got to look here’. It’s about doing what you think is best for kids and making adjustments. (2-11-10 p. 5)

Ideally, the grade-level PDSA plans should have aligned with a schoolwide school improvement plan. Sharon describes the purpose of the state-mandated Educational Plan for Student Success (EPSS) to members of the Instructional Leadership Team,

What an EPSS plan basically does is state what we are going to do this year. We’re going to tell the district and the state what we are doing to improve student achievement. (9-16-09 p. 12)

She explained further,

The EPSS plan really is going to spin off of two areas…It spins off of what your grade level goals are somewhat. It’s a schoolwide plan. We have to look at the areas of need which this year stand out pretty easily, which are ELL population because we had that big dip and then our Special Ed. (9-2-09 p. 1)
Sharon described how her own Professional Development Plan (PDP) related to the schoolwide Educational Plan for Student Success (EPSS) as she requested feedback on the draft EPSS that she had created,

> The fun thing about my PDP is that it actually looks like our EPSS plan to be honest with you. So if you look on the sheet, there’s a sheet in the back that kind of has our AMO [Annual Measurable Objectives] goals and it’s also on my PDP. The reason I am showing this to you is because we need your input too on what happens…Basically my action plan is our action plan. So take a look at it. If you want to make any comments or if you see any concerns, please let me know. (9-16-09 p. 14)

In an effort to respect the teachers’ time, Sharon chose to draft the plan by herself and then ask for feedback. She explained mid-year before an update was due to be submitted to the state,

> I want you guys to be really grateful to me because really what I’m supposed to do at the beginning of the year is you are supposed to be a part of every sheet of this but I knew you would kill me so I did this by myself. (2-11-10 p. 6)

Sharon provided a copy to members of the Instructional Leadership Team and asked them to provide feedback if they so desired at a later time,

> The major thing that we need to look at is what’s been working so far that you think from your vantage point and your grade level and what’s not from our plan…So look it over real quickly and see if you think we’re missing something or something that we need to add or subtract. We’re not going to go through the whole thing today because it will take forever, but I just wanted your input on some stuff. (2-11-10 pp. 7 & 9)

While Sharon’s approach minimized the amount of time that Instructional Leadership Team members needed to spend on debating and drafting schoolwide goals and actions, the siloed approach to continuous improvement contributed to the lack of alignment between initiatives.

One mismatch was between the type of goals in the grade-level PDSA plans and the schoolwide EPSS plan. The grade level plans focused on particular skill areas for all
students whereas the EPSS plan goals focused on the improvement of particular subgroups of students, particularly Students with Disabilities and English Language Learners. Three of the nineteen action steps in the 2008-2009 plan specifically referred to these subpopulations. One strategy was to offer professional development to help parents of English Language Learners support their children. Another strategy said that “ELL and SE students would receive ‘triple dose’ intervention in math and reading if their scores warranted this support”. Finally, the plan called for the creation of a Special Education PLC.

Like the grade level teams, the ELL, bilingual, and special education programs were also characterized by teacher autonomy and a lack of alignment. In an interview, Sharon commented that the ELL support to students in a fifth grade class was disjointed. [The classroom teacher] would be teaching something and [the ESL teacher] would be pulling them for something else so it wasn’t aligned. And that’s the thing that I keep thinking in my head. It’s got to line up and the arrows have got to go the same way. It’s to make sure what we’re doing is not like popcorn. You know, we’re shooting it here and we’re shooting it here and we’re shooting here. (6-5-09 p. 5)

Sharon also described the Spanish bilingual program,

[The bilingual program has been] pretty traditional as long as I have been here, even as a teacher. It has been an enrichment model. There’s not been a lot of vertical alignment. There’s not a lot of collaboration. It’s just there. And you’ve got different philosophies. (6-5-09 p. 1)

One of the teachers in the bilingual program concurred,

There’s no consistency at all. We have got to get on the same page. (11-4-09 p. 5)

The additional financial resources available to the bilingual and special education programs did not necessarily contribute to alignment and consistency. Sharon described how funds were allocated for the bilingual program in 2009,
They’ve got this pot of money…This teacher wants this and this teacher wants that. I asked, ‘What’s the goal here?’…It’s the magic fairy godmother thing. I’ll get you this because you want it. (6-5-09 p. 6)

In 2008 she used very similar language to describe how resources were allocated in the Special Education program prior to 2008,

The Special Education program was all over the place. It was kind of like Tinker Bell. ‘You want this? Sure, we’ll give you this. You want this? Okay, we’ll try this.’ But they weren’t pulling together.

Sharon elaborated,

These Special Ed teachers have got to pull together. They have to get aligned for a common vision. They need time to talk and meet and talk about what’s going on. So we formed a Special Ed PLC and there was a little bit of resistance to that at first. They were like blah, blah, blah, blah and we said ‘No, it’s crucial’…Something happened last year. We not only didn’t meet [AYP], we dropped thirteen points and that alarmed me terribly. I could understand not going up, but why did we drop so far?...I think they need to examine and figure out what’s the root cause of this drop. (9-30-08 pp. 21 & 22)

However, none of the grade-level PDSA plans specifically addressed the ELL, bilingual, or special education populations.

Given the deeply ingrained culture of individual teacher autonomy at the school, the desire to create more alignment placed the administrator in the role of change agent. The delicacy of that role was heightened by the need to foster ownership within the staff members while nudging them into alignment. Sharon described how she constantly was alternating between pushing and pulling back. Part of the administrators’ role was to set expectations and make a case for why change was necessary. At the beginning of the study Sharon said,

I think we’ve had to take a harder stand and go, ‘That’s not acceptable’. We’re not micro-managing them but I think we’re more active participants in what goes on with this than we used to be. (9-30-08)
The administrators tried to be actively involved and nurturing of the process but also tried to maintain a respectful distance so that teachers would feel empowered and would take ownership of the process. Sharon described what happened in the first year of the study when the administrators tried to rotate through and observe teachers’ PLC meetings,

[One team] was very, very offended if [the assistant principal] or I even stepped foot into their meeting. They felt like we were evaluating them or monitoring them. So in that area I had to pull way back…It was a big mistake that we made. We tried to back off and leave those grade levels alone and at some grade levels we know it doesn’t work. (6-5-09 p. 18)

At the end of the first year, Sharon informed the Instructional Leadership Team members that the administrators had decided that they needed to readjust their role when the PLCs met each week,

Part of what [the Assistant Principal] and I are supposed to do is sit in on PLCs. Not from an evaluator’s standpoint but just being a part of it and what kind of happened last year was we got a lot of pushback from groups. So we kind of stepped back a little bit. I would rather us not step back on that. Please know that we are not in there to go ‘Are you doing your job’? We are in there to listen. (5-28-09 p. 3)

Sharon and a new assistant principal, William, who replaced Wanda in November of 2008 struggled to define what expectations they considered to be “tight”, or inflexible, and which they considered to be “loose”, or flexible. In the end, they communicated that while they were willing to be flexible in some areas so that the teachers were empowered, the expectation that teachers would engage in the process in good faith was a nonnegotiable. Sharon expressed a willingness to terminate the school’s association with teachers who were unwilling to meet their minimum expectations regarding this shift,

There’s also going to be these people that if they’re not on board, they need to go. (9-30-08 p. 13)
Like a tree, Sharon believed that she needed to bend but ultimately must retain some rigidity in order to remain standing.

**Personal characteristics of the school principal**

It was not an easy process for the administrators. Sharon described her own personal evolution,

If somebody had told me six years ago that I would be in this position doing the things that I have been doing and the way that I have been doing them, I would have laughed at them. Because I didn’t think I was that kind of person. People said two things about me: Number one, I was too nice. Number two, I wouldn’t be firm and be able to make tough decisions because I was so nice. (6-15-10 p. 8)

Sharon related how some of her personal characteristics helped her in her role as a change agent,

I’m diplomatically manipulative. I’ve always had a good way of telling people what I want them to do without really ticking them off. But it’s shifting. There are some conversations that I’ve had to have with people that I couldn’t do that with. (6-5-09 p. 7)

She explained how her personal philosophy had given her the courage to have the difficult conversations with staff members,

The important thing is what is best for the children here at this school. If I can live with that at night, even if I have hurt somebody’s feelings, even if things have not gone the way somebody wanted them to go, then I am okay…In the past I would have thought, ‘It’s best for children but I might hurt somebody’s feelings’. So [with] the things I’ve done this year, I kind of look at myself and go, ‘Holy Cow’! (6-15-10 p. 9)

Sharon admitted that,

I thought that it would be easier than it’s been. It’s been tough because we’re really kind of yanking at them ‘Let’s go this way. We’ve got to go this way’. And it’s hard to change. (9-30-08 p. 21)
By the end of the first year of the study, some changes had been made at the school but much work was left to be done. Sharon evaluated their progress at the end of the first year,

I think the pushback will probably come more next year because what kind of happened this year is we set the expectations. This is where we are as a school and this is where we are going to be. (6-5-09 p. 8)

She also noted,

Some grade levels were really good [about class profiles and PDSA assessments] and some weren’t. They weren’t consistent…Some teams were very good about showing us what was going on and some teams I would be going, ‘Yoo hoo, haven’t seen you lately. Is something going on’? (6-5-09 p. 17-18)

The lack of consistency and shared vision was also apparent in a voluntary, anonymous survey that was administered towards the end of May in 2009. Staff members were asked to assess the school against the same continuums from Education for the Future as they had at the start of the school year. This time, however, thirty-four staff members responded individually via SurveyMonkey instead of putting dots on posters. Staff members identified the school at all five rubric levels for six of twenty-one, or 29%, of indicators. Staff members identified the school at four of five rubric levels for an additional 9, or 43%, of the indicators. Thus, there was significant disagreement about the status of the school on 76% of the indicators which indicated that different practices were occurring in different classrooms and at different grade levels. Open-ended comments also indicated the lack of consistency across the school. For example, one comment regarding the current implementation of student achievement processes was “Needs to become more widespread—still isolated to certain teachers and not a full team approach”.

Status of the initiative at the end of the first year of the study
The average ratings on each of the twenty-one indicators ranged from 2.7 to 3.3 on a five-point scale. Two indicators for which the mode was a level 2 (with five being highest) were implementation of leadership processes and implementation of professional learning.

Two-thirds of the respondents identified the school at the following leadership stage:

School values and beliefs are identified; the purpose of school is defined; a school mission and student learning standards are developed with representative input. A structure for studying approaches to achieving student learning standards is established.

However, the majority of respondents did not believe that the school was at the third stage yet:

Leadership team is active on study teams and integrates recommendations from the teams’ research and analyses to form a comprehensive plan for continuous improvement within the context of the school mission. Everyone is kept informed.

Furthermore, two-thirds of the respondents also identified the school at the following stage with regards to professional learning:

Teacher professional development is sporadic and unfocused, lacking an approach for implementing new procedures and processes. Some leadership training begins to take place.

Interestingly, staff members rated the school highest in Information and Analysis approach and implementation and the approach to Continuous Improvement and Evaluation. Eleven of thirty-four, or 32%, of staff members identified the school’s information and analysis practices as:

There is a systematic reliance on hard data (including data for subgroups) as a basis for decision making at the classroom level as well as at the school level. Changes are based on the study of data to meet the needs of students and teachers.

Eleven staff members also described the school’s continuous improvement and evaluation practices this way:
All elements of the school’s operation are evaluated for improvement and to ensure congruence of the elements with respect to the continuum of learning students experience.

These high ratings (“systematic reliance on hard data including data for subgroups as a basis for decision making at the classroom level” and “all elements”) were surprising - given the lack of consistency seen in the evidence - and indicated that staff members may have had different visions of the goal and, therefore, rated the progress of the school towards the goal differently from each other.

**New expectations for School Year 2009-2010**

School Year 2008-2009 was a year of developing expectations and of letting grade levels respond in their own ways. However, Sharon felt that more changes were necessary for School Year 2009-2010. She started to make the case for further change at a retreat at the end of the first year with Instructional Leadership Team members,

> We didn’t get [PLC reporting forms] back consistently [in SY08-09]. Also, there was a disconnect for some people. You were kind of looking at the data but you didn’t look at a common purpose for what was going to happen for those kids that weren’t learning and what you could do together for those kids that weren’t learning. (5-28-09 p. 11)

She continued to lay out her case for change at a retreat with Instructional Leadership Team members at the beginning of the second year,

> I think we did a fabulous job last year in grade levels tightening up and doing focus groups and really trying to look at student data. We really are starting to focus in but I think we have got to look at the next step to see what we can do to intervene before they intervene for us…I think we’re going to have to look at some different ways to service our children that may not make us comfortable but might be better…We have a choice. We can take care of the issues before they [the state] make us take care of the issues their way. (8-5-09 pp. 1 & 6)
She gently introduced her vision and set the parameters,

In talking with other principals we thought about this last year. Maybe we do an intervention block as a whole school. I don’t know. It’s not my call. It’s our call together to look to see what we need to do. Especially for those ELL students and especially to help Special Ed. But also to help our kids because the target is going to go up again next year…It’s tight and loose. You have to help all kids, but how that is going to look is different [among grade levels]. (8-5-09 p. 8)

The plan was for every grade level to utilize the common formative assessment data they already used to measure student progress towards PDSA goals to also create flexible groups for targeted intervention for students. The interventions would have a duration of one hour per week each for reading and math on Wednesday mornings. All teachers, educational assistants, and administrators were to be assigned to grade levels to provide small group instruction. Students would be assigned to groups with homogeneous ability levels for the intervention period. Members of the PLC were to discuss what students were expected to learn and were to commonly plan instruction appropriate for the ability level of the groups.

While two grade levels, grades three and five, had voluntarily done something similar the previous year, this proposal was a significant change for the other grade levels. Sharon knew that there would likely be considerable resistance to the proposal. She told the Instructional Leadership Team members in September of 2009,

It’s like we are moving a huge cow this year. I don’t mean this disrespectfully, but it’s like one big cow because everyone is comfortable with what they are doing. And it works for you. But the thing about it is we have to look…at what is not working and what we can do to make a difference. If we stay status quo, we’re going to have the same stuff that we had this year. It’s not that good teaching isn’t going on…It’s a reflection of our kids and their needs. (9-16-09 p. 5)

In a different meeting she elaborated,

We are past the days where we can close the door and say, ‘Okay, we’re just going to take care of this’…It’s not an option anymore. So if you have pushback [the Assistant Principal] and I are going to be rotating. We can stand there smiling. We
can do that. But it’s hard for people…to think ‘I’ve got to give up two hours a week for somebody who may not take care of my students the way I want them to’. But I think once you start doing it, you can see the power in it. (9-2-09 p. 17)

While Sharon had initially presented the idea as just a suggestion (“It’s not my call. It’s our call together to look to see what we need to do.”), it was clear that participating in a grade-level wide intervention was actually non-negotiable. Sharon discussed the roles of Instructional Leadership Team members in getting other teachers to participate,

You are not the mom of that group. And you’re not in charge. I don’t want you to feel like your role is to make sure they are doing what they are supposed to do and run to us if they are not. But you also need to understand if you need help, that’s what William and I are here for. And William and I don’t have a problem coming and going, ‘This is what we’re mandating’. (8-5-09 p. 10)

Not surprisingly, many questions and concerns were voiced by staff members. Special Education teachers had several concerns. Instructional Leadership Team members relayed the concerns to Sharon in a meeting in September of 2009. One concern was the impact on Special Education students if they had to get used to a different teacher,

I was told the Special Ed teachers were going to keep their own kids. We had kind of discussed that since [the students] are moving around so much. (9-16-09 p. 1)

After a discussion, it was concluded that decisions would be made on a case-by-case basis. If working with a different teacher for one to two hours per week would be considered detrimental to a particular student, such as a student with autism, then the student would not be assigned to a different teacher. Other students with an IEP would be assigned to an appropriate ability-level group for intervention.

Special Education teachers were also concerned about students having fewer contact hours with them as the teacher. As one ILT representative explained,
I know our Special Ed inclusion teacher is really worried because she is held accountable. She follows the IEP and she doesn’t know if she is going to be able to...She like kind of wanted to stop. (9-2-09 p. 11)

Charmaine, the Site Specialist, responded to the concern,

All their goals go with the NM standards. So even in your group you are hitting these goals – [increased] reading levels, [increased] math levels. Those are all their goals too. So one hour a week is not going to be a detriment to them. It is only going to help. (9-16-19 p. 2)

Even general education teachers expressed concern about losing some control over how their students were taught,

I know there are a lot of teachers that are hesitant to send their kids to these groups because they think that they are not going to be teaching them the way that they would teach them. (9-16-09 p. 8)

Several teachers who had had experience with grade-level wide interventions then discussed the benefits that their students had experienced either hearing the same thing from a different teacher or hearing something slightly different that “clicked” for them.

One of the challenging aspects of the intervention was that it was based on standards, not on the curriculum. For some teachers, this was a new way of thinking. One ILT member noted,

There are still a lot of teachers that think they are going to be teaching from the curriculum. And then they were surprised when I told them that we were focusing on our PDSA. (9-16-09 p. 7)

The need to pick and choose materials to support student progress on particular standards rather than following a preset sequence of materials and the need for collaborative planning required some teachers to transform their instructional practices.
A few people were so hesitant that they openly said that they planned to sabotage the intervention so that it wouldn’t work and would be dropped. One ILT representative relayed what she had heard from a teacher,

I do know of one person who plans to just basically do ditto puzzles and things like that during that time to kill that one hour because they don’t want to do the intervention. (9-16-09 p. 10)

Sharon responded,

That’s where William and I come in…It’s not like we are going to be the intervention police but they need to know clearly that this is not optional. (9-16-09 p. 10)

**Communication and the role of Instructional Leadership Team members**

One factor that contributed to teachers’ confusion and hesitation was the lack of consistent and reliable communication. Sharon had mentioned the intervention plan directly to all staff members in the beginning-of-the-school-year orientation in 2009 but had left it to the Instructional Leadership Team representatives to communicate the details as they evolved to the other teachers. This led to staff members hearing different things from different people. By mid-September ILT members were asking Sharon to clarify expectations directly to all staff members,

I feel like there are a lot of inconsistencies [with plans for implementing interventions]. And there’s lots of ‘Why are we even doing this? Why do we need to do this? We don’t need to do this…I was wondering if you could send something out so that everyone is on the same page. We know exactly what to expect. (9-16-09 p. 4)

Clearly not all staff members agreed that there was a need to change.

The process of introducing the intervention proposal to the staff brought up questions about the role of the Instructional Leadership Team. The role of the Instructional Leadership
Team had evolved over time at the school without those changes explicitly being articulated.

Sharon acknowledged,

What they had to do on ILT versus when I was on ILT is a whole different kind of story…Instead of the dissemination of information, it was more of what are we doing and what are we going to do next? (6-15-10 pp. 16 & 17)

To Sharon, it seemed natural to empower the Instructional Leadership Team to make decisions. She stated in one meeting,

Isn’t the purpose of [the grade level representatives on the ILT] to make those decisions for the grade level? (2-11-10 p. 14)

However, some of the ILT members relayed concerns that their grade-level colleagues had expressed about the way that the ILT was functioning during School Year 2009-2010. For example, Anne articulated the following concern from first grade teachers at a meeting in February of 2010,

My grade level [is] a little concerned that decisions are coming out of ILT, are just being made and it comes across as a mandate and less like decision-making. They said that it used to come back to the grade level for discussion and then come back [to ILT] and then a decision was made. Now it just seems like decisions are made. (2-11-10 p. 13)

One of the team members who had been at the school for a long time noted that the purpose of grade level meetings had also evolved over time,

I think they’re thinking about how it used to be like when we used to have grade-level meetings where you would meet a lot and they would get to talk about things. Now we have PLCs…and interventions. (2-11-10 p. 14)

Sharon pointed out,

If we go back to what they want then there are going to have to be two sets of meetings. There’s going to have to be a grade-level meeting every month and there’s going to have to be a PLC meeting every week. (2-11-10 p. 15)
She also noted that the decision making process had previously been extremely inefficient,

What frustrated me was that…it bounced back-and-forth for six months and then maybe a decision might be made the next year. (2-11-10 p. 14)

The changes in the role of the Instructional Leadership Team, however, had an emotional impact on many teachers. One teacher explained,

One of the differences about ILT this year was this is the decision-making committee and we’re going to make the decisions and then tell you what to do. Part of that was because in the past you’ve had some principals here that would let us discuss something to the nth degree and it is a dead horse but we are still beating it and then we finally make a decision in April. I think part of it was to avoid that. So, as a result, things that were coming down to me…felt like a mandate. You will do this. And Civics 101 says if you have a democracy, just say ‘We’re thinking of doing this. What do you think?’ Take it back once and then this is the decision that is being made. That didn’t happen this year. It was all coming down from ILT. (5-17-10 pp. 11-12)

Elizabeth noted that the Kindergarten teachers felt like the concerns of the teachers regarding the intervention initiative were being disregarded,

It’s one of those things where they feel like they are not being heard. It’s just like, ‘No, [interventions] are not going away. You’re going to do it. You have to do it. So do it.’ And they are feeling like, ‘Well, why do we even have to?’ They are just like not even caring anymore because their voice isn’t being heard and if it is being heard, they are still being told you have to do it. (2-11-10 p. 16)

Elizabeth also articulated to Sharon in an Instructional Leadership Team meeting the impact that the changed role and communication structure had on her as the grade-level representative to the Instructional Leadership Team,

What happens is we’re getting the brunt of that. Like my grade level is saying, ‘You go tell them blah, blah, blah’. And then I am dealing with it and it feels like it was my decision as opposed to my grade level’s decision. So they feel like they are not important…They’re pretty hostile. (3-3-10 p. 7)

Sharon frequently acknowledged the difficult role that grade level representatives had. She told them,
It’s a big thing to herd a group of teachers. You are walking that tightrope between being a facilitator and being a mother. (6-5-09 p. 27)

The grade level representatives were put in the role of facilitating processes about which they themselves might have been hesitant.

One of the contributing factors to some teachers’ unease with Instructional Leadership Team members making decisions was the fact that Sharon appointed members instead of having grade levels select their representative for the 2009-2010 School Year. While there had been some positive comments made about the ILT in SY2008-2009, Sharon was concerned that not everyone was a good team facilitator. She stated at the end of September, 2008,

We’ve got a couple of grade level chairs that are not good facilitators to run a team…It’s real important that you have leadership that can [not only] guide but also listens. (9-30-08 p. 6)

At the end of School Year 2008-2009, Sharon appointed members of the Instructional Leadership Team for the following year. Some members stayed, but others were replaced. Sharon explained,

It was a conscious decision on my part…to appoint people [to ILT] this year. I needed a team that would help us move forward…Not ‘Yes People’ but people who will say to you, ‘This isn’t going to work and this is why’. I’m not the kind of person who expects everybody to clip clop along with me. I don’t have all the answers but I want people that are committed to doing good work with kids and who want to do what’s best for kids. (6-5-09 p. 17)

Professional development for Instructional Leadership Team Members

The new members of the Instructional Leadership Team were not provided with training per se. Rather the principal and experienced members described “the good, the bad,
and the ugly“ about the ILT membership as part of a leadership retreat. Some experienced members of the ILT had attended a training previously. Leslie stated,

The PLC training attended in 2007-08 ‘was more general than specific’ and ‘it wasn’t data driven which changes your dialogue’. (10-14-08 p. 18)

She added,

People who went to the PLC workshop last year aren’t necessarily the leaders of the PLC this year. I think we do need maybe some more training in how to run the PLC and make it more effective. Protocols and that kind of stuff. (10-14-08 p. 18)

Sharon acknowledged that the professional support of members as leadership team members had not been fully provided,

One thing that didn’t happen as well as it should have was checking in on their growth and knowledge – what they need and the support that they need. (6-5-09 p. 27)

She felt that professional development for members would also contribute to greater alignment of efforts. She stated to ILT team members at a leadership retreat,

I would really like to try…[to] have a little bit of professional development among us…because sometimes you branch off and you know what you are doing but it’s kind of that alignment thing. Are we all doing it together? (5-28-09 p. 12)

In fact, some members of the Instructional Leadership Team did participate in professional development. However, members attended different conferences and did not hear a consistent message. In addition to the initial data training with Dr. Bernhardt attended by seven staff members, several other team members attended three other workshops. Anne and Diane, the first and third grade representatives attended a two-day workshop on managing professional learning communities sponsored by Summit Tree with Sharon in September of 2009. Keynote speakers included Rick DuFour, Rebecca DuFour, Robert Marzano, and Tom Guskey. Natasha, Eleanor, Leslie, and Sharon participated in a “data-
driven dialogue” training in Connecticut. Cheryl and several Reading Resource teachers attended a training sponsored by the University of Arkansas Center for Literacy regarding a Comprehensive Intervention Model (CIM). While all of the trainings were beneficial, there was no professional development that all team members received first-hand. Although Sharon frequently expressed a desire for Instructional Leadership Team members to participate in a book study, the sessions never took place.

Sharon also felt that restructuring the agenda of the ILT meetings could also provide better support for the ILT representatives,

I need to ask them the four basic questions in the sense of ‘What is working? What do we do if it is not working? What do you need from me to get you there? What do we need to do for people? How can we better support you?’ I need to make the shift more that way because I think in the past the principal was pretty much the disseminator of information. We did have conversations. We listened and adapted but I don’t think that I gave them the support that I need to give them in those kinds of terms. (6-5-09 p. 27)

Thus, the grade level chairs have been put into new roles without consistent professional development to help them learn how to facilitate teams under the new paradigm of evaluative inquiry.

Part of the challenge of supporting Instructional Leadership Team members is that each of them facilitates a grade-level team that has its own history and team dynamics. Some of the grade-level teams have a history of collaborating smoothly and some of them do not. Profiles of four grade levels and the bilingual program illustrate the diversity of issues that need to be addressed.
The Third Grade PLC

The third grade PLC had a widespread reputation for being one of the most cohesive and productive teams in the school. Diane, the grade level chair for third grade, described her team,

Third grade has the benefit of having mostly the same members over the last few years. So I think our vision is a pretty shared one, a pretty commonly developed one because we have been working together for so long…There’s a lot of trust in our group because we’ve been around together for a long time. But also even the people who just joined us this year have felt very open to say, ‘I don’t get it. You know, I’m having a problem with this.’ I’ve seen a lot more of that happening this year from the group that joined us but also from our own teachers. (2-22-10 pp. 1 & 9-10)

Karen, an ELL teacher who had been a member of the third grade PLC for the past two years concurred,

We have a very effective PLC. They communicate with each other. They listen…We all agree to disagree…It doesn’t get personal. It’s all about the students and what the student needs are. (5-19-10 p. 11)

Even ancillary staff members who worked with students from different grade levels noticed the effect of their collaboration on their students. Natasha, who provided math intervention to students in grades three, four, and five noted,

Third grade is very, very cohesive…It seems to me that they do a lot of common planning. Or, at least, ‘We are all going to work on this unit next.’ You know, ‘What are you doing? What works?’ Like they are really having those good conversations in their PLC…They are, for the most part, strong teachers but they are also not afraid to take ideas from each other and just talk through what works. (2-22-10mi p. 2)

The third grade team also examined assessment data a bit more thoroughly than some of the other grades. Diane described to Patricia, the independent consultant, how the third grade team used assessment results to inform instruction,

We look at the standards and we look at how they do in 3rd, 4th, and 5th as well. In 4th and 5th they have some test data. In third grade we’re stuck where we’re looking at last year’s kids to define this year’s children and what they need to work on but we go
back and we look at how they performed on the test and then what are the state standards they are required to know and where are the biggest gaps. (2-24-09a p. 8)

Diane explained how they also used other data to plan instruction for their current students,

The [SBA] data that we have is from the previous year’s kids…Take that with a grain of salt. Then we really focus on the DRAs and the NWEAs and those beginning of the year assessments that we do collectively…We say where they are and where we need to get them. (2-22-10 p. 1)

The need to use the data to plan for weekly interventions changed the way the team looked at its data,

[Before] we would wait until that day in October when we finally got our SBA results in and then we would develop [PDSA goals] from that. Since we knew these intervention things were coming…we knew we needed to be defined earlier. So I think that led us to using the data that we had on hand more and going more with what we had as hard numbers versus like ‘I feel like we should do X’…I think it helped too that we had done the math interventions last year as a grade level so we knew kind of what we wanted it to look like. (2-22-10, p. 2)

The third grade team voluntarily worked together to implement grade-level wide interventions in School Year 2008-2009. They were one of two teams to implement the intervention model at their own initiative. Diane explained the origin of the interventions,

It came about from many multiple conversations at the grade level where we just said we don’t have enough time in our instructional day for differentiation. Our highest highs aren’t getting what they need and our lowest lows aren’t getting enough time with the teacher. (2-22-10 p. 2)

She continued,

Five or six years ago we tried this on a daily basis with language arts. And there were some really good and bad aspects with it. And we thought maybe we’re not ready to jump in every day again that deeply but let’s try it once a week…We were really pleased with the results that we got from it last year. Their test results, their improvement in their overall math, their confidence level. The kids did really well on the SBA too in math last year. But you could also see the change in the kids from day to day. The lower kids seemed to have more confidence. They would be a little more participatory in class and those high kids just adored it because they got to go do something special and different. (2-22-10 p. 3)
The success of the third grade with the intervention model helped convince the administrators and the Instructional Leadership Team that it might be beneficial to expand the model schoolwide. At the leadership retreat at the beginning of the second year of this study, Diane gave the following testimonial to other members of the Instructional Leadership Team,

I got a lot of positive feedback from my grade level [with the interventions] being contained to once a week for an hour. Much more than that I think it would have been really intrusive on what they were doing for their kids. It was nice to take my kids that I had told 18 times to do it this one way and have another teacher tell them something just a little bit different and then I get them back…We got to know all the kids in the group…and it kind of brought us together as a group more. (8-5-09 p. 13)

She added at an Instructional Leadership Team meeting a few weeks later,

One thing I noticed is when we all started working together and having to share the kids, everybody became everybody else’s problem. Or everybody else’s success. And so it was a lot less like ‘My kids do this’ or there was a lot less of that going on. It was ‘We need to get everybody to…’ (9-2-09 p. 16)

Thus, the interventions helped build ownership for the success for all students at the grade level.

While the third grade team had instigated the interventions in SY2008-2009, they modified their model to match the schoolwide model in SY2009-2010. Some of the changes included utilizing ancillary personnel and aides to provide small group instruction and providing interventions in reading as well as math. The changes did provide some challenges to the team members. Diane explained,

[One challenge has been] including non-classroom teachers [in the intervention] because they don’t have the training. That’s taken up a huge amount of time and planning for them…Now I have to teach the teacher how to teach it. (2-22-10 p. 4 & p. 9)
Although the students were able to be in smaller groups, the team felt that bringing in non-grade level personnel also had a negative impact on the students. Diane elaborated,

> We plan a lot more together now for the intervention times but we’re actually doing less differentiation than we were last year because we spend so much time planning with the people that need the help. (2-22-10 p. 11)

Providing both interventions back-to-back on the same day also was a challenge for the students. Diane noted,

> Last year we just did it with one subject. Doing both subjects together on the same day has been long. It’s long and it’s hard on the kids [because] they are getting intense teaching for basically two hours straight and then their poor little brains are fried by the time they come back. (2-22-10 p. 4)

Thus, changes to the intervention model affected students as well as teachers.

One difference between the two years of this study was the amount of common planning that was required. The additional common planning did have some benefits. It helped to define what was important and to narrow the focus of what was considered to be essential. It also helped to shift the culture around data at the third grade level. Diane explained,

> It’s less compliance and more using it as a formative tool. It’s more of, ‘What did you find out and how are you going to change what you are doing based on what you found out?’ I think it’s coming naturally…because we have to do stuff every single week. If you have to do stuff every single week then you have to know where the children are and you have to have good solid grounds to stand on versus just ‘I feel like’ or ‘I noticed that’. (2-22-10 p. 18)

She added,

> We [looked at the data together] more explicitly this year than we have in the past. (2-22-10 p. 2)

Diane recognized that while her team had reached a new stage in using data, they still could improve. She stated,
The purpose of data is never just in reporting it. But in the daily grind, data gets reduced to that...At first [using data] is overwhelming. And if you’re not practiced in it then it’s completely mind-boggling. But the more you do it, the better you get at it. I know that we still have a long way to go. Like we don’t use our NWEA anywhere near like we could and should. (2-22-10 pp. 8 & 9)

The need to plan for weekly interventions prompted an increased awareness of and use of data, but additional training and time might be necessary to help teachers reach an even higher level of data use.

The team also discovered some unexpected benefits from teaming with ancillary teachers to provide the instruction. Diane stated,

It has kind of forced us to look more at the whole child and get a more total picture. There’s been more of discussion [about how the child is doing in related arts] than I have ever seen before...So I know it was painful for the poor ancillary people to come in and do it but I think it was really good at the same time because it brought us together in a different way. (2-22-10 p. 7)

Diane also felt that expanding the intervention concept campus-wide despite resistance from some teachers also positively impacted the culture of the school. She stated,

I think the really good thing that has come about with all these intervention days, as hard as they were for everybody to implement and as much struggle as we have had, is that the entire campus has now come together and seen that there is a real need to meet these kids and provide them with interventions. I mean I don’t think it was universally recognized...Everybody knows that it is going to affect and impact their life now. (2-22-10 p. 7)

While the third grade team had already recognized the need for interventions, the mandated expansion of the intervention model may have helped expand the view that interventions were necessary beyond a few isolated teams.

Members of the team attributed its success to the leadership of their grade-level chair and the creation of a psychologically safe environment. When asked what made the third-grade PLC effective, Karen stated,
I would say the leadership of the PLC is first and foremost the most important thing. We have an excellent leader…She’s very focused, very structured, very knowledgeable. She has control of the group. If we start getting off on tangents, she brings us all back… And then everyone feeling safe. Comfortable enough to talk and express their ideas and opinions [about] various teaching strategies [and] best practices. (5-19-10 p. 12)

Diane attributed part of the willingness of third grade teachers to work together to the interdependence that the accountability testing created,

In a way 3rd, 4th, and 5th have it a little bit easier because we all either collectively succeed or collectively fail through this unifying test that we have to take. And so we all have this push. It doesn’t matter if I’m great and the person next to me isn’t because then we all fail. (2-24-09b p. 15)

Thus, the externally mandated testing did, under certain circumstances, encourage teachers to work together collaboratively to take ownership for the success of all students. However, this was not the result at all grade levels where students took the accountability tests.

**The Fifth Grade PLC**

The fifth grade PLC was the other team that had voluntarily initiated grade-level wide interventions in School Year 2008-2009. While there were similarities in the interventions themselves, the team dynamics were very different within this PLC which greatly impacted the creation of a culture of evaluative inquiry.

Natasha, who served as a math interventionist for students in grades three, four, and five in School Year 2009-2010, was the grade-level chair in School Year 2008-2009. She recounted how the team initiated the interventions,

We were talking about what we can do to help those kids that were struggling so much. And I talked about how when I was over at [another elementary school] a couple of the grade levels had started these groups. I asked, ‘Would you be interested in breaking the kids up into groups and doing a rotation [where] we can concentrate on whatever their lowest area was?’…[Another teacher had experience at another
school and said it was good.] At first we were talking about doing reading and math
and then we got bogged down in all the reading stuff. It was just overwhelming. So
we decided as a grade level to just try it for math because it is more cut-and-dry,
black-and-white. (2-22-10mi p. 18)

There had been some question as to whether the team should focus on reading or math
because the data showed a greater need in reading. Sharon explained the dilemma,

Wanda (the assistant principal at that time) was kind of upset [that the fifth grade
created a focus group for math] because she said they really need it in reading. I said,
‘I realize that but the fact that they’re even doing it and taking the time to try to meet
the needs of kids and work with them, that’s a huge step for them’…The grade level
chair said, ‘I know we need some areas of reading but the math is easy and concrete
to grasp on to right now and that is a good step for that group.’ And I agree with her.
(9-30-08 p. 11)

In this case, the ownership of the process was more important than focusing on the area of
greatest need.

Natasha described how the team set up the math interventions and the team members’
reactions,

We set up weekly rotations where all the teachers get together and identify their kids,
our lowest, middle, and high and divided them up into groups. Three teachers have
low, and they are limited to ten students each…We don’t keep any of our own math
kids just to give them all a different perspective…We do an hour a week on
Wednesdays. We do specific instruction on computation, estimation, and the four-
block…As the teachers report that in class they are really doing much better and
scoring higher on the three-week assessments, then we move the kids…So we just
kind of started it and went with it and they really liked it. (2-24-09b pp. 3 & 4 & 2-
22-10mi p. 19)

The interventions appeared to also lead to higher student achievement by the students,

Last year we…concentrated on computation and estimation because that was our
weak area. And we went up 18 points or something so we were really, really proud
of that. (2-22-10mi p. 11)

Natasha was referring to the percentage of points fifth grade students earned in these strand
areas on the state accountability test.
Similar to the third grade, the fifth grade also had to modify their intervention planning process for School Year 2009-2010 in order to accommodate the schoolwide model. Natasha described how the inclusion of ancillary teachers and aides affected their planning process,

> We started kind of planning towards the middle because all of the people that worked with the middle groups are not classroom teachers. And so they were the ones who really needed the specific ‘step-by-step tell me what to do and say and script it for me’. So we were planning towards that and the teachers that had the high groups would adjust as they needed and those of us with the low groups would adjust as needed. (2-22-10mi p. 15)

Unlike the third grade, however, the fifth grade team appeared to have less of a collaborative culture and the work of the team was driven by an individual, Natasha. When asked how the group planning process was going, Natasha responded, “I plan, they listen.” (2-22-10mi p. 15) A district instructional coach, Kelly, who worked with the fifth grade team members in SY2009-2010, also noted that there was not a collaborative culture within the team. She stated,

> I couldn’t even get them to share their lesson plans. If we’re all working on this standard right now then that’s what we need to be looking at. You know, this is what we are teaching, so how are we teaching it? They are not there yet. (6-3-10, p. 7)

The lack of collaboration and alignment was evident throughout all instructional practices. Sharon invited Kelly, a district instructional coach, to work with an inclusion team and the ancillary staff members who worked with the same students. Sharon explained her rationale,

> The reason we targeted that fifth grade team was because looking at the data they had the double dose of ELL and Special Ed in there…and those were the kids with the greatest area of need. (2-11-10 p. 7)

The administration had chosen to place almost all students with special needs into one classroom with both a general education teacher and a special education teacher with pullout
assistance from the math interventionist, special education interventionist, ELL teacher, and bilingual program coordinator with the hope that alignment would be easier if the students were in one homeroom. After working with this team, Kelly observed,

There was no collaboration [within this team]. There was no direction. It was just kind of haphazard. They weren’t planning together…There was no cohesiveness of schedule. Kids were just coming out left and right and there was no protected time block. There was no intervention block. (6-3-10 pp. 1-2)

This lack of alignment was not new. Sharon had noted the previous year,

The fifth grade bilingual teacher said it really worked well when the ESL teacher came in and supported what we were doing vocabulary wise and classroom wise. I was wondering ‘How did that disappear?’ It didn’t come from me. It was the exact thing [the ESL teacher] was saying from her end. [The classroom teacher] would be teaching something and she’d be pulling them for something else so it wasn’t aligned…It’s got to line up and the arrows have got to go the same way… So what we’re going to do for [the ESL teacher] is we’re going to build in time where she sits down weekly and collaborates with that teacher and asks what’s coming up, looking at the curriculum map. So either if she pushes in or she pulls out she’s supporting them with what [the classroom teacher] is doing. (6-5-09 p. 5)

Sharon and William, the school administrators, therefore, asked the district instructional coach to work with the team in order to try to create better alignment. Kelly related her experience with the team,

My suggestion was to start with something very simple like…starting focus groups with NWEA data. The teachers said that they understood NWEA data…and that they would pull the data. The first meeting after that they still hadn’t pulled the data. I asked them why and it was, ‘We didn’t have the time’. So I said, ‘Okay, by the next meeting will everyone have the data?’ and they said, ‘Yes’. And they still hadn’t…At this point they just weren’t going to have it.

So I asked the assistant principal if he could help…He called me on the phone and asked me to walk him through the steps on NWEA so I showed him at that time how easy it was…Once he saw the numbers, he became much more aware. He had a very frank conversation and said, ‘Half of these kids are not reading anywhere near grade level. This is a priority’.

We pulled the Class by RIT and we went down to the strands. We pulled all the math and all the reading so we could really look at where those kids were…We looked at
the groupings and that really gave them a focus and they understood…which kids were really significantly behind [and] they restructured their schedule a little bit.

I couldn’t get them to the next step of the DesCartes and really focusing in on the skills…It was like a block…They still wanted to go through the teacher’s manuals, business as usual so to speak. I kept trying to explain to them that [the basal] is a tool, not the curriculum…I think there was a big issue with ‘I need to teach them out of the basal because that is what they are going to see on the test’. (6-3-10 pp. 3-5)

The interpretation of the student achievement data also became an issue. Since the team lacked a common vision, there were different interpretations as to the acceptability of the growth that students made across the school year. Kelly explained,

The average growth for this class who is 50% below grade level was 3-4 points [at mid-year]…My experience is that they are going to grow more…I did an analysis of the RITs and where they were in comparison to the grade level so we could take the growth point out of it…I thought the meeting was going to be a real A’ha for those teachers and be like, ‘Oh, my gosh. We’re not making the progress’…But instead it was exactly the opposite where they were lauded for the progress they were making [because it] met the national average [for growth]…They kind of came back with, ‘Well, we are doing everything right’ and all of the instructional tools that I brought to them after that they were like, ‘We already do that’. It was like they had kind of shut the door. (6-3-10 pp. 8-9, 12-13)

Natasha described the reluctance of team members to engage in discussions of vision and determining areas for collaborative continuous improvement,

The vision came from ILT…Pretty much we need to use data to improve student scores and student achievement…It really should have been tweaked, looked at, discussed, and dissected but unfortunately they are not that kind of a group…They think of it as ‘Here’s my list of scores. You want them? Here they are.’ Our grade level chair asks, ‘What does this data tell us?’ and they all look at me. I’m always the first one who says something and I always try to make a positive. ‘Well, we did go up in this.’ You know, that kind of thing. Then, ‘It looks like we need to work on this’. ‘Yeah, okay.’ And that’s about all the discussion we ever have [laughs]. (2-22-10mi p. 10)

It may have been that fear inhibited the team members’ willingness to collaborate and expose what was actually happening in their classroom. Kelly commented,
They feel that if they make their scores public and they really look at that data in a reflective way that it is a judgment on them and themselves as a teacher. (6-3-10 p. 6)

Natasha agreed that team members may not have been supported enough with how to use data in a psychologically safe manner. She stated,

I think people have been trained to collect data. They have not been trained sufficiently with what to do with it once they have it. (2-22-10mi p. 1)

She noted that team members had also been sensitive to previous efforts to provide support in this area,

They are one of those typical groups that if you tell them what to do, they get mad because ‘We’re always being dictated to. It’s always coming from the top down’. But if you ask them, then they feel like it is a waste of their time. ‘Just tell me what to do and I’ll give you the scores and then you can make your graph and we’ll move on.’ That is not the intent of data. It’s not to make a pretty little graph that we can post on our wall. (2-22-10mi p. 6)

It may have been that the teachers’ resistance to thoughtfully analyze the data and use it to drive their instruction was based on insecurity. It was less threatening for a teacher to take a teaching-focused approach and say ‘I covered the curriculum and the students chose not to learn it’ than to take a learning-focused approach and say ‘I did not adequately facilitate their learning on this standard’. The analysis of student achievement data provided evidence of student learning which could have been used to compare teachers to each other and some teachers may have feared that they would not look good in that comparison. Natasha stated, “Teaching is not a competition…But not everybody buys into that.” (2-22-10mi p. 3)

While progress was slow in developing a culture of evaluative inquiry at the fifth grade level, there were some promising signs. The PLC members, with the assistance of Natasha, made some significant realizations through the analysis of their data. Natasha
explained how they learned to better define what students were expected to know and be able to do in certain standards,

Historically we always score low in geometry and measurement. So then I started doing a little research and looking because in our minds we [think] measurement and whip out the ruler. So I started checking and brought to them documentation [that] it’s not just measuring with a ruler…Those lovely perimeter and area questions come in both of those sections. And they were like, ‘Oh!’ [laughs] Nobody had ever looked to see. (2-22-10mi p. 11)

She also shared how she had modeled the analysis of a formative assessment for the team members,

We hit our goal or just about for computation and estimation and so we decided to start a new goal third quarter which was geometry and measurement. So we gave all the kids a pretest and then rearranged [groups] based on that pretest…I did an item analysis to see what type of questions most of them blew and which ones we seemed pretty strong in. (2-22-10mi p. 13)

There were also some signs that other PLC members were beginning to shift their attitudes and were more willing to look at the data. Natasha related how participation in a district-wide task force to develop the Math Quarterly Assessments (MQAs) had affected the attitude of one of the team members and helped her provide constructive feedback,

She usually is a real negative Nelly and everything is horrible, bad, terrible, etc. Now she has been much better in the PLC…[After analyzing the item-plot report at the MQA training she said to the PLC], ‘Yes, historically our scores are lowest in geometry and measurement but they are also only worth this many points. We shouldn’t drop [them] and not teach [them]. We need to concentrate and do them but we need to spend more time in this other area because that’s where we have so many more [possible] points’…They were all intrigued by what she was telling them and they wanted to see [the item-plot report]. So I think that’s kind of sinking in to them a little bit more. (2-22-10mi p. 10)

Thus, while the third grade PLC had been perceived to be highly functioning, the fifth grade PLC had had a more inconsistent record. Most of its data analysis and data-driven instruction had been driven through the leadership of one individual. However, there were
signs that other team members were beginning to shift and better appreciate a culture of evaluative inquiry.

**The Fourth Grade PLC**

In contrast, the fourth grade team had acquired a reputation for being one of the most dysfunctional teams at North Mesa Elementary School. Natasha, who worked with students from grades three, four, and five, gave her observations of the alignment of instruction at the fourth grade level,

Fourth grade is where I see the biggest disparity...Teaching methods, amount of time, and talking to the kids...Not everybody is on board with the ninety minutes [of math] a day. Not everybody is on board with differentiated instruction in their classroom. It’s still a lot of whole class lecture. ‘Here’s an assignment, do it’ versus some teachers that pull the kids aside and really help them. Obviously they’re not having those conversations and using data in PLCs either. Otherwise I would be guessing that they would be doing something about this and it doesn’t seem like they are. (2-22-10mi p. 4)

Furthermore, the PLC meetings, where some of these alignment issues could be discussed and negotiated, were sidetracked by conflicts between some of the team members. Eleanor, an ancillary literacy teacher who had worked with the fourth grade PLC for the past several years, described the meetings,

There was a lot of friction. And a lot of it was friction between personalities...There was personal criticism rather than being objective. (3-2-10 p. 7)

Marion, the grade level chair, agreed that the team had had a dysfunctional history. She noted, “They sweat the small stuff.” (3-11-10 p. 1)

The conflict between the team members had to be addressed before they could productively focus on the analysis and use of the data. Marion described some of the team dynamics during the first half of School Year 2009-2010 and a turning point for the team,
They [were] just not a cooperative group until we had a big blowout and I think that is what helped all of us change and develop into better teachers. We finally got to the point where instead of whining and complaining…we try to find solutions to concerns and issues…

We had one teacher that just kind of felt that she was the only one that knew how to teach and she couldn’t understand why everybody wasn’t doing things her way. We couldn’t get through to her that we had a lot of fine teachers here and that if we all worked together we could get more done…and it would just be a better blend of ideas…Everybody was scared to approach her and so they were building up these feelings of resentment that every time she came in you would hear these sighs…

Finally, when everybody started to ignore her and let her know that we have an opinion too, she blew up at all of us. She didn’t want to attend meetings anymore or anything. We just let her know that if she wants to be in our group that we all have to work together, that we have to work as a team because that’s what we are. We are a team….

She is a real good team member now…and she wants to help which is nice. Now whenever she realizes that she’s been either obsessive or annoying about small stuff…she says ‘Oh, never mind. I’ll just bring it up another time or I’ll think about it’. Now we’re working as a team and can get so much done. (3-11-10 pp. 2-3)

The blowout was one catalyst for the change. In fact, Marion stated, “I think that blow out had to happen in order for there to be some change.” (3-11-10 p. 3) It helped to clarify the acceptable norms of behavior for team members. However, the blowout was not the only catalyst for different team dynamics. Another catalyst was the use of conversation protocols to guide looking at data.

In October of 2009, Eleanor, Sharon, Natasha and I attended a three day workshop entitled “Data-Driven Dialogue: Practical Strategies for Collaborative Inquiry” with Bruce Wellman and Laura Lipton. The purposes of the training were as follows:

- To develop practical structures for using data to focus a group’s attention and energy;
- To understand and apply a three-phase model for guiding data-driven dialogue;
• To extend a repertoire of tools for mediating productive group learning, planning, and problem solving; and

• To explore methods for surfacing multiple perspectives and frames of reference.

Sharon felt that the skills taught in the training were so beneficial that she sent Leslie, the bilingual coordinator, to attend a different session of the same training to help her facilitate change in the bilingual program.

Eleanor used some of the protocols she learned at the training to consciously try to change the fourth grade team’s dynamics. She explained how the training helped,

The really good thing about the training that we had was that I was able to lead a few meetings and bring some of that core knowledge that we got to the group and help them depersonalize it and look at the data and start getting serious about how we are going to move our kids forward rather than going off track and getting personalities involved…I took a couple of quotes. One of them was on listening because that was a huge issue. (3-2-10 p. 8)

She described how members of the team responded to the structured activities,

We went through one of the activities [protocols for looking at data]…It was not real comfortable for them. There was some, you know, squirming and a few faces made and that sort of thing but they got the point. And the really cool thing was about halfway through the body language was changing. They were leaning forward and listening rather than sitting back and [crossing arms]. (3-2-10 p. 9)

The protocols did appear to help team members develop more productive interaction norms. Furthermore, they provided an unexpected benefit. Eleanor, who had been a fourth grade teacher at another school for a number of years before becoming an ancillary literacy teacher, noted,

The other thing for me is that [the use of protocols] allowed me to be in that circle rather than an outsider. (3-2-10 p. 9)
The fact that Eleanor felt like she had not been accepted by the members of the fourth grade team despite having participated in their meetings for more than a year was indicative of the culture that had dominated their team.

Eleanor also assumed the role of data coach within the team. She explained,

I have been able to access NWEA and get the information for fourth grade so that we can get it all together. I probably have more snippets of time than the regular classroom teacher has and so I have made that my job to help gather that data and present it to them. Get it in a chart form...so that they can see it... and can make decisions. (3-2-10)

She described her impression of how the team had previously interacted with data.

They really didn’t know what to do that much. There was going through the motions because you have to but not really doing a lot with it. I did not see that they were doing a lot with data...This year we’ve had to get more serious with it. (3-2-10 p. 5)

By collecting and organizing the data, Eleanor helped the team meet requirements which team members did not either have the skills, time, or desire to meet. This also helped develop Eleanor’s relationship with team members. She clarified,

[Gathering the data] has helped ingratiate me to the group because I am doing my part. (3-2-10 p. 15)

Given the lack of trust and psychological safety within the team, Eleanor needed to be careful in how the data was presented. She related how she needed to organize the data to bring the risk level down to a level where the teachers could engage with the data instead of putting up walls for self-protection,

One of the mistakes that was made at the very beginning was individual teachers showed up on the board. Not good. They are not ready for that. [Now] I compile the information from each of the teachers to depersonalize it because we are looking at it as a team. (3-2-10 p. 16)

Given depersonalized data, the team did then begin to examine the students’ achievement data.
The structure set up by the district and the school administration created pressure for the team to examine a variety of data. Eleanor explained what type of data they looked at for various purposes,

We started out the year looking at our SBA data and using that information. I helped fourth grade this year with interpreting that data and we got our first PDSA started using the SBA data. (3-2-10 p. 6)

In order to monitor progress on the PDSA goals, the teams were expected to create common formative assessments. The process of collaboratively creating assessments brought some awareness of the need for greater assessment literacy. Eleanor described what happened,

It took a little bit of going to get the short cycle assessments in place...When we’re making up short cycle assessments, we’re not sure whether they are going to be reliable or valid or whatever. We’ve learned from that too because we’ve made mistakes on making our test based on what our kids need. (3-2-10 pp. 6-7)

Because of the interventions and the need for updated information, the team persisted. Eleanor cited the progress that the team had made,

The good news is that each cycle has gotten better. And the ball is really rolling now... There’s more buy-in now than there was. (3-2-10 pp. 6-7)

The team also looked at student growth on the NWEA. While kids had always been tested to determine growth targets, their teachers had not always analyzed the scores by skill. Eleanor described how they analyzed the NWEA information in SY2009-2010 to monitor their progress towards their focus areas,

We took the winter scores and the fall scores and compared them to see the growth...Our focus was numbers and operations and critical thinking in reading. We have made growth in numbers and operations but every other area had grown even more and was at the norm. We were right where we needed to be. And that was like a big ‘Yeah!’ Numbers and operations was still a little low so we used that information and did some brainstorming and problem solving about what was going on. Well, part of it is the kids don’t know their facts…and steps. And so we decided we needed to focus on multiplication and division relationships. A lot of the kids don’t understand that there is a relationship there. (3-2-10 p. 16)
One of the goals was to use the information gathered throughout the year to differentiate the instruction. NWEA provided lists of specific instructional skills to be reviewed, instructed, or introduced in each category (e.g., numbers and operations, measurement, algebra, geometry) based on the students’ scores. These lists, called DesCartes, could help teachers create flexible groups for targeted instruction. Eleanor related how teachers reacted when she demonstrated how to access this tool to her teammates,

A lot of people will do the NWEA test, take their kids to it and then not doing anything with the data because they don’t know what to do with the data. So at that meeting I was able to show them DesCartes and it was like ‘Whoa!’ They pulled specific things from DesCartes for Intervention Wednesdays…It just made the job so much easier and that is part of the efficiency piece. (3-2-10 p. 18)

The teachers often had resources available to them that they either did not know how to access, did not have time to access, had not yet learned how to integrate into their instruction, or did not see the value to their instruction. The tools related to the NWEA Measures of Academic Progress assessment were one example of a tool that had been there for years and despite initial training had not been fully implemented and used to impact instruction.

A review of the team’s PDSA worksheet and interview with the grade-level chair revealed that team members had acquired some data and assessment related vocabulary and were striving to apply the concepts but were not yet fully clear on how to analyze and report the data. When asked how they were measuring student progress, Marion, the grade-level chair, responded,

We’re using the SCAs (Short Cycle Assessments), the MQAs (Math Quarterly Assessments), the NWEA. You know, we’re using all of it because we are triangulating those so that we can focus in on the weakest areas and try to develop that. (3-11-10 p. 5)
However, the team’s PDSA worksheet revealed that no one measure was reported consistently and different measures were mentioned in different review periods. At the baseline review dated August 27, 2009, the team noted in the PDSA plan that “students scored an average of 1.9 out of 5 points (38%) in Number Sense” on the SY2008-2009 Standards Based Assessment. It was not clear whether the data was from the then current cohort on the 3rd grade assessment taken the previous year or the data of the previous cohort of 4th grade students. No other data was mentioned. The goal was stated as “4th grade students will achieve 80% accuracy on bimonthly assessments by 10/15/09” which implied that 100% of the students would achieve 80% accuracy on the unnamed assessments.

At the October 20 review, the team cited the same SBA data as they did in the baseline window. However, in a different section on the PDSA worksheet, they stated “Baseline data showed 24% of the students are proficient. Current data reveals that 54% are proficient. This is a 30% increase in the number of proficient students.” The plan did not state on which assessment or specific skills this statement had been based. The review did not mention the students’ average fall NWEA scores for number sense although the testing window had been completed by that time.

At the January 25 review, the team stated, “Based on the 09-10 NWEA, students made a 4.2% gain from fall to winter in number sense. This, however, is still below the average norm by 1.4 points.” Most likely the team meant that the mean student score had risen by 4.2 points. Furthermore, the PDSA worksheet stated “Baseline data showed 24% of the students are proficient. Fall to winter NWEA and MQA scores show that there is a 70% in the number of proficient students.” The worksheet did not specify the percentage of students who achieved proficiency for each assessment. Considering that the assessments
were not well aligned, it was highly improbable that the same percentage of students would be deemed proficient on both assessments. Furthermore, the NWEA did not specify a proficiency level. It was unclear what RIT level the team was using to define proficiency. The team might have used the mean score from the national norm group or it might have used a score that represented a high probability of passing the New Mexico Standards Based Assessment. It was also unclear whether the team looked at the percentage of students that met their proficiency criteria in number sense or in mathematics overall. Finally, the Math Quarterly Assessment in SY2009-2010 tested different standards each quarter based on the curriculum map. Thus, it would have been possible for students to be considered proficient in one quarter and not be considered proficient on the different standards tested the following quarter.

At the March 15 review, the team stated “SCA and NWEA assessments reveal that fourth grade students are still 5% below the proficiency level in number sense and basic facts…NWEA and SCA data shows that 75% of fourth grade students are now proficient in number sense. This is a 5% increase from last quarter. Overall, there has been a 21% increase in proficient students since the beginning of the school year.” A 21% increase to 75% would mean a baseline of 54% which is the number cited in the October review period. It is unclear whether the 80% proficiency goal referred to 80% of the students would score 80% or higher or a different goal of 100% of students reaching 80% proficiency. Furthermore, the team could not have received new NWEA data by March 15 since the end-of-the-year testing window did not begin until mid-April.
During the end-of-the-year review, the team stated,

There has been a big improvement in the number of students reaching proficiency in number sense. However, 4.5% of the students are still below the 80% proficiency level. The grade level feels that the root cause may be that the students were just over-tested and they were tired. We will make the 5th grade team aware of the progress made in number sense. With only 4.5% of the students’ non-proficient in this area, the 5th grade team can help with interventions that will move them from 75.5% proficient to 80% proficient. From the assessment results the fourth grade students showed increase gains from 38% proficient at the beginning of the school year in number sense to 75.5% proficient in the 4th quarter.

The 38% mentioned in the baseline window was the average percentage of points earned by students on the SY08-09 SBA in number sense. Since the scores from the SY09-10 SBA had not yet been received, the comparison must have been to different unnamed assessments. Thus, the document review showed that the fourth grade team needed more support in learning how to analyze and report the progress of their students towards meeting their common goal.

Like other grades, the fourth grade team implemented grade-level wide interventions in SY2009-2010. Unlike the third grade and the fifth grade, the fourth grade team did not have prior experience with grade-level wide interventions and the idea did not originate from within their team. This is how Marion, the grade-level chair, described their intervention process,

This year we had a big change with interventions. Because we didn’t make AYP, we had to figure out how we were going to focus in on the children who were being left behind. We decided that we would level them off and have teachers teaching low groups, medium groups, and high groups. And we would just focus in on what they really needed help on. (3-11-10 p. 5)

Eleanor described the reaction of many of the fourth grade team members to the requirement that they collaboratively provide intervention to the students,
I would say the majority of people don’t like doing it but what I see as the big learning experience and benefit from it is common planning…They are making some baby steps towards that. There still is resistance but not nearly as much squawking as there was when we started…They’re just accepting that this is what we’re going to do and they have gotten efficient with it. You know, when you do something new there is a lot of circling and missteps, but we’re much more efficient with it now. (3-2-10 pp. 12-13)

By March when the teams were reviewing the benefits and drawbacks of the current model and possible changes to the intervention model for the 2010-2011 School Year, Marion was able to report this reaction from the team,

This was the first time doing a plus/delta that we had like twelve pluses and only two deltas which was wonderful because every other time it was always the other way. (3-11-10 p. 7)

Marion also believed that the interventions prompted some teachers to shift towards a greater alignment between instruction and grade-level goals. She stated,

We’re seeing some progress. We’re also seeing that the teachers that weren’t really taking teaching seriously, if I may say that, I think they were forced to follow a path that they needed to go down instead of ‘Let’s do this worksheet today’…It had nothing related with the curriculum. So we’re seeing some progress. (3-11-10 p. 6)

Thus, while the mandated intervention process was painful for team members, it did produce some beneficial shifts in instructional planning and team members slowly began to accept the practice of interventions.

Marion was a seasoned teacher with more than fifteen years of experience as a teacher. Yet facilitating this team of adults without prior training was challenging for her. Some of the challenges she faced were common to all team facilitators. She stated,

This year has been a learning experience for me…I have had to learn how to work not only with different personalities but also different teaching styles. (3-11-10 p. 7)

However, some of the challenges may have been unique to the dynamics within this team. Given the lack of collaborative history, some of the team members were recalcitrant to
cooperate which placed the grade-level representative in the role of enforcer although the representative was a peer and not an administrator. Marion described how this made her feel,

I have to push this group because if I don’t push them, they don’t hand things in. They don’t get things done. So I have kind of felt like a mother hen this year and I hate that. (3-11-10 p. 7)

Despite the challenges, the fourth grade team did undergo a significant transformation over the course of the school year. They began to engage in collaborative planning and looking at how best to measure the progress of their students towards common goals. However, the most significant change was in the team dynamics. Marion stated,

For the most part at every PLC I have seen more and more teamwork…It was a bad situation, but I can take a deep breath now and actually smile without grinding my teeth like I used to. (3-11-10 p. 11)

Hopefully the shifts in SY2009-2010 significantly changed the culture within the team so that they might continue to collaborate to produce higher student achievement in the future.

The First Grade Team

The first grade team had also encountered many challenges in shifting to a culture of collaborative evaluative inquiry. The teachers had long operated as independent operators. This was how a member of the first grade team, Liz, described the way the team functioned in February of 2009:

We have a big group of strong teachers who do things the way that they do them. [laughs] Do you know what I’m saying? (2-24-09b p. 16).

She elaborated that the responses of teachers regarding students who had not yet mastered the standards varied among the teachers and that team members were not aware of what strategies other teachers used (2-24-09b p. 9).
The culture of independence was so strong that the team members had questioned the need to collaborate. According to Liz,

I think part of it too is the shift to the PLC in general. It’s the shift to having conversations all the time. I don’t know that everybody wants to have a conversation all the time about what to do. (2-24-09d p. 2)

The hesitation stemmed from more than just annoyance at having to take time away from their personal preparations for instruction. The deprivatization of practice was very threatening to team members who fear they may be judged as incompetent. Liz explained,

Teaching is very personal (2-24-09b p. 13)…Instead of looking at the PLC as a group that’s supposed to help everybody, they don’t see it that way. They are afraid to talk about what they do or they are afraid to ask for help because they think that that reflects badly on them…They don’t ask ‘What are you doing? Your scores look really good’. They don’t want to know. They just say, ‘Well, this is what I am going to do’. (2-24-09b pp. 11-12)

Looking at data was especially threatening to the team members. Liz elaborated,

Unfortunately there are a lot of the teachers who feel like when they put forth their data that it looks punitive to them. So when a lot of times if somebody is not going to be at a meeting, they don’t want to leave their data without being there because they are afraid that data is going to make them look bad. (2-24-09b p. 10)

Changing the manner in which teachers interacted with each other tapped into deep emotional issues like self-efficacy and values such as autonomy of practice.

A culture of evaluative inquiry required the acknowledgment that no matter how good a teacher was, there were specific areas of instructional practice that could have been improved. Even if most students in the class had met expectations, the teacher must have believed that a tweaking of practices could have led either to all students meeting expectations or some students performing at a higher level. Liz further described the first grade team in February of 2009,
It’s not that what they do is bad, but they don’t see what they do as needing anything else...They still want to look at the big picture...They can’t take out a piece and say ‘What can I do to fix this piece? The picture is good but we can make the picture better if we work on this piece.’ (2-24-09b p. 17)

Even if team members were to see the need, it cannot be assumed that they had the assessment literacy to know how to do so.

However, conversations in a vertical articulation meeting had prompted Liz, and possibly others, to begin to see the need for collaborative dialogue.

It was quite interesting to hear the different opinions about what people thought number sense was. As a group we didn’t all necessarily have the same view about what number sense is. So when you’re trying to align things and do things, if you are not even talking about the same thing, how can you fix that? (2-24-09b p. 22)

Defining vaguely worded standards was important in the common planning process.

Several factors combined to prompt significant changes in the culture of the first grade PLC during the 2009-2010 School Year. First, the principal appointed a new grade level chair. Anne was a relatively young teacher in her twenties with less than five years of teaching experience. However, she had obtained a master’s degree and was eager to learn how to facilitate the team and create a culture of evaluative inquiry.

Second, there was significant turnover in the summer of 2009 and several teachers from different grade levels moved to the first grade team. Wendy, one of the teachers who moved from a different grade level to first grade said that the year started as “very old school versus new school”. (2-24-10, p. 6) Anne described it as a brand new team that did not start with an existing foundation of trust. The requirement to collaboratively assess and provide grade-level wide interventions brought philosophical and personality differences to the forefront. The first half of the year was very challenging for team members. Anne said,
It seemed like anything presented from anyone would be shot down. It seemed to get really personal, like personal attacks... People were just really scared to talk and people from the grade level would come to me afterward and tell me how they felt. I went to the administration saying how everyone was feeling and how we weren’t going anywhere... I think that subconsciously they were doing it to stop the process because it was so different. (3-22-10 p. 14)

Over the course of the year the dynamics were able to be shifted through a combination of unplanned as well as planned forces. First, there was a shift in who attended the PLC meetings which changed the interpersonal dynamics. One team member described what happened,

One particular member was gone for quite a bit of time and I think that helped begin to develop discussions for other people feeling like it was safe and then once this person came back in, it was, you know, well, I can talk over this. (5-17-10 p. 8)

Also, school administrators and other grade level colleagues talked to some of the most domineering personalities and helped them recognize the impact that they were having on the team dynamics. Second, the new grade level chair, Anne, became more assertive. One team member described how her leadership style evolved,

She had to stand up to some people. And that was really hard for her but when she did and realized that the world was not going to come crumbling down and that it created... something that was beneficial, it was like ‘Oh, I can do that too’.

Finally, the school principal and grade-level chair invited me as an outside facilitator to come in to facilitate a series of meetings in November and December. Anne also invited Sharon to attend the team’s meetings as a participant observer so that she could provide immediate clarification if clarification questions arose.

The goal of the outside facilitation was to depersonalize the conversations around the data and goal setting. Utilizing tips from the Data-Driven Dialogue training, I compiled data from all of the classrooms and presented only grade-level wide data. The first session
focused on math achievement since the team was less divisive when discussing math goals and assessment. Teachers, student teachers, ancillary teachers, and aides were purposefully mixed within three different groups to ensure a mix of roles within a group to provide different perspectives within each group and to change the normal conversation dynamics. Each group predicted the percentage of students across the grade level who scored at the Advanced, Proficient, Nearing Proficient, and Beginning levels during the baseline, Week 6, and Week 9 testing windows. Then teams were asked to list the assumptions behind their predictions. After the assumptions were discussed, the teams compared their predictions to the data and discussed salient points, surprises, and possible inferences based on their observations.

At the next session, the team started by using a text protocol to discuss the differences between advocacy and inquiry and the differences between dialogue and discussion. The team then discussed sample norms of collaboration from a variety of sources including MBTI personality preferences and the implications of those preferences for collaboration styles. Finally, the team analyzed the first quarter Developmental Reading Assessment data and identified student needs. The team was supposed to identify the three to five most important needs and prioritize them based on importance and implications for teaching those skills during the intervention block.

At the third session, the team had to decide whether to focus on comprehension or word analysis skills for the PDSA goal and common formative assessments. It was decided that word analysis would be easier to manage both in terms of differentiated instruction and in terms of formative assessment. The team then discussed what proficiency would look like for a student at the end of the year. Although the team did create a vision of proficiency for
reading and decided how those skills could be assessed, they did not follow through with either implementing the reading intervention or measuring student progress towards the common reading goal.

While the team did not follow through with setting reading goals because the process was considered too overwhelming, the process of having an outside facilitator did appear to help the team dynamics. Anne stated,

> I think it did make a difference, just slowing everything down and breaking it back down to the basics, and you know, just the basic professionalism. I think that did help in the long run…It definitely helped that it was an outside person. You know, just changing the whole tone. (3-22-10 p. 27)

Wendy concurred, “It was really nice having other people come in sometimes and we needed that.” (5-17-10) By the spring semester, the focus appeared to have moved beyond interpersonal dysfunctions to collaboratively focusing on how to best meet the PDSA goals and provide interventions in a manner appropriate for first grade students.

The idea for the grade level-wide interventions did not originate with the team members and they struggled to find a way to meet the mandate in a manner that they felt was beneficial for first grade students. Wendy explained their perception:

> There are some major concerns about not making AYP and we have had some knee jerk reactions…This year’s focus was on interventions which had worked very well for a couple of grade levels. So it was ‘Let’s do it for all grade levels’ without really taking the time to sit and think, ‘How does this really work?’…It was trying to massively make everybody do the same thing at the same time. (5-17-10 p. 3)

The lack of buy-in from team members led to subtle passive-aggressive sabotage. For example, one teacher admitted that,

> There were some people who just said, ‘Oh, you guys are here. Go play. Go play Bingo. Go do this.’ (5-17-10 p. 5)
Anne had the unenviable task of trying to move the team forward,

I think at first there was a lot of push back and I was trying to keep us going ahead because it wasn’t a choice and we had to make the most of it. (3-22-10 p. 10)

One frustration for team members was the lack of clear parameters from the administration as to what extent the model for the schoolwide intervention block could be modified for the first grade students. Wendy said,

I think we ran around in circles a lot chasing the thing we were supposed to do rather than how this is meeting the needs…Everybody was frantic with what was going on. (5-17-10 p. 4)

While the school administrators did allow the team to modify the intervention model to better meet their needs by letting the first grade students meet for shorter periods of time, the first grade team still felt that the model was an outside mandate that was not in the best interest of their students. Anne explained,

We all understand that whatever we’re going to do is not going to please everyone. But it’s just, I think, a general feeling that the needs of the students are being left aside for the sake of us pursuing interventions even if it is not working…I don’t think [students] benefitted directly from interventions. I actually think that it hurt them more because there was not enough time to fully establish the system so a lot of it was teachers trying to come up with something last minute. (3-22-10 pp. 11 & 21)

She elaborated on their frustration,

I think everyone got to the point of acceptance and really did give it an honest try...We know how hard we are working. And we know what we’ve tried. And we know what hasn’t worked. And we don’t know what can work because we haven’t experienced anything that can yet except for what we have done in class. (3-22-10 pp.10 & 21)

Ultimately, the first grade team stopped doing grade-level wide interventions altogether by early March.

While the first grade team struggled to find value in the grade-level intervention model, by the end of the year, many members of the team did find value in having to
commonly assess student progress towards the shared PDSA goal of learning math-related vocabulary. Anne shared her view of the experience,

In previous years it wasn’t really a shared writing of the PDSA. And if we did talk about it during a PLC, it was more of a quick ‘What do you think about this?’ It just seemed like putting something together that had been in place for years and years and we didn’t really think about, have a dialogue and discussion about it. But this year we all came up with it together and looked at the data and looked at our own kids and saw where the needs were... We talked about [the PDSA goals] every week. And we would look at the data every three weeks. It held everyone accountable. Everyone felt I really do need to teach this vocabulary and target it in the classroom. I think that the teachers knew exactly what to focus on. (3-22-10 pp. 12 -13)

Anne described the process that they used to review the students’ progress towards the PDSA goals and its impact on the teachers,

We’ve looked at the data [from the common formative assessments] every time that we said we would. Everyone stayed on track with reporting their data to me. And then I put it into a graph the way that you showed me and that has been really effective. Because we can see where the kids are and we’ve had good discussions about is this a good goal. It’s really more of us looking at our teaching and if that is a good goal. (3-22-10 p. 3)

During the first quarter, Anne had created spreadsheets after each common formative assessment had been administered anonymously showing all of the scores of the students in each class in rank order. The percentage of students proficient in each class was summarized at the bottom of the spreadsheet. The spreadsheets did not show proficiency levels from prior testing windows. After the outside facilitation in November and December, Anne utilized a new format which showed the percentage of all students within the grade in a stacked column graph where each column represented a testing window and the percentage of students at or above the proficiency level over time could be interpreted quickly.
Wendy further described how the PDSA collaboration process helped the first grade teachers monitor students’ progress once they succeeded in developing assessments that accurately measured the goal,

Our PDSA goal for math really revolved around the importance of number sense [and] the importance of vocabulary. It was hard because people wanted something fast and easy. When we really sat down, [we discovered] this test does not accomplish that…When we really developed a test that spoke specifically to the vocabulary, then people began to say, ‘I didn’t realize that my kids didn’t know this’…People were saying, ‘You know, I really saw a difference with this’.

The discussions also had an impact on the team dynamics. Wendy explained,

It worked out in the end and I think in some ways for our grade level it pulled people together more…People had to start using the same vocabulary. People had to start saying ‘What do you mean by that?’ Um, ‘I’m not sure how to accomplish that.’ ‘What is it that we are specifically looking at?’ So I think it had to get more detail-oriented. Did it help the kids? No. Did it help people begin to speak a little bit more? Sure. (5-17-10 p. 5)

While schoolwide student data suggested that reading may be the area of greater need, the choice of focusing on math vocabulary helped the team’s productivity given its early stage of development. Wendy explained,

Because we had something so focused to talk about, which vocabulary is necessary to make this connection, it was safe ground for people…Vocabulary was very specific, very precise, and people could go ahead and open up on that…rather than just how do we teach reading. (5-17-10 p. 8)

For this team, having a successful experience with one content area was more important than covering both content areas and continuing to struggle with the process.

While the first grade team struggled to create a manageable PDSA goal for reading and to find a grade-level wide intervention model that they believed would benefit the students, the team did make significant progress in their team dynamic and willingness to
collaborate. By the spring, both Anne and Wendy were optimistic about the future of the team. Anne shared their status as of March, 2010,

Everyone can listen to one another now and everyone seems to want to participate now. It used to be more that people really just sat back and didn’t really say anything unless they were more aggressive naturally in their personality and they would always say something more negative toned instead of productive or solution oriented. And now it seems like everyone is providing feedback that is more on the productive side and less on the negative…It’s not like it used to be where it was just trying to stop things. (3-22-10 p. 1)

At the end of the year, Wendy concurred with that outlook,

In the end I saw a real growth in the grade level. We still have several people who are willing to sit back and say, ‘I don’t care, do whatever you want to do’. But I saw some people that I haven’t heard from for quite some time step up to the plate and say, ‘You know, I have always wanted to do this’ regarding whatever it was we were talking about…It wasn’t just status quo…There has been some peace. There has been some quality. There has been some ideas of this is a direction that I can go in. (5-17-10 pp. 7 & 8)

She concluded,

We’re not seen as the team last year [where] you could feel the tension…That’s not to say it is still not present, but it’s not the team that nobody wants to be on anymore. (5-17-10 p. 8)

**The Bilingual Program**

The performance of bilingual and ELL students had long been a concern at the school. In 2009 Sharon noted that bilingual and ELL students at another elementary school in the district had achieved higher gains and higher levels of achievement. Sharon and the school’s bilingual coordinator, Leslie, were concerned that the school’s bilingual program was not as effective as it could be. Sharon noted some of the issues in June of 2009,

There’s not a lot of vertical alignment. There’s not a lot of collaboration. It’s just there. And then you’ve got different philosophies. (6-5-09 p. 1)
The school had utilized an enrichment model in which both native speakers of Spanish and
native speakers of English could volunteer to participate in a classroom with a bilingual
teacher and receive some Spanish language instruction from the bilingual program
coordinator. Sharon explained some of the unintended consequences of having an open door
policy,

People were doing teacher jumping. They’d start off in Kindergarten and they’d like
the teacher. When they went to first grade, maybe they didn’t like the first grade
teacher so they didn’t want to be in [the bilingual] program anymore. Then they went
to second grade and liked the teacher. So the bilingual coordinator was dealing with
people who got to fifth grade as brand new enrichment speakers because they had
never had it before. (6-5-09 p. 4-5)

Both Sharon and Leslie reviewed research on best practices. They concluded that a different
model might be more effective. Sharon described what they envisioned,

[We are looking at] the additive model where you would start with kindergarten and
you would have truly monolingual speakers with their core reading and math
instruction in Spanish. It would be kind of like a bilingual inclusion model…You
would have a TESOL endorsed teacher in one classroom and would have a bilingual
certified teacher in the other classroom. Both the groups would mix [for the non-core
subjects]. (6-5-09 pp. 1-2)

Sharon and Leslie discussed how they might be able to begin shifting towards this model for
the 2009-2010 school year. Sharon described how they proceeded,

We thought about this really carefully and we thought if…we do this whole scale
right now, it’s not going to work because you’re going to have the pushback. We
really need to see if it is going to work so we targeted two teachers…They’ve got
good developmental backgrounds. They are skilled teachers. Their scores
improve…[One is] TESOL endorsed and they team well together. So we approached
them with what we wanted to do. We talked to them about how we would support them
because it would also include both our ESL and our bilingual person supporting them
in materials and things like that.

They needed to think about it. They asked us questions; they asked us for research.
[The bilingual coordinator] spent time with them. They said they would like to try it
but they’re not comfortable doing it this year because they would like time to look at
it and get ready. So here’s the quandary. We’ve got two teachers who really could
do it but they’re scared to stick their toes in the water. And then we’ve got a group of
kids who really cannot wait a whole year. (6-5-09 pp. 2-3)

The teachers had had several experiences which made them hesitant to embrace a
maintenance model for the bilingual program. Sharon explained how the poor
implementation of a previous maintenance program at the school had shaded teachers’ views
towards this type of model,

[Under the tenure of a previous principal] they did a 90/10 model where the
Kindergarteners were in Spanish 90% of the time and in PE, art, and music in
English. But these weren’t just monolingual kids. These were just kids and then on
top of it, the teachers that they got spoke Spanish but they weren’t skilled. They
didn’t have a good curriculum. It was a nightmare. We’re still seeing the residual
results…So [the bilingual teacher we were targeting] said, ‘I saw it was a nightmare’.  
(6-5-09 pp. 3-4)

Furthermore, the teachers had also had some negative experiences themselves growing up as
bilingual Hispanic students. Sharon described her conversation with one of the teachers
about her experiences,

She is probably in her mid-twenties. She’s not that old but she said that when she
went to school, ‘If we spoke Spanish we got made fun of. And I didn’t transition
over to second grade [where English became the primary language].’ She said, ‘I just
don’t want to do this to these kids. I don’t want to hurt them.’ And I said, ‘I realize
that, but what we’re doing now is hurting them because they are not getting the gains
that they should and so we’ve got to do something different’. (6-5-09 p. 4)

Given the hesitation of the staff members who would be involved, no changes were made to
the bilingual program for the 2009-2010 school year.

The need for change became even more apparent once the scores on the statewide
standardized test came back in July of 2009. For the first time, the ELL subgroup did not
make Adequate Yearly Progress. Not only did the percentage of proficient students fail to
meet state targets, but the percentage of proficient students dropped by over 20% from the
previous year. Only four of thirty-one students identified as current ELL students, or 12.9%,
were proficient. Five of ten students who had exited the ELL program within the past two years were proficient.

As the Instructional Leadership Team discussed the data at a retreat in August of 2009, it became apparent that the purpose and design of the ELL program and the bilingual program were not clear to the majority of members on the Leadership Team. The team asked questions about which students were members of the ELL subgroup, which students received instruction in English, and which students receive instruction in Spanish. Diane, a third grade teacher in the bilingual program, mentioned that only one of the twenty-two students in her class the previous year was considered to be an ELL student. The others were considered enrichment students. The team also discussed the competing demands on the time of students who qualify for ELL, bilingual, and literacy programs. Some students also had an identified disability and received Special Education services.

The concern about students in the bilingual program continued as the Instructional Leadership Team examined the progress of students on the intervention wall in November of 2009. As Sharon looked at the cards of students who had stayed at the same reading level she noted,

> Whatever is going on with bilingual, we are doing something really wrong…I think that is where we need to get our arrows aligned. You look at those children and they have had four or five different interventions and they are still stuck. Somehow we’re not meeting their needs. (11-4-09 pp. 3-4)

Thus, data from several different sources consistently confirmed that the needs of at least some of the ELL and bilingual students were not being well met.

In order to be prepared to make significant changes in the 2010-2011 school year, a bilingual/ESL PLC was created in December of 2009. Sharon used bilingual funds to pay for
substitutes so that fourteen staff members could meet for a half day once a month. At the first meeting, Maria, the district Director of Federal and Bilingual Programs presented a “big picture” profile of trends of ELL students nationally. Then teachers presented district and state assessment data which, according to Leslie, showed the school’s ELL students “tended to be lower in class ranking in terms of their academic performance”. (5-19-10 p. 1)

At the second meeting, an outside facilitator who works at a government research facility facilitated a causal analysis using the Apollo Root Cause Analysis method. The stated problem was “There are too many English Language Learners in grade 3 that are not meeting grade level academic proficiency standards” (1-11-10 p. 1). The team identified the following costs if the problem was not addressed: children’s long term success, failing to meet AYP, negative reflection on the district, loss of control over the bilingual program, loss of teacher freedom, mandated curriculum required, and reduction in working conditions. The team identified five root causes: program structure and staffing, student placement, curriculum, parent and community involvement, and communication.

According to Leslie, the bilingual program coordinator, the team initially felt that there were not enough staff members to meet the needs of the students. However, through their discussion they realized that some TESOL-endorsed teachers were not servicing ELL students. Therefore, the team sought to create a structure to better utilize the existing staff members on the campus. Their goal was to place one bilingual and two TESOL teachers at every grade level. One bilingual and one TESOL-endorsed teacher would work together as a team to provide services for students who spoke Spanish as their first language. The third person would provide TESOL support for ELL students who had a language other than
Spanish as their first language. This structure would allow for Spanish speakers to receive their academic instruction in their primary language first. Leslie explained,

One of the big shifts is that we want to be able to provide them their reading instruction, their math instruction, their main academics in their primary language [during] their first couple of years [when they have] all that cognitive overload…So we are going to start that with our incoming Kindergarteners and then gradually phase that into our program. (5-19-10 pp. 3-4)

Furthermore, placing ELL and bilingual students into three classrooms per grade level would assist the ELL teacher and bilingual coordinator with collaboration and communication with the students’ homeroom teachers.

In addition to these structural and placement changes, the team decided that greater alignment in curricular materials was necessary. Leslie stated,

Some people are using the Comprehensive Literacy Model. Some people are using the core literacy program. So students weren’t getting necessarily sequentially based skills, vocabulary development, and all of that. As a group we decided that the most important thing was that we get streamlined with that as well. (5-19-10 p. 5)

The district agreed to use bilingual grant funds to purchase both English and Spanish versions of the newly adopted language arts curricular materials for the kindergarten bilingual teachers although the district could only afford to purchase the materials for full adoption in grades three through five. The use of one language arts series from kindergarten through grade five would allow for better vertical articulation and would facilitate conversations regarding how best to use the curricular materials and how best to instruct the students.

Another area of concern was parent and community involvement. The team’s goals included teaming more with the Title I staff’s parent outreach efforts, making sure that letters to parents were translated into Spanish, recording voicemail messages in Spanish as well as
English, placing welcoming Spanish language displays on bulletin boards near the entrance to the school, and a long-term goal of making sure that at least one front office staff member spoke Spanish.

Finally, a smaller committee met several times to create a plan to improve communication. The plan included ideas to improve communication within members of the bilingual PLC team, between the bilingual team and the school staff, between staff members and parents and students, between bilingual and non-bilingual students, and between the school and other schools within the district.

The process of creating a shared vision was not easy. Leslie described how one teacher struggled to reconcile her personal beliefs with a vision of teaching core subjects in the students’ native language.

We did have that discussion [about philosophy and shared vision] and there was one teacher who went through the entire year with us with these meetings and everything but at one point during one of the meetings she philosophically said, ‘You know, I just can’t agree with that’. A lot of it has to do with her personal experiences that she had and she can’t support the primary language component of our program…She sees the Spanish more as an enrichment opportunity…She continued to come and I felt like we were really making progress in terms of her participation in the group. She contributed a lot of things…and I felt like her opening up and sharing was a huge step because in the past she has been sort of been behind the scenes not in agreement with our program but not really saying it out front. She is a bilingual certified teacher. [It] finally came to the point where since she is not buying into what the rest of the team feels is important for these kids, she decided that she didn’t want to be a part of this particular part of it. She is also TESOL endorsed…and she might just be that third TESOL teacher. (5-19-10 p. 7)

Teachers’ personal experiences and beliefs were so intertwined with their educational philosophies that sometimes collaborative conversations were not enough to develop a shared vision. In this case, the school leadership recognized that not all staff members could support
the vision and chose to strategically place the staff member in a role which the staff member could philosophically support.

**Status of the initiative at the end of the study**

*Tension between inquiry focus and other duties of teachers*

As the profiles of these five PLC teams showed, the dynamics within each team in the school varied significantly. However, there was some shift in the purpose and function of all of the profiled PLCs over the past few years. The shift was not complete and the teams were at different stages of development in making the shift. One challenge was to create a shared vision regarding the purpose of the PLCs. Even at the end of the two years of this study, not all staff members agreed that data should be the focus of the team collaboration time. In an Instructional Leadership Team meeting in March of 2010, Sharon said to the team,

> We got feedback that said ‘PLCs were doing too much data’. That’s what PLCs are about. It’s looking at what the needs are of your kids. I was a little confused about that because I was thinking, ‘What else are you supposed to be talking about besides the kids?’ (3-3-10 p. 4)

One of the frustrations for teachers was that they felt that time had been taken away from discussing necessary logistics in order to focus on data. The teachers’ schedules were so busy that they resented losing the time to discuss “business” items. When asked during a retreat in 2009 if there was anything that could be done to make sure that business and housekeeping items weren’t addressed during PLC time, the response of a teacher was,

> Not without a lot of anger. You would have a lot of pushback because then we would have to have another meeting. That can’t be done through e-mail. They won’t read it. (5-28-09 pp. 13-14)
The degree to which business or housekeeping items needed to be discussed varied throughout the year. According to a teacher,

There were times where none of [PLC meeting time] was [discussions about housekeeping or business items] and there were times like coming up on a field trip where most of it was. I mean it varied depending on time of year and what was going on. (5-28-09 p. 13)

Observations of PLC meetings showed that the entire meeting was sometimes taken with topics like how to spend grant funds or trainings on how to use the new electronic grade book.

Indeed, even the principal had a hard time keeping business and housekeeping items off of the agendas of the Instructional Leadership Team meetings. A review of agendas from SY2009-2010 revealed announcements about clubs and programs (e.g., Mileage Club, student government, art program, Family Backpack program, Reader’s Club) and events (e.g., Walk/Jog-a-thon, PBS assembly, Literacy Night, Turkey Bingo, Donuts for Dads, staff party, Multicultural Day); distribution of recess equipment; and queries about textbook and supply needs. Furthermore, the ILT meetings provided a structure for grade level representatives to bring up grade-level concerns to the administration and other grade levels. A review of ILT meeting minutes from SY2009-2010 showed that concerns about duty schedules (e.g., playground and parking lot), custodial work, administrative memos, fire marshal reports, and other items also were discussed in ILT meetings in addition to concerns and questions about the intervention program and continuous improvement plans.

The Instructional Leadership Team discussed how to structure PLC meetings in order to make them most productive. Suggestions included setting a timer to discuss improvement issues for the first fifty minutes of the PLC meeting and then limiting discussion of
housekeeping items to the last ten minutes. The team also discussed the suggestion of the outside consultant, Patricia, that PLC teams should create a cycle of meeting activities to include:

- Data meetings for review of relevant data in meaningful ways for instructional planning and decision making at each grade level;
- Instructional planning and sharing of strategies and interventions between grade level teachers;
- Resource investigations and sharing regarding additional strategies, approaches, interventions, for use in classrooms;
- Quarterly review of the PDSA plan. (2-24-09)

However, some teachers had concerns that mandated structures would hinder the autonomy of the teams and would ultimately hinder progress. During the discussion one teacher asked,

Do we structure it so much that it is no longer useful?...If we try to unify everybody to do the same thing at the same meeting, I don’t think it’s going to work. (5-28-09 p. 15)

Time was a very limited resource at the school and it was hard to create the time for evaluative inquiry without negatively impacting other functions within the school.

*Systems and structures to support a culture of evaluative inquiry*

Another challenge was developing systems to manage the data in an efficient and user-friendly manner. Traditionally some staff members had relied on Excel to manage their data. In fact, Sharon described how the school planned to provide background data to teachers about their incoming students,
[The secretary] is really good at Excel…We’re going to pull in the class profile data so when teachers walk in in August they are going to have their class profile data for their kids from spring. (6-5-09)

The school district recognized the need for a district-wide data management system. In the fall of 2008, the district offered to provide training on a new locally-developed web-based data warehouse called the Student Assessment Portal. In addition to being able to access SBA and NWEA data, teachers were also expected to enter DRA and Math Quarterly Assessment data so that a longitudinal profile of student growth could be developed. Due to a lack of time, Sharon and Wanda decided that they would train the teachers how to use the system themselves. The tool was only partially implemented during the 2008-2009 school year. Diane, one of the most data-savvy teachers, stated in February of 2009,

All I do is input data [into the Student Assessment Portal]. I don’t have any access to manipulate that data to tell me things about my classroom this year. (2-24-09a p. 6)

This statement indicated that teachers had only learned enough to comply with district requirements of inputting DRA and MQA data, but they had not learned enough about the system to use the data to make instructional improvements.

In the 2009-2010 school year, the district shifted to a more robust, commercially-developed web-based data warehouse and assessment management system from Data Driven Classroom, Inc. The data warehouse provided a variety of achievement and growth reports for both current and former students and provided longitudinal profiles of individual students. The assessment management system allowed teachers to quickly score assessment results and analyze student achievement by standard and by item. Although the system was available by December of 2009, training was not scheduled for teachers at North Mesa.
Elementary School until April of 2010. None of the teachers truly utilized the system until after the scope of this study.

In addition to finding the time for data analysis and becoming proficient with a system to manage data, the teachers also needed to learn how to integrate data analysis into their work in a psychologically safe manner that honored the colleagues’ needs to relate to each other. Leslie, the bilingual program coordinator, commented,

The touchy-feely part I have down. It’s the cold data part [I have troubles with]. (5-19-10 p. 9)

Leslie described how the protocols she learned at the Data Driven Dialogue conference could help her team,

In the ‘real world’ we don’t have time for sharing…[like] telling about our pets…[It’s] just get to work, but there is a way to do the work in a very inter-relational kind of way…I felt like I really learned a lot and am kind of ready to take the team forward next year. (5-19-10 p. 9)

Given the low levels of trust within many of the teams, tools needed to be provided to help team members develop relational resilience while examining and using data.

Communication and relationships

Communication and relationships were recurring themes in many staff members’ comments and the two concepts were intricately connected. Addressing staff members’ concerns about perceived insufficient communication from the school administrators, William, the Assistant Principal, explained how the administrators sometimes consciously thought about the message that their actions would send and how that could contribute to relationship building,
The communication thing. I understand, but it’s not for a lack of effort. [For example], Sharon and I were talking about Krispy Kreme versus Wal-Mart donuts. I’m a guy. You bring donuts, I’m happy. But apparently Wal-Mart donuts are not quite good enough. You know, Krispy Kreme says we love you and Wal-Mart says we kind of like you. We talk about that stuff and it sounds silly, but it’s what we do. (2-11-10 p. 23)

However, some of the choices that administrators had made over the previous few years may have inadvertently contributed to changes in the dynamic among staff members. Some teachers noted,

[We] don’t feel as close to people as in other years…Even when we were a big campus, you at least communicated a little bit more than we seem to now. We rarely see each other as a group whether it is for social [or] creating vision and policy or anything…There’s not that camaraderie. (5-19-10 pp. 20-21)

Sharon concurred that there had been a shift in the dynamics. She stated,

Something that has bothered me greatly is the lack of social connections that we have had here and I’m not sure why…I know things have changed. People are financially stretched. Time-wise they are stretched. But you need to make connections with people here and I think we’re at a stage where we are small enough that we can. (6-15-10 p. 12)

She continued,

Honestly I don’t know the root cause. I mean, I think we’ve offered things. I think if you go back and look, the expectations have ramped up. That is systematic and also deliberate. The school under [a former principal] is whole different school that it is under me. The expectations have changed. In some ways I think it is good because the needs [of children] are clearer. When you are really looking at stuff and you are examining data, the paperwork is tremendous and exhausting. I can’t be that happy person anymore that says, ‘It’s okay. Don’t worry about it’…I would like to do it sometimes. I look at their faces and I know it’s hard. It’s hard on me, but we’re doing what’s best for children. (6-15-10 p. 14)

Sharon’s desire not to infringe upon grade-level PLC meeting time or teacher preparation time led her to try to minimize whole staff meeting time and Instructional Leadership Team meeting time. Some topics that could have been discussed by the ILT team or the entire staff which potentially could have led to increased communication and a more
commonly shared vision were not widely considered. For example, a member of the Instructional Leadership Team noted in 2008,

> Everybody has gotten [the climate survey results taken by staff members and students after the training by Education for the Future] but we haven’t discussed [them] and we haven’t done anything with [them]. (10-14-08)

In 2009-2010, the school administered several different surveys as part of the state-mandated NM CLASS continuous improvement process for schools designated as “In Need of Improvement”. After collecting and analyzing perception data from staff, students, and parents, analyzing student achievement data and classroom observation notes, and self-reflecting, a team was supposed to rate the progress of the school on a number of criteria and set priorities for continued improvement. Sharon ended up analyzing the data by herself and reporting her conclusions to the members of the Instructional Leadership Team. When asked why she did not include the Instructional Leadership Team in the review process, Sharon stated,

> When I thought of NM CLASS I thought, ‘Okay, this is my review that I have got to do’…[The NM CLASS data] is probably something we should talk about at the ILT retreat and go through. I didn’t even think about that. I did talk to them about it but I told them what I saw. (6-15-10 pp. 21-22)

Sharon was aware that some staff members had expressed a desire for better communication about many issues including schoolwide improvement efforts. In a meeting with Instructional Leaders in 2009, she commented,

> Every year I have seen on evaluations on me [that] communication could be better and I think unless I stand on my head and run around naked in the courtyard, I can’t do any more. (5-28-09 p. 14)
She pointed out that despite the desire to focus meeting time on collaborative discussion rather than “business” or “housekeeping” items, some teachers often did not pay attention to the content of written communication,

> In a professional world, housekeeping could be taken care of through e-mail and read your staff memo and stuff but the problem is it doesn’t happen that way. (5-28-09 p. 14)

She further explained her strategy for communicating with teachers in interviews,

> I have tried [communicating] every way possible. We do e-mail. We do things in their boxes. We talk to them. But it’s just like with the kids. Not everybody gets the same message all the time or the way that you are delivering it. It just doesn’t come across the same way… I feel secure [that] I have tried every single way I can to communicate and validate what people need to hear. (6-5-09 p. 28 & 6-15-10 p. 13)

She concluded by comparing communicating with staff members to teaching,

> A lot of it is repeating and repeating and reiterating and check to make sure for understanding. It’s the same thing you do in a classroom…Teachers are a lot like children. They take a while to attain stuff too. (6-5-09 p. 28 & 6-15-10 p. 13)

The lack of time was also a factor with the quality of communication throughout the school.

---

*Planning process for continuing the initiative in School Year 2010-2011*

Given the number of concerns that had been raised about how the schoolwide intervention block had been introduced in the fall of 2009, Sharon and William decided to take a more inclusive approach to decisions about how to revise the intervention block schedule for SY2010-2011. They asked each of the PLC teams to provide feedback about the benefits and drawbacks of the current intervention model and to consider whether a different model might work better for their students. Some teams were skeptical about the process. One member of the Instructional Leadership Team asked Sharon,
When we get our grade level’s feedback, is that going to be taken into consideration? Because I think that’s where it comes from with the grade levels with the whole ‘We’re not included. Our voice doesn’t matter.’ Because it’s like if you have an idea in your head, that’s going to be it whether we give feedback or not. (3-3-10 p. 5)

The grade level chairs discussed the interventions with their teams and provided written feedback to Diane, the third grade chair, who compiled all of the comments.

At the next Instructional Leadership Team meeting, Diane shared the feedback with her colleagues,

I took everybody’s notes…where you gave the EPSS feedback and feedback on the intervention. I just made a list of what was on there…There were things like ‘It was not appropriate at Kindergarten but it was great at fifth grade’. It was very interesting the two sides of the coins. Some people said they got great positive feedback about it from the kids and then other people said the kids gave very negative feedback about it. There’s the list. Share it out to the staff so everyone can know what everyone else is saying. (3-3-10 p. 1)

Diane summarized three main concerns the teachers had about the current model. The first concern she cited was, “It was difficult to use the same model across many age and ability levels.” (3-3-10 p. 2) Natasha, the math interventionist, noted that too much focus may have been placed on the model and not enough on the goal of the interventions. She commented,

It’s like they have a model in their head and they think that everybody has to fit into it instead of coming up with what works for their kids. (2-22-10mi p. 21)

The second concern Diane cited was, “We did not have complete staff buy-in, agreement, or implementation.” (3-3-10 p. 2) Natasha believed part of the reason some staff members did not buy in to the intervention model was because,

You have people that were just digging in their heels because it was like, ‘I don’t want to do that.’ They are still of the mindset [that] ‘They are my kids, my room, my way’. (2-22-10mi p. 22)

Finally, Diane stated that teachers found the intervention model to be “very time consuming”. Sharon agreed that the interventions were very time-consuming. She stated,
It took a lot of time, something that I didn’t really anticipate at first…I felt they were working harder but maybe not working in the right way…I think it was worth the work but then sometimes I think it wasn’t…in terms of morale because they were tired. (6-15-10 pp. 6 & 7)

While the interventions did come at a cost, they also produced some benefits. Sharon noted,

I think the good thing about it was it got teachers talking about the children [and]…about the needs… As they started talking and they started thinking, the planning got deeper and more involved. (6-15-10 pp. 5 & 6)

After discussing the feedback regarding the current model, Diane introduced a different model for the teams to consider:

We’re looking at Model 3 for next year. Model 3 is a schoolwide daily intervention schedule. What that entails is…grade levels plan interventions and implement them according to their grade level needs and the data that is pertinent to their group. Okay? Some benefits to everybody using the standardized schedule are that there is protected instructional time every day for the teaching and then there is protected instructional time for the intervention as well. Every day. Every kid would get to go to daily specials five days a week instead of having Wednesdays off. We would have intervention blocks at every grade level during the day in addition to the daily instruction… It is not taking the place of instruction like it is for some people right now. (3-3-10 p. 2)

One issue that the school faced was how to provide intervention services such as language support or reading support without diminishing the amount of time the students received core instruction in their classroom. Cheryl, the Literacy Instructional Coach, explained the dilemma of how to best support students with multiple needs,

What has happened in the last couple of years with us too is that sometimes they ask us as literacy teachers not to pull [ELL students] as well. There’s the language issue but we want to help anyone who we feel is below [in reading] but then we have been told at times not to pull those kids. Getting pulled too much is part of the problem too. So we have to figure out what we are going to do. (11-4-09 p. 5)

With the proposed model, most intervention teachers would have worked with the homeroom teachers to provide targeted, differentiated remediation or enrichment to students during a set time so that students would not miss instruction in core skills while receiving pull-out
services. The intervention teachers would either push-in to a classroom or pull-out based on
the needs of the particular grade level. While pushing in to the classroom might have
contributed to greater alignment of instruction between the homeroom teacher and the
intervention teacher, not all teachers welcomed that model. As Eleanor explained,

Part of the issue is getting over people feeling threatened about having that other set
of eyes in there and feeling like maybe they might be critiqued when really you are
going to be focused on whatever your job is within the classroom. (3-2-10 p. 14)

Another potential drawback to the push-in model was that it would become more difficult for
the intervention teacher to service students from different homerooms at the same time.

The team discussed, based on their experience with interventions in SY2009-2010,
what parameters would be necessary in a new intervention model. Diane stated,

We have to have some common prerequisites and a lot of this we have built in from
all the stuff we have done this year. We have to have common pacing guides across
the grade levels before we could ever do that. Well, we pretty much do in Suburban
School District. We have common curriculum and pacing guides. We pretty much
have common formative assessments across our grade levels. Common standards and
definitions of proficiency and a way to approach and analyze the data that we get so
that we can make these groups and figure out who needs interventions. Questions
that we are going to think about as we move forward and there are lots more but,
- How have we and how are we going to provide additional support for
  students who are struggling in a way that is timely, systematic, and direct?
- How do we enrich and extend knowledge for those that need it and who is
  available to assist our teams?
So, um, who is available to assist our teams is a different question but those other
questions come from the four essential questions that we are supposed to be asking all
the time.” (3-3-10 p. 2)

Sharon added,

The one thing that can’t happen though is that you can’t…go ‘Okay, it’s 45 minutes
so I am just going to work in my group and do whatever.’ You still have to work
toward common goals. That’s the purpose of it. It is still looking at formative data.
It’s still looking at what your needs are. It may not be your PDSA goals. You may
look at some of the other data and say, ‘Okay, if we hit this hard and heavy’. It is
going to depend on what you decide as a grade level. (3-3-10 p. 3)
One of the big differences between the existing and the proposed models was that teams would not necessarily need to plan the interventions together.

The proposed model was met with some skepticism by some staff members and alternate models were discussed. For example, Anne, the first grade representative, shared this concern,

It seems very inflexible to how much we can change or how much it can be modified. The time frame, if we go ahead with this, doesn’t take into account the different age levels. Low performing students only need about 20 minutes and higher kids need about an hour. (3-3-10 p. 5)

Many staff members were eager to stop doing grade-level wide interventions so that collaboration would not be required. This prompted one teacher to state,

They were talking about doing focus groups within their own classroom which to me is just teaching. [laughs] Differentiated instruction. (2-22-10mi p. 22)

While the discussion was inclusive and everyone was provided an opportunity to give input, the nature of the questions sometimes disappointed Sharon. At the end of the year, she lamented,

Their first questions weren’t ‘What will work best for the kids?’ but it was ‘How much planning time am I going to get with each model?’...I was really sad that that was the major criteria for a couple of [the teachers]. (6-15-10 p. 2)

Despite some reservations and incomplete buy-in, the Instructional Leadership Team decided to adopt the proposed model for SY2010-2011.

**Conclusion**

Looking back over the two-year span of this study, Sharon noted that many changes had taken place and that the downsizing of the school had provided some opportunities. She stated,
Honestly, it was easier when we had 1,250 kids than it was last year. I think the difference was the depth of what was going on…As I go deeper, it gets harder. It’s steering something that on the surface seems easy but as you go down it gets harder…We’ve learned a whole lot more about what’s going on here than we ever would have learned when we were 1,250. (6-15-10 pp. 1 & 4)

Part of the reason why the changes were so difficult was that they required a cultural shift.

Sharon described her perception of the beginnings of the shift,

When I was a teacher here, we did our PDSAs, sat down, and looked at the DRA levels. We wrote up a plan and then we’d meet about it…But this has probably been the first year where we really, really have gone more from looking at it at a surface level to really looking at driving instruction…One thing that I saw coming into this job was that we were all over the place. We were just flinging arrows everywhere. And when you tighten things up on people, it’s hard. Especially those people who want to be left the heck alone to be in their classroom doing what they want to do. It’s been a huge cultural shift. (6-15-10 pp. 4 & 15)

The shift, however, was not completed in the two-year timeframe of this study. While some people appeared to embrace a culture of evaluative inquiry, other staff members either grudgingly complied with mandated requirements or were either quietly or openly resistant. Sharon acknowledged that even after two years of the initiative, not all staff members shared a common vision. She stated,

I know already that we are going to talk about [vision] at the beginning of the year. A little bit about where we are going and why we are going where we are going. Not that we have to spend two days on it but for them to have some understanding and feedback into it. (6-15-10 p. 12)

Sharon also shared lessons that she learned through this change initiative. The first lesson was that data, and by extension instructional practices, were very personal. From the beginning, Sharon had noted some people were hesitant to share that some students were not growing at the desired pace for fear that they might be perceived as ineffective teachers. At the beginning of this study Sharon discussed the need to create a culture where teachers would be comfortable sharing their student achievement data,
Some teachers don’t like to do SATs [Student Assistance Team meetings where teachers ask for strategies to help struggling students] because they think it’s a reflection of their failure. That’s not true. Smart people know how to ask for help…So it’s working to kind of build a community where everybody can ask for what they need and get help with what they need. (9-30-08 p. 18)

Throughout the process, Sharon tried to convince staff members that decisions about teaching practices should be made based on student achievement data, not personal beliefs or philosophies. She stated in 2009,

I’ve always worried about [whether the students were receiving effective teaching or not], but the difference is now we’re really looking at the data. See, that’s the difference…Instead of going on hunches, you go on what you are really seeing [in the data]. The other thing that is good about the data is that it makes your case…This is what needs to change because this is what it looks like. It’s not against you. It’s not anything personal. (6-5-09 p. 23)

However, Sharon discovered through the teacher pushback to the initiative to develop a culture of evaluative inquiry just how deeply teachers’ identities were associated with their instructional practices and their students’ achievement. At the end of the study she stated,

It has been hard because for some people it is personal. It is a reflection of what they do and who they are. I think that the driving message is that it’s about looking at the needs of that child and what is going to work best for them. (6-15-10 p. 5)

The definition of “what is best for the child” differed, however, depending upon one’s personal belief system.

Another lesson that Sharon shared had to do with managing the change process. She declared, “The big lesson was not involving everybody.” (6-15-10 p. 2) She elaborated,

As a teacher, I never liked to go to meetings where you sat there and you were told stuff that you could read very well…I know what it is like at the beginning of the year when you are a teacher. You are sitting there and the principal is talking to you about something and you are thinking, ‘I have got 25 things to make here and what am I going to do here?’ Your mind is just all over the place. So I have tried really hard as a principal to value people’s time. On the other hand, I think I did that a little bit at the expense of some people. In trying to keep it as efficient and streamlined as I could, I think I unintentionally left some people out…I felt like if you have a
representative, then it’s okay. But that’s not the way a lot of people felt. (6-15-10 pp. 9-10)

By not including all staff members in discussions about why change was necessary, an opportunity was lost to develop a shared vision and to affect individual teachers’ schemas. Furthermore, some teachers believed they either did not have the opportunity to express their feelings or that their views were not valued.

Related to the lesson of involving more people in the process, Sharon also decided that some flexibility was needed with short term goals in order to maintain momentum towards long-term goals. She explained,

I made a major blunder first quarter. We were going to do this and we weren’t going to make any excuses whatsoever. The week [to print report cards using a new system] was coming up and we had several things coming up. The teachers asked if we could not do Intervention Wednesday and I said, “No, we are going to do it.” We won the battle, but I think we lost the war on that one because I think we sent a message that we didn’t value what they needed…I always promised myself that when I go into [administration], I wasn’t going to forget what it was like to be a teacher. And I did that week. Sometimes in your zeal to do the right thing, you do the wrong thing...We have to value what their needs are. Because if we don’t, we’re going to lose them. (6-15-10 pp. 7-8)

In response to this realization, Sharon planned to do several things differently the following year. First, she planned to identify times when the teachers would be allowed to focus solely on their classrooms and would not need to participate in collaborative activities. At the end of the case study she said,

One thing that I have done for this next year is I am making sure that as we look at the calendar…we identify ‘No fly zones’ when they want to be left alone. (6-15-10 p. 17)

Second, she planned to spend more time talking to teachers within a particular grade level alone. She elaborated,
One thing that I thought about doing…is maybe inviting [the grade level teachers] to lunch maybe once a month. Sitting down with them and having a chance to see how things are going and what their needs are. (6-15-10 p. 19)

She added, “I probably need to [also] spend more time with [the grade level chairs] individually.” (6-15-10 p. 19) By taking the time to meet with smaller groups of teachers, Sharon hoped to further develop individual relationships, which in turn could lead to greater trust and greater willingness to engage with change initiatives.

Finally, the greatest lesson learned by both the administration and the teachers was that the change process was challenging. While the concept of using data to inform instructional and administrative decisions on a regular basis sounded easy, the process of changing the organizational culture was actually quite complex and non-linear. Diane, the grade level chair for third grade who initiated grade level wide interventions first, warned other members of the Instructional Leadership Team,

It wasn’t like a nice neat little package the first time we tried it…You might tell everybody you just jump in and get your feet wet and then you start to sort it out from there. (9-16-09 p. 7)

It was not just the PLC teams who had to improvise and constantly adjust their course.

Sharon noted,

It would be nice to have a road map where you knew exactly what you were going to do and everything would go exactly the way that you wanted to do. But that’s not real life. I think I learned a lot in terms of what it really takes in terms of school reform and school growth. (6-15-10 p. 9)

While the school accomplished much within the two years towards developing a culture of evaluative inquiry, the staff members needed to continue to strategize and adjust to reach the long-term goal of improved student achievement.

272