This is to certify that the
dissertation entitled

At the cliff's edge: Utilizing evidence of student achievement for instructional improvement in a school district

presented by

Matthew Militello

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in K-12 Educational Administration

Major Professor's Signature

12/13/04

Date

MSU is an Affirmative Action/Equal Opportunity Institution

AT THE CLIFF'S EDGE: UTILIZING EVIDENCE OF STUDENT ACHIEVEMENT FOR INSTRUCTIONAL IMPROVEMENT IN A SCHOOL DISTRICT

By

Matthew Militello

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Educational Administration

2004

ABSTRACT

AT THE CLIFF'S EDGE: UTILIZING EVIDENCE OF STUDENT ACHIEVEMENT FOR INSTRUCTIONAL IMPROVEMENT IN A SCHOOL DISTRICT

By

Matthew Militello

The 2001 reauthorization of the Elementary and Secondary Education Act (ESEA), No Child Left Behind, brought the utilization of student assessment data to the fore. This mandate-based, sanction-ladened, and publicly exhorted policy ushered in a new era of accountability rooted in the collection, analysis, and use of student assessment data for educational improvements. This study explores how a school district responded to the new assessment accountability pressures.

This study seeks to understand both the reactive plans of the school district central office staff, vis-à-vis federal policy mandates, and the nascent actions of the individual actors (principals and teachers) that followed. The study's overarching question is: *How are school districts using evidence of student learning to guide decision-making and instructional improvement? If so, what sorts of data? And, if not, why?* As a result, this research investigates a district's strategies (long term vision), tactics (management, support, and resources), and operations (actual use) to infuse data into decision-making and pedagogical practices.

A review of the extant assessment accountability literature highlights problems in the readiness of district actors and tenuous findings on the impact of teaching practices and student achievement outcomes. While the study reveals a well-intentioned plan of action by district elites it uncovers contentious findings including: pedagogical issues of

teaching to the test and didactic teaching; problems with the metrics and utility of assessment data collection and analysis; and the most prominent use of assessment data for student placement and curricular adherence. In addition, the study exposes promising leads for assessment information guiding decision-making and instructional practice. Such promising assessments embody characteristics of proximity to administering assessments, correlation of assessments and the curriculum, on-going, specific professional development opportunities, and content specific, on-site professional resources.

This study also considers the predictive nature of alternative theoretical frameworks. Institutional isomorphism, organizational capacity and coherence, and the privacy and power of the teaching profession provide an alternative to the federal, state, and a district's theory of assessment accountability. When assessment information was filtered through this alternative framework its use became highly predictable and equally ineffective. Finally, this study offers implications for learning policy, knowledge-managed professional practice, and assessment accountability research anchored in the black box of teaching and learning.

For Elizabeth, Dominic, Luke, and Gabriel

ACKNOWLEDGEMENTS

In 1994 I was a member of the first Michigan State University masters cohort in West Michigan. There I was fortunate to cross paths with a number of influential faculty members. When the department instituted the first Ph.D. cohort in 2000 it was not a difficult decision to join the EAD family once again. While the masters cohort experience heightened my practical knowledge of educational administration and prepared me for my work as a school administrator, the Ph.D. cohort propelled me into a new academic career. As I reflect on my ten-year affiliation with M.S.U. I am reminded of the supportive and stimulating friends, family, colleagues, and mentors who have surrounded me.

To begin, members of my dissertation committee contributed to this work in important ways. I thank Michael Sedlak for helping me realize the importance of primary source research. This aided the development of a rich case study through the voice of the educators in the district. Michael was also able to assist me in procuring financial resources when I left my administrative post and came to MSU as a full time student. I thank David Plank for his analytical and honest assessment of my writing. David's ability to question my findings in a thoughtful manner provided an additional view of my work. I thank Maenette Benham who has been my mentor and friend throughout my graduate work. This dissertation would not have been completed without her. Working as a graduate assistant with Maenette has taught me about research and has ingrained in me an indelible work ethic. Finally, I want to thank Gary Sykes, my dissertation and guidance chair, for his intellectual ability to press my work into a new

and improved dimension. Gary's patience, thoughtful suggestions, and support contributed to my scholarship.

I am grateful to the administrators and teachers who took time from their practice to speak with me about their work. Without their stories I could not have written this dissertation.

I owe a great deal to my colleagues at MSU. The 2000 Cohort embarked on this journey together and always found time to visit the "library" after class to encourage and support one another. I want to also acknowledge my office mates, Joyce Piert and Patrick Halladay, who put up with my neurosis during the final phases of the dissertation. Finally, my fellow cohort member, office mate, and dear friend, Betty Murakami-Ramalho, thanks for walking, running, and stumbling through the Ph.D. journey with me.

I want to also recognize the colleagues in my professional life as a teacher and administrator. From my early teaching experience in Holland Public Schools I am indebted to Rich Ryzenga, Jon Manier, and Kathy Cook—friends and mentors. My administrative experiences were supported and shaped by a number of people. Specifically, from West Ottawa Schools: Dr. Charles Muncatchy, Jerry Klomparens, and the entire ninth grade team of teachers with whom I experienced the absolute highs and lows of life as a building leader. And, from my work in the Mason Schools I developed two of my best personal and professional friends: Chris Janson and Cheri Meier.

While all of the aforementioned individuals have supported my work, the opportunity to pursue an advance degree began at home. The unconditional love and support from my family anchored me emotionally throughout my work. I want to thank the two most important educators in my life, my mom and dad. My mom who instilled in

me waves of ambition and compassion, while my dad embedded an educational opera in my head where the school administrator could indeed be the protagonist. In addition, my extended family has been instrumental in my life: John and Michelle Militello, Martha and Dan Thorell, Susan and David Fultz, and John Goldner.

Finally, the dedication of this dissertation cannot summarize the love and admiration I have for Elizabeth and our boys. You are my sun, moon, and stars. Elizabeth, without you, I am not; with you, I am complete. Dominic, Luke, and Gabriel, I have learned so much from your youth and sense of wonderment. I am a better person because of you. I hope you remember the adventure of the past four and a half years with fond memories and the lesson that learning never ends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: FRAMING THE ISSUE

Economics [and by analogy psychology] is all about how people make choices; sociology [and by analogy anthropology and political science] is all about how they don't have any choices to make. ~ James Duesenberry

Introduction

Early hunting practices involved non-technical tools. Killing a woolly mammoth was a mental exercise of deception. That is, because early humans did not have the necessary tools to wound or kill the mammoth, humans had to back the huge animal to a cliff's edge. Soon, the animal learned to engage the humans and force a retreat. However, as new technologies evolved (rocks, arrows, fire, gunpowder, etc.) humans developed new ways to achieve their ends. Similarly, various technologies have been used as weaponry as a means to modify the behavior of school organizations.

The utilization of hard, scientific data has become an important theme in educational settings. Utilization of data in education implies use in two areas: assessing growth in student achievement and the implementation of scientifically based strategies for intervention. While assessments and accountability instruments have a long history in education, today's accountability is now synonymous with requirements to test students and the sanctions that accompany the results (Darling-Hammond, 2003). Consequently, schools have been cornered at the edge of a cliff. Resist jumping over the edge and school actors are castigated as unwilling or unable to change. Jump, and actors surrender their autonomy. The new federal demands to utilize student achievement data have brought a new set of advanced weapons: monetary sanctions, threats of school reorganization, and professional licensure certification to name a few. How this new mandate-based, sanction-ladened, and hortatory policy (McDonnell, 2004; McDonnell &

Elmore, 1987) targeted the utilization of evidence of student learning to guide decision-making and instructional practice in today's school districts was the focus of this study. More specifically, how a school district and its actors reacted to the latest assault guided this study. Educators, like the woolly mammoth have been placed in a precarious position. Understanding their actions will help determine the outcome of the initiative.

Problem Statement

Student assessments have become the measurement tools in the new accountability era. Historically, student achievement data have not played a significant role in the operations of schooling; however, recent trends have catapulted the collection and utilization of student achievement data into the daily operations of states, schools, and teachers. Policy makers today want schools to make decisions more rationally based on problems "uncovered by empirical data and ... programs proven effective by research to raise student achievement" (Massell, 2001, p. 148). As a result, understanding how raw data are contextualized into usable information and then transformed into new knowledge is now a critical component of the school organization's vision and strategic plan (Earl & Katz, 2002; O'Day, 2002; Petrides & Nodine, 2003). Under the current press for accountability, schools are faced with the task of becoming an organization where *data* are collected, examined, and transformed into meaningful *information* in order to make *knowledgeable* educative decisions (Petrides & Guiney, 2002).

Policymakers and educators alike agree that data are a powerful policy lever to change schools in productive ways (Earl & Katz, 2002; Learning First Alliance, 2003; The White House, 2001). While the business community has used the evolution of data, information, and knowledge to develop strategic plans, schools and data continue to have

a disjointed relationship (Petrides & Guiney, 2002). The complexity and uncertainty of the organization combined with the unclear and weak technology of schooling make discriminating simple statistics difficult in complex organizations like schools (Ogawa & Collom, 2000). Even though educators work toward fulfillment of federal and state mandates (Creighton, 2001), many often feel inadequate and threatened by the use and publication of achievement data (Earl & Katz, 2002). Making matters worse, educators are provided with little training in the use of data (Cizek, 2000; Creighton, 2001a; Holcomb, 2004; Popham, 2001) and feel disengaged from the planning process of collection and inspection of achievement data (Earl & Katz, 2002; O'Day, 2002). This created in a number of consequences, intended or not. Earl and Katz (2002) posited that the high demands to use data, coupled with the inadequate training and pervasive fear, result in the phenomenon of pedagogical practices geared toward tests and less on good instructional practice. Even more damning, Jones and Egley (2004) depicted the theory of action of NCLB as being undemocratic (centralized policy), oversimplified (schools are complex not simple organizations), and undermining (de-professionalizing the teaching profession).

The utilization of student assessment data was formalized by the 2001 reauthorization of the Elementary and Secondary Education Act (ESEA), No Child Left Behind. Consequently, the philosophical debate of *why* data should be utilized has morphed into a procedural and operational quandary of *how* to utilize evidence of student learning to guide instructional improvement. Important new questions have emerged. In what ways have data impacted decision-making in schools? Are educators participating in shaping data to change their practice or is their practice shaping the data? And, will

the coercive, exogenous pressures, and mandates have a simple axiomatic impact on administrative decision-making and pedagogical practices? The study examined the implied and explicit (if not sanctioned) demands placed on school districts and the individuals that comprise them to utilize data in their practice. Utilizing theories of the institution, organization, and information, this study began to explore the impact and utility of the implied theory of action and outcomes explicit in the No Child Left Behind (NCLB) legislation.

This study began by seeking to understand how a school district develops an internal action plan to collect, analyze, and use student achievement data. And, more importantly, how data are transformed into information and ultimately knowledge for practitioners in the district. As a result, this research investigated a district's strategies (long term vision), tactics (management, support, and resources), and operations (actual use) to infuse data into decision-making and pedagogical practices. While the study is a single site case study, multiple units in the case will be examined including: levels of the organization (central office administration, principals, and teachers) and grade levels (elementary and high school). In the end, this work embarks on understanding both the reactive plans of the school organization vis-à-vis federal policy mandates and the nascent actions of the individual actors that followed.

The insistent forces to utilize student assessment data are mounting from a number of fronts including the coercive (political forces to improve education writ large (The White House, 2001)), normative (to subscribe to the next wave of school reform boosterism (Fullan, 2001)), and the mimetic (economic credentialism and parental choice (Labaree, 1997; Meyer & Rowan, 1991; Plank & Sykes, 2003)). As a result,

understanding how a school district reacted to the sanction-based mandates to utilize achievement data is crucial. Specifically, the organizational artifacts that are created, technical support mechanisms provided, and the individual actors' responses at multiple levels of a district provided insights into the utilization of student achievement data to guide instructional practice.

Purpose Statement

The purpose of this single site, embedded case study was to discover both the organizational and individual impact of the institutional pressures to utilize student achievement data in educative decision-making and pedagogical practices. As a result of the recent sanction-ladened federal policy mandates as well as the exhortations of poor student achievement results in schools, school districts have been forced to create and mediate new policies around the utility of student achievement data. This study explored the collection, analysis, and use of multiple assessment data streams in one school district. More specifically, the study began to unpack the complexity of transforming student achievement data into knowledge to guide administrative decision-making and classroom pedagogical practice.

Significance

Today, we live in a chaotic, fast paced society. Moore's law of technology doubling every eighteen months has implications on society and social capital (c.f. Putnam (2000), *Bowling Alone*; Gleick (1999), *Faster*; Brown & Duguid (2000), *The Social Life of Information*). If institutions previously faced a famine of data, today's datum "has quickly turned into glut" (Brown & Duguid, 2000, p. 12). Data have infiltrated our society because of the advancements of the ubiquity, temporal proximity,

speed, access, and interactivity of technologies today. Recently, data have formally entered school organizations via strong policy levers. As data based policies have emerged in schools, understanding assessment data policy and practice must be a priority for educators.

Previous educational reforms have had a mostly superficial impact on teacher pedagogy and student achievement (Elmore, 2000; Ogawa & Collom, 2000). The growing body of literature on the effects of comprehensive school reforms (Borman, Hewes, Overman, & Brown, 2003), the current standards movement (Malen, 2003; Ogawa, Sandholtz, Martinez-Flores, & Scribner, 2003; Sipple, Killeen, & Monk, 2004; Spillane, 2004) and accountability efforts (Elmore, 2004; Fuhrman, 2004; O'Day, 2002; Ogawa & Collom, 2000) provide little evidence of the effects on student learning. While a preponderance of the research has atomistic units of analyses-- teachers (Elmore, Peterson, & McCarthy, 1996) and district leaders (Spillane, Halverson, & Diamond, 2001)-- little has been done to study the district as a whole (O'Day, 2002; O'Day & Smith, 1993; Spillane, 2004; Sykes, Thompson, & Skrla, under review). However, because the school organization (school districts) has a large influence on reform efforts (McLaughlin & Talbert, 2002), a district perspective was needed to better understand the use of data in schools today. Ultimately, this study will contribute knowledge to education on a number of fronts: policy architects, practitioners, and researchers.

Significance to Policy

This work had implications on the form and function of policy. Specifically, the study investigated the functionality of the current policy typology that is best defined as a

mandate-based, sanction-ladened, and hortatory policy (McDonnell, 2004; McDonnell & Elmore, 1987). March and Olsen (1988) posited:

> Policy analysts interested in designing organizations that can learn intelligently and organizational theorists interested in understanding the dynamics of organizational choice share the need for an effective model of organizational learning under conditions of uncertainty about what events happened, why they happened, and whether they were good or bad. Such situations are common in a wide variety of organizations; they are conspicuous in most public organizations. (p. 357)

Moreover, Elmore and McLaughlin (1988) stated that, "Policy can set the conditions for effective administration and practice, but it can't predetermine how those decisions will be made" (p. 10). As a result, taking an internal look at how local policy was being developed and enacted in the midst of a complex and uncertain environment is important work.

Significance to Practice

Understanding the organizational features that foster or inhibit the use of data is both significant and timely to practicing educators. The current dilemmas associated with utilizing student achievement data are creating pressures for practitioners in, out of, and through school districts. The practice of punitive measures (monetary and publicly reported scores) has raised the stakes of practice as well. As a result, educators must understand assessment data as a prerequisite to effective, meaningful utilization of such data. This work begins to unveil both the implied and tacit knowledge that educators must possess to utilize assessment information in an effective, meaningful manner. Understanding the work of educators is an important element to connect theory, policy, and practice. English (2002) stated, "The *theory-practice* gap has been to urge more emphasis on application in the form of fieldwork and the internship. This is

wrongheaded. The *theory-practice* gap will be removed when we construct different and better theories that predict effects of practice" (p. 3).

Rather than balkanizing the work of research from the work of practitioners, this study took advantage of research *in* practice "to understand which contextual factors [were] critical and which [were] not" (Burkhardt & Schoenfeld, 2004, p. 7). Uncovering pedagogical insights is an important function of this study. Because educators are conspicuously absent from both policy design and the analyses of the research community, this work provides substantive promises for leadership, technical mechanisms of support, pitfalls and pratfalls, and teaching and learning practices. Practitioners would benefit from a better understanding of the organizational and individual features of data collection, analysis, and use.

Significance to Research

Current studies fail to theorize and inform practice in regard to the vision, operation, and support of effective and meaningful use of data in schools. This work began to analyze the use of data in schools in the early stages of assessment accountability policy implementation through multiple theoretical lenses. The transformation of raw student achievement *data* to analyzed and disaggregated *information* and ultimately tacit *knowledge* to guide instructional improvement is a hefty task to be sure. The current rational frameworks of the bureaucratic reactions and outcomes of student achievement have not uncovered the detailed picture of school organizations that is warranted. Mistakenly, calls for additional organizational and agency explanations of schools have gone generally unanswered (Rowan, 1995). Consequently, this case-based research was an effort "to replace traditional theories of

technical contingency or strategic choice with alternative models that are more consistent with the organizational reality that researchers have observed" (DiMaggio & Powell, 1991a, p. 3). Furthermore, the research seeks to uncover, "some systemic variation in the information behavior of organizations and the individuals in them" (Feldman & March, 1988, p. 424).

Meyer and Rowan's (1991) institutional model of organizations as open and natural systems bounded by multiple rationalities and focused on legitimacy and efficiency will be tested in today's high stakes environment. Bidwell (2001) stated, "A prime objective of an organizational theory of schools should be to show how organizational structures and the formal specifications of the work of teachers and administrators affect school production" (p. 101). Without such micro level research, "we risk treating institutionalization as a black box at the organizational level, focusing on content to the exclusion of developing a systemic explanatory theory of process... and neglecting institutional variation and persistence" (Zucker, 1991, p. 105). While it is important to understanding the effects of the rational perspective on the outputs of education (e.g. achievement scores), it is vital to analyze the work inside the black box of school personnel through alternative frameworks (Black & Wiliam, 1998).

Research Questions

To better understand the role student achievement data plays in educational decision-making this study's overarching question was:

How are school districts using evidence of student learning to guide decision-making and instructional improvement? If so, what sorts of data? And, if not, why?

In addition, the use of assessment data in a district was explored by the study's research sub-questions:

a) *How did the district respond to the assessment accountability demands explicit in the No Child Left Behind legislation?*
b) *To what extent did the responses effect the organizational learning environment, administrative decision-making, teachers' pedagogical practice, and student outcomes?*
c) *What are the strategies, tactics, and operations that foster or inhibit the use of student achievement data?*

Conclusion

This study explored a district's response to the federal assessment accountability. This study also investigated the actions and perceptions of the actors in the district. As a result, the purpose of this chapter was to provide an overview of the current assessment accountability policy terrain. The next chapter summarizes the conceptual and empirical literature in the field of assessment accountability. In addition, Chapter Two provides a framework of competing hypotheses (normative and behavioral) to analyze and predict the utility of student assessment data in schools. Chapter Three provides the context and setting of the school district studied. Chapter Four describes how the district responded to the assessment accountability pressures. Chapter Five summarizes the utility of assessment data through the eyes of building principals and teachers at different grade levels (secondary and elementary). Chapter Six sets out to discuss the response of the district and its actors based on the conceptual framework and the extant empirical literature. In addition, the chapter attempts to provide insight to an emergent question: "Is the assessment accountability plan working in the school district?" The study concludes with Chapter Seven that provides a set of explanatory propositions and implications. Finally, a detailed account of the study's methodology can be found in the appendix (see Appendix A).

CHAPTER TWO: THE ASSESSMENT ACCOUNTABILITY PUZZLE

Darwin did not change the islands, but only people's opinions of them... That was how important mere opinions used to be in the era of big brains... Mere opinions in fact, were as likely to govern people's actions as hard evidence, and were subject to sudden reversal as hard evidence could ever be. ~ Kurt Vonnegut, *Galapagos*

Introduction

Historically, school accountability efforts have incorporated the use of curricular standards and student achievement data. At different times the accountability efforts focused on student testing (IQ tests in the 1920s, Minimal Competency Tests in the 1970s, and the testing provisions in the reauthorization of Elementary and Secondary Education Act in 1989, 1994, and 2001) and the creation of common standards (Committee of Ten, 1892, Cardinal Principles, 1918, and the standards movement of the1980s) as a means to evaluate the educational system. Current accountability reform efforts are attempting to combine both standards and assessments. That is, accountability today resides in a school's ability to utilize student achievement data from a common set of standards. Like previous reform efforts, there are disagreements on many fronts. There is a palpable disagreement about the *idea* of accountability-- what the mandates and sanctions are—and about the *ideals* of accountability-- what specific set of standards will be utilized. With little agreement on the standards, the next wave of accountability that is anchored in student assessment data marches forward.

The purpose of this chapter is to set a context for the inquiry into the current accountability reform efforts. The review begins with a survey of the existing empirical work in assessment accountability. Next, the study's conceptual framework outlines two analytical perspectives utilized to elucidate the case study findings.

Impact of Assessment Accountability

To stem the tide of apparent mediocrity in schools a number of measures have begun to emerge or reemerge on the educational landscape including: teacher compensation based on student achievement, new teacher preparation standards and testing, and most recently, a systemic effort to use student achievement data to grade the success of both students and schools. The rhetoric that has accompanied these reforms is bound in a call for both excellence and equity. While these measures resemble the hyperbole of the external efforts of the past, with the addition of monetary sanctions and school choice options to the reforms, schools have been forced to begin unprecedented efforts to reconceptualize educational decisions and practice.

As a result, there is a clear call to tap the technical core of schooling- the curriculum, teachers' pedagogy, and student learning. The impacts of these efforts are still unknown. This portion of the review will address the important questions of how this new wave of assessment accountability is affecting education as well as the questions that assessment accountability are raising. To this end, the review begins by exposing the three major questions raised by the assessment accountability debate- excellence, equity, and metrics. Subsequently, the review summarizes the extant empirical and conceptual research on the effects of assessment accountability on educational leaders, districts, teachers, and students.

Question of Excellence

Establishing a causal relationship between assessment-based reforms and student achievement is complicated and tentative at best. Nonetheless, a number of studies have reported gains in student achievement since the inception of assessment accountability.

In a study of the high stakes states of Texas and North Carolina, Carnoy and Loeb (2004) report that students are making improvements on NAEP scores. Carnoy, Loeb, and Smith (2003) also found significant improvements in four states where high school exit exams are required for graduation. Additionally, student achievement gains have also been attributed to assessment accountability at the elementary level (Carnoy, Loeb, & Smith, 2003; Supovitz & Taylor, 2003). Supovitz and Taylor (2003) concluded that assessments are improving elementary test scores in reading and math in a study of Florida counties. A recent Manhattan Institute report states that the stakes of tests are not distorting results. That is, the report validates the results of high-stakes tests by factoring out external features such as teaching to the test and cheating (Green, Winters, & Forster, 2003).

The student achievement gains attributed to the new pressures of assessment accountability have not gone unchallenged. Amrein and Berliner (2002) found that student achievement gains were inconclusive in an eighteen state study. The research found that independent test scores (e.g. ACT, SAT, NAEP, AP) did not improve along with state assessment results in each state. This finding is even more potent when considering that the population taking these independent assessments is not as representative as those who are administered state assessments. In addition, utilizing the same Texas TAAS[1] data set as Carnoy, Loeb, and Smith (2003), Haney (2000) found inconclusive results of achievement gains. Haney (2000) cites the gains are based on the exclusion of students taking tests. The mixed results in Texas led Haney (2000) to state, "The Texas miracle is more hat than cattle" (p. 1).

[1] The Texas state assessment is called the TAAS- Texas Assessment of Academic Skills.

Question of Equity

Fifty years after the Brown v. Board of Education decision, equity in schools continues to be an important issue. While the access to schooling has undoubtedly improved, questions remain in regard to the inequalities from school to school and the achievement gap between sub-groups of the population. Like the debate of the impact of assessment accountability on educational excellence, the effects of assessments on equity are at issue as well.

A number of researchers cite the positive effects of assessment accountability on the issue of equity. The researchers believe that this wave of accountability can and has had a positive impact on student achievement scores between groups. Utilizing disaggregated data within and between groups of students provides tools to address inequalities (Scheurich, Skrla, & Johnson, 2004). Huebert (2004) states, "Low-achieving students need high-quality instruction more than anyone else, and there is little question that data from large scale assessments, if used properly, can help improve instruction, hold schools accountable for improvement achievement, and identify and address students' learning needs" (p. 239). In a study of four school districts in the high stakes state of Texas, researchers found that gains were made to shrink the achievement gaps between certain sub-groups of students (Skrla, Scheurich, Johnson, & Koschoreck, 2004).

Proponents of assessment accountability for educational equity also point to the utilization of student achievement data as a means to move away from deficit thinking-- a deterministic view of achievement based on only contextual factors (Massell & Goertz, 2002; Skrla & Scheurich, 2004; Valencia, Valencia, Sloan, & Foley, 2004). Massell and Goertz (2002) found that data are able to debunk superstitious beliefs about achievement

(e.g. family circumstances). Moreover, in a study of four Texas school districts Skrla and Scheurich (2004) determined that the accountability efforts persuaded the district to look at assessment data, not just socio-economic and socio-cultural demographics, in regard to achievement. Similarly, Valencia, Valencia, Sloan, and Foley (2004) report that student assessment data help educators focus on the root causes of assessment inequalities rather than the symptoms. That is, the inputs (resources) and processes (quality of instruction) should drive instruction, not solely the outputs (student achievement assessments) or context for each class of student (Valencia et al., 2004). The research purports that assessment accountability showed evidence that the public was not serving all students well. As a result, data has put racial and socio economic achievement inequalities to the fore, thus increasing pressure for educational leaders to focus on the instruction of all kids (Skrla & Scheurich, 2004).

Conversely, a number of researchers have countered the equality gains that have been reported. Orfield (2004a) states that low achieving students categorized in sub-groups have to make more progress under NCLB because sub-groups are not mutually exclusive. That is, students can be counted in more than one sub-group (e.g. Hispanic, limited English Proficiency, and low income). Additionally, low achieving students must make larger gains to meet the common proficiency goals (Linn, 2003). Consequently, the groups of students that are the focus of the assessment accountability equality component are subject to the highest performance pressures (Madaus & Clarke, 2001) and are failing to meet AYP more than middle class students (Kim & Sunderman, 2004). Of greater concern is the possibility of data use as a subversive tool to eliminate groups of students

who are likely to perform poorly or to pervert the educational process by altering instruction (Ogawa & Collom, 2000).

Empirical studies support these conclusions. Tests of verbal and quantitative aptitude have already been associated with socio economic class (Haney, 2002). Utilizing state assessments, studies show that gaps between groups were not being closed in Texas (Klein, Hamilton, McGaffrey, & Stecher, 2000). Additionally, studies utilizing NAEP and TAAS assessment results come to similar conclusions (Klein, 2004). Amrein and Berliner (2002) found that 80% of students in affluent districts in Michigan qualified for the Merit Award[2] while only 6% of the students qualified from the Detroit schools. And, spending to support the Merit programs has outpaced general educational spending 2.5 to 1 (Amrein & Berliner, 2002). Armein and Berliner (2002) conclude that, "high school graduation exams affect students from racially minority backgrounds in greater proportions than they do white students" (p. 9).

In addition to the issue of achievement on standardized tests, researchers have uncovered disturbing effects on sub-group populations as a result of high-stakes pressures in a number of states. The Carnoy et al. (2003) study that reports overall improvements in student achievement scores also found that high school completion rates were on the decline in the high stakes states studied. Similarly, Haney (2000) concludes that there was a continuing decline of high school completion rates in Texas. Haney (2000) also purports an increase in 9th grade retention (only 50% of minority 9th graders have been progressing to graduation since TAAS was implemented in Texas) and lower college admission rates as detrimental effects of the accountability efforts in Texas. Finally,

[2] The Michigan Merit Award provides students with post-secondary scholarship monies if they attain proficiency status on the high school tests.

research has uncovered that the sub-group populations most in need of interventions are becoming victims of "teaching to the test" and "skill and drill" pedagogical practices as a result of the assessment accountability pressures in Texas (Klein et al., 2000; L. McNeil & Valenzuela, 2001). As a result, a number of theorists have tried to focus the achievement gap debate away from assessment measures alone and toward conceptual discussions of race, social justice, and social inequalities (Anderson, 2004; Parker, 2004).

While the debate is being waged on whether or not assessment accountability is improving the overall excellence and equality of student achievement, another battleground has formed over the metrics of the assessments themselves.

Question of Metrics

The fluidity and validity of assessments are limiting the efforts of those studying the effects on excellence and equity (Linn, 2003; Linn, Baker, & Betebenner, 2002; Rothman, 2004). Schools have problems identifying data, managing data, and choosing the metrics that best serve educational needs (Ogawa & Collom, 2000; Thorn, 2001). Schools possess both antiquated data warehousing means and archaic analysis measures (Petrides & Guiney, 2002; Thorn, 2001). In addition, comparative and longitudinal data are difficult to utilize as a result of the constant development and deployment of new tests. The data that are utilized by states to adhere to NCLB can be either norm-referenced (percentage based on national sample) or criterion referenced (performance based on standards). Because each state may devise their own plan of evaluation to meet the AYP mandate, comparing scores from state to state is not an option (Linn et al., 2002). Orfield (2004a) states that the lack of common metrics between states "punishes

schools in one state for achievement levels that are defined as great successes in another" (p. 5).

Because standardized state assessments are being utilized for student placement and to determine monetary sanctions and incentives, the test designs have come under great scrutiny. School assessments are not currently held to the same high standards as high stakes tests in other professions. The questions of validity, reliability, fairness and test scrutiny are all important factors that should be considered for school assessments if such high stakes are to be placed upon the results (Klein, 2004). A number of factors, other than cognitive ability, play important roles in assessment results including: cultural and socio economic test bias (Popham, 2001), dubious technical quality of tests (Haney, 2002), over-reliance on norm-referenced test items which are selected based on difficulty (Popham, 2001), lack of diagnostic items on assessments (Amrein & Berliner, 2002), changing pool of test candidates (Amrein & Berliner, 2002; Haney, 2000; Kornhaber, 2004), test-prep invalidation (Popham, 2001), timeliness of test results (Amrein & Berliner, 2002) to name a few confounding variables. The current demands to improve test scores have led to gaming techniques in schools. That is, test results can be easily skewed by student exclusion from tests (Amrein & Berliner, 2002; Kornhaber, 2004). Haney (2000) reports that improvements that were made in Texas scores could be attributed to such gaming techniques[3].

Performance indicators of everyday formative assessment efforts that are ubiquitous for classroom teachers are not considered in today's assessment accountability reform (e.g. student experiments and portfolios). Linn, Baker, and Betebenner (2002)

[3] Haney (2002) purports that tenth grade improvements on the TAAS were illusory as a result of the increase of students designated as "special education." The number of special education students doubled from 1994 to 1998 in Texas.

suggest that other evaluative possibilities should be considered including "the use of index scores, composites across grades, and rolling averages" (p. 16). Others advise that criterion-referenced tests should be used as well in order to compile data which focus on a more general body of knowledge (Haney, 2002; Popham, 2000). "Quality of life" data and context are often missing including: citizenship, art appreciation, and character education, to name a few. As a result, educators often question the current restructuring efforts that are based solely on standardized tests (Thorn, 2001). Teachers have echoed concerns over the psychometrics of assessments. Researchers learned in a survey of 708 elementary teachers that assessment metrics were not contextualized to the students (background and developmentally appropriate) or their curriculum (Jones & Egley, 2004).

Assessment metrics also must take into account reasonable standards. In many cases the level of attainment is set " too high to be held as a reasonable expectation for all students" (Linn et al., 2002, p.15). Rothman (2004) states, "Test items tend to measure the least cognitively complex of the expectations outlined in a standard" (p. 111). Too often the tests do not match the curricular standards that have been chosen (Herman, 2004). Students should not be held accountable for learning content they have not been taught (Elmore, 2004; Popham, 2001). And, if students are performing well in a setting where the subject-content and assessments are not aligned, test results can be attributed to teaching test skills as much as teaching content (Herman, 2004).

Schools should be held accountable to individual student learning over a period of time (Elmore, 2004). However, the current assessment designs are geared toward quick snap shots of cohorts (Asp, 2000). Evaluating good assessments entails not only

questions of psychometrics, but also the influence of assessments on teaching, learning, and the organization of schools (Chakwera, Khembo, & Sireci, 2004; Sykes, 1997). To be sure, the questions of excellence, metrics, and equality are left unresolved and highly contested. The promises of excellence and the utility of assessment accountability are retarded by the questions of metrics and equity. The next portion of this literature review continues the discussion of assessment accountability in terms of its impact on the district, leadership, teachers, and students.

Impact on Districts

Firestone (1989) described district responses to testing as: fragmented (no procedural "know-how"), communicative (relay only), or coordinated (aligned goals and activities). A number of studies have focused on particular district efforts to utilize student achievement data. McLaughlin and Talbert (2002) studied three California school districts and use of assessment data in schools that were internally held accountable and supported. In these high accountability schools the researchers found that "the schools collect and examine data to set goals and identify improvement plans throughout the system- from the district to the school, teacher, and individual students" (McLaughlin & Talbert, 2002, p. 186). In addition, these schools "regularly disaggregate data 'to set targets' particular to subgroups of students and schools" (McLaughlin & Talbert, 2002, p. 186). The timeliness and saliency of student achievement data were found to be crucial to inform resource allocation and decision-making (McLaughlin & Talbert, 2002). However, Massell (2001) purports that data appeases "political interests rather than... innovations for their proven ability to improve teaching and learning" (p. 148). Massell's study of standards based reforms in school districts in eight states over a

three-year period found mixed results in the use of data in districts. While meaningful decisions were made with student achievement data, too often decisions were made in a hasty fashion. Massell (2001) states that data are often treated primarily as "an accounting mechanism, not as a text to guide educational improvement" (p. 153).

Massell and Goertz (2002) report key elements and a growing body of strategies utilized in schools in their study of twenty-three school districts. To begin, aligning the curriculum and instruction to test outcomes was an important element. In many schools, data became a conversation point for educators to interact about their practice. Focusing on student achievement data help districts create programmatic changes in professional development opportunities for teachers (e.g. action teams) and support mechanisms for students (e.g. tutoring) (Massell & Goertz, 2002). However, other studies indicate that short-terms solutions were sought for assessment policy compliance rather than as a process-driven capacity building catalyst (O'Day, 2002).

The complexities of data generation and analysis generated a new market demand. And, educational software companies were quick to respond (Asp, 2000). Wayman and colleagues (2004) report on over thirteen different software companies that are contracted by schools. The companies each claim collection assistance and the ability to make multiple data streams operational. Easton and Luppescu (2004) report on the use of the Grow Program software that a number of schools have contracted. The program reports data to parents, teachers and administrators. In addition, individual reports for each group based on state standards and specific strengths and weaknesses are provided specifically for the Iowa Test of Basic Skills. Parents have access to a website and teachers get access for resources plus specific strategies based on the item analysis of the

test results. At a cost of $2 million a year, Grow is proving beneficial to teachers and administrators in the Chicago Public Schools[4] (Easton & Luppescu, 2004). The research reports that 54% of the teachers utilized the data reports to plan instruction while 50% used the additional resources. Principals overwhelmingly approved of the program. 80% of principals said they looked at the data and 72% stated they shared data with parents. Recently, Pennsylvania signed a $9.98 million deal with Grow to help districts in the state (Olson, 2004). Similarly, many districts have utilized the Schools Interoperability Framework (SIF). SIF is a non-profit membership organization comprised of over 100 software vendors, school districts, state departments of education and other organizations active in primary and secondary (K-12) markets, who have come together to create a set of rules and definitions to enable software programs from different companies to share information. Sadly such resources are not available to all school systems. In fact, research is indicating that the "accountability has not been accompanied by... [an] increase in resources" (Carnoy & Loeb, 2004, p. 192).

Schools failing to meet the assessment guidelines set by NCLB must provide supplemental educational services. In a six state study of available supplemental services, Sunderman and Kim (2004b) found that such services were not being offered in schools even though thousands of students were eligible. The report found that only one to eighteen percent of the eligible schools were actually offering services. The enormous financial burden to administer and fund the services along with the lack of working models or researched interventions are proving to be too much for schools (Sunderman & Kim, 2004b).

---

[4] Research is underway to see effects of the Grow program on student achievement.

Massell (2001) warns that data does not inject a "super rationality" into districts without a change in teachers' belief systems. Data management in schools means finding appropriate stakes of usable information for educators at the most atomistic level, student learning (Thorn, 2001). Additionally, the local knowledge and skill of teachers, the organizational culture, professional norms, and mobility of staff and students are all instrumental in the utilization of data (Massell, 2001). In order to shake up the complacent bureaucracy, data needs to be usable, incorporating a feedback loop between data and the practice of "real" schooling (Massell, 2001; Massell & Goertz, 2002). Massell (2001) states, "Districts and schools are applying data to make decisions in a mix of old and new ways. But it is clear that many respondents perceive a real change in the way they think about data to make decisions and often attribute it to the pressures of assessments and accountability" (p. 164-5). While the pressures of assessment accountability are emphasizing the use of student achievement data in schools, it is the district and school environment that are the key factors if assessment data are to impact teaching, learning, and leadership (Hightower, 2002).

Impact on Leadership

As a result of the external mandates to improve schooling, leaders have become mediators of school improvement efforts. Administrators have been left to lead the ambiguous means to achieve lofty goals set forth by the assessment accountability movement. There is little evidence that leadership training, compensation, and support have made notable gains in conjunction with the new demands. As a result, there is a great variation in the leadership response to accountability. Spillane (2000) reports that district leaders utilize behavioralist approaches to mediating state math policies. That is,

their leadership practices and decision-making reflect the compliance-based mandates of assessment accountability. Because the district administrators are overwhelmed with an inordinate number of tasks, a behaviorist approach may be a necessary survival function for change agents. Spillane (2002) states, "Relationships between the district change agents and classroom teachers may work against a situated or cognitive approach to teacher change" (p. 410). Administrators have tremendous burdens to symbolically comply (Ogawa et al., 2003) with mandates which may lead to highly centralized behaviors (Lemons, Luschei, & Siskin, 2003). This behavioralist approach is antithetical to understood needs to develop capacity (Elmore, 2003) and learning opportunities from policy (D. K. Cohen & Hill, 2000).

Assessment accountability has again placed the school administrator in a precarious position. Administrators are beginning to pay attention to data because of NCLB (Fusarelli, 2004) even though most administrators do not know good assessment from bad (Popham, 2001)[5]. In a survey of over 1900 school district administrators, money and mandates are listed as their primary concerns (Farkas, Johnson, & Duffet, 2003).

Capacity to implement reform begins with the building administrator (Farkas et al., 2003). Leithwood and Prestone (2002) state that, "both the pressures and supports needed to overcome the sheer inertia of the system as a new policy is introduced may be leveraged most effectively by those in administrative positions" (p. 51). In other words, the principal must mediate the school climate and the instructional organization (Bossert, Dwyer, Rowan, & Lee, 1982). Leithwood and Prestone (2002) describe four leadership

[5] Popham (2001) reports this is a result of administrative preparation programs not requiring enough courses on evaluation and design.

approaches for accountability: market or exit approach (salesperson), decentralized approach (empowerment and distribution of voice), management approach (rationale perspective plan with data), and the professional approach (site-based management). In order for educational leaders to build capacity for assessment accountability in schools, they must integrate new reforms into current improvement efforts and utilize a bundle of approaches (Adams & Kirst, 1999; Leithwood & Prestine, 2002).

Despite the pressures and lack of resources leaders face, there have been promising leads of effective leadership practices as a result of the assessment accountability reforms. District leaders have begun to leverage accountability pressures to improve schooling (Grogan & Roland, 2003; Skrla et al., 2004). Grogan and Roland (2003) conducted an in depth study of teachers and principals in Virginia. They report signs of distributed instructional leadership as a result of the new accountability demands in the state. Similarly, Hannaway and Kimball (2001) concluded from two national studies that district efforts were moving away from solely managerial roles and into the instructional domain. Another multi-state study reports that high capacity schools demonstrated "stretched" leadership between administration, department heads and teachers (Lemons et al., 2003).

Not surprisingly, principals have increased tension and stress and report a need for instructional skills required to lead with student achievement data (Grogan & Roland, 2003; Tucker, 2003). The pressures to manage the organization of schools forces school leaders to become reactive and hierarchical especially in times of chaos. As a result, leaders rely on behavioralist-management models at best and unethical practices at

worst[6]. Too often school reform standards and legislation outpace the implementation process causing the change efforts to become chaotic.

Leadership in a high-stakes era of assessment accountability hinges on a number of factors: organizational and accountability coherence and capacity building (Elmore, 2003; Lemons et al., 2003), data literacy and sensemaking (Earl & Katz, 2002; Lemons et al., 2003), data-based decision-making (Creighton, 2001b; Holcomb, 2004), and the development of a culture of inquiry (Earl & Katz, 2002). The development of coherence between pressures and enactment of policy are crucial for leaders (Lemons et al., 2003). Data literacy includes: knowing the purpose of data, recognizing sound and unsound data, being knowledgeable about statistical measurement concepts, making interpretation of data paramount, and paying attention to reports and audience (Earl & Katz, 2002). The development of a culture of inquiry is based on inclusive engaging in data with others, stimulating a sense of urgency, and making the time to converse about the data. Earl and Katz (2002) conclude, "Data provide a ready-made vehicle for engaging staff in planning. Leaders can broaden the base of inquiry by distributing leadership and developing a cadre of people who are competent and confident with using data" (p. 1020). In the end, leaders must know more about data and community building. With only the former, leaders are overly rational and objective, cold and mechanical. However, with only the latter, the leader is unconcerned with interpretation and inquiry of data. A balanced

[6] Recent media reports have emerged regarding ethical practices by district leaders in the quest to fulfill the new accountability requirements. In one case a Houston school principal and technology coordinator were suspended for falsifying data reports to the state (Dobbs, 2003; eSchool News staff, 2003). And, Michigan school districts have been accused of inflating school performance assessments, which make up one-third of their state grade, to boost their overall score (Fuller, 2004; MacDonald & Feighan, 2004).

approach of both data literacy and relationship building will be the hallmark of leadership in this era of assessment accountability.

Pressures to improve test scores in this age of assessment accountability are squarely placed on the shoulders of the school administrators. Educational leaders are asked to implement reforms that they must learn simultaneously. Spillane (2000) states that "carrots and sticks of various sorts can get district leaders to pay attention to state policies, but they cannot ensure that district leaders learn what state policy makers intended them to learn" (p. 158). As leaders become evaluated more and more on performance, they cannot rely on intuition or trial by error any longer. Data continues to be needed to understand and demonstrate what is wrong and what works. Leaders must be artistic with data and develop new capabilities (Earl & Katz, 2002). While there have been promising strides in the area of leadership for the new wave of accountability, critical issues of administrative will (motivation and approach) and capacity (knowledge) have yet to be investigated. The question of whether or not assessment accountability will press leaders into an instructional role remains to be seen. As Hallinger (1992) states, "The reality of the principal as instructional leader continues to lag well behind the rhetoric" (p. 45).

Impact on Teachers and Teaching

The current assessment focus on accountability has placed burdens on today's teaching profession both in what they teach (curriculum) and how they teach (pedagogy). Previous accountability reforms did not lead to drastic changes in how teachers taught. According to a study conducted by Wilson and Floden (2001), teaching did not change based on the standards based reform efforts of the 1990s . In fact, teachers continued to

teach in traditional ways with a bifocal vision, one eye on the policy and one on the students (Wilson & Floden, 2001). Similarly, Firestone, Mayrowetz, and Fairman (1998) determined that the effects of professional based state assessments in Maine and Maryland on teaching were minimal (some effects on content alignment, but not pedagogical practices). Even when teachers have been provided professional development opportunities to enact new pedagogical practices, the results have been wrought with misconception (D. K. Cohen, 1990). While there is more attention on the saliency of tests, the alignment of the curriculum, and some indication of utilizing assessments for instructional planning (Duke, Butin, & Sofka, 2003), there is little indication that pedagogical practices are changing as a result of assessment accountability.

There is some indication that the focus on student achievement data has affected teaching practice. Duke, Butin, and Sofka (2003) utilized data from Virginia high school English department chairs to better understand the effects of the state's new high stakes accountability system. The study reports an overall change in instructional planning and approaches to their practice of assessment and remediation. In addition, the study determined that tools or resources were being created in districts as a result of the accountability pressures: 53% of the schools had implemented tutoring programs, 31% enacted special testing programs for students, 72.3% of the teachers began sharing standards with kids, and 78% of the teachers began to be involved with analysis and interpretation of test results (Duke, Butin et al., 2003).

In a study of twenty-four schools in twelve states, the high-stakes accountability systems did help teachers focus on content to be taught; however, professional

development opportunities were misaligned and teachers and principals had different interpretations of assessment data (Berry et al., 2003). As a result, the ability to connect assessment data and student learning is difficult for teachers (Berry et al., 2003). Teachers lack the developmental opportunities to collect and analyze data (Cizek, 2000; Holcomb, 2004) not to mention use the data in their pedagogical decision-making (Asp, 2000). Supovitz and Klein (2003) study of five *America's Choice* schools supported the finding that teachers do not currently have the ability nor spend the time to utilize data. Consequently, teachers marginalize data by demanding access to more data because they don't trust the information or they are skeptical of one source of data (Mason, 2002; Massell, 2001).

Because teachers listen to policy signals, teaching content and pedagogy mimics the testing policy (Herman, 2004). As a result of the strong policy signals of assessment accountability, there is a fear that teachers are beginning to "teach to the test" (Amrein & Berliner, 2002; Carnoy et al., 2003; Earl & Katz, 2002; Earl & Torrance, 2000; Haney, 2000; Jones & Egley, 2004; Kornhaber, 2004; Massell, 2001; L. M. McNeil, 2000; Popham, 2001). Grogan and Roland (2003) purport that too often test preparation is becoming synonymous with teaching. In an *Education Week* teacher survey, 66% of the teachers who responded report a concentration on test information at the expense of other learning. In addition, 79% report spending time teaching to the test that included teaching students how to fill in bubble forms and answer multiple-choice questions (Doherty, 2001). Similarly, Jones and Egley (2004) found 23% of the 708 elementary teachers surveyed in Florida report teaching to the test.

The pressures of the external accountability efforts that rely on assessment data make it problematic to measure higher order thinking skills or performance assessments in a systemic fashion. As a result, the current accountability forces drive curriculum that focuses on knowledge and skill and not on the creation or use of knowledge (Elmore, 2004; Popham, 2001). Teachers resist standardized assessment data for fear that the tests rely on rote-memorization and low level thinking skills (Grogan & Roland, 2003). A recent international study reported that as the country of Malawi moved toward a system of assessment accountability, the previous project based methods of classroom assessment were discontinued (Chakwera et al., 2004). While deep and civic learning are more distal and difficult to enumerate (Amrein & Berliner, 2002), Smith, Lee, and Newmann (2002) found that interactive teaching and learning modalities proved more beneficial on Iowa tests than didactic teaching.

Giving primary focus to a standard set of basic skills inhibits not only deeper learning and higher order thinking skills, but also to the curricular reforms of integration and multiculturalism (Grogan & Roland, 2003). Amrein and Berliner (2002) specify the dangers of narrowing the curriculum and eliciting a training model of education, especially for disadvantaged students: "Poor, more than their advantaged peers, need not only the skills that training provides but need the more important benefits of learning and education that allow for full economic and social integration in our society" (p. 12). It appears that the pedagogical means to prepare students for tests will subjugate instruction to more traditional means. Popkewitz (2004) posits:

> Pedagogy needs intellectual tools that consider the relation between knowledge (concepts, generalizations) and the cultural practices that enable the production of that knowledge. Such a reading of science or mathematics for pedagogy, among other school subjects, would require a way of thinking that does not crystallize the

> conclusions and propositions of a field or produce a psychological reductionism. (p. 27)

The assessment systems in place have not focused on all subject areas. Siskin (2003) purports consequences on the non-tested subject areas. Her study of non-tested subjects in Kentucky created issues of legitimacy and marginalization. Teachers reported strong departmentalization as a result of the curricular isolation (Siskin, 2003). Moreover, non-tested subjects lost legitimacy by becoming perceived dumping grounds. Elementary teachers report teaching non-tested subjects as a "break" (Berry et al., 2003, p. 16). The end result is the steady decline and marginalization of the non-tested (e.g. vocational and humanities) courses in schools today (Fusarelli, 2004). Reports also indicate that issues teachers want to teach in grade levels unaffected by assessments have surfaced (Pedulla, 2003).

Assessment accountability may also have an effect on the perception, recruitment, and stability of the urban school teaching corps. Darling-Hammond (2004) rejects the notion of a teaching shortage in the United States. Rather, she indicates a shortage of good teachers in poor school settings. Poor urban and rural schools have been subjected to a succession of reforms imposed upon them. Another reform would overburden and demoralize staff and continue the exodus of good teachers from schools in need (Orfield, 2004b). Darling-Hammond and Sykes (2003) posit that this is indeed the case as a result of new demands on teachers in the current wave of accountability. Popham (2002) warns that teaching in a school with high-test scores does not mean one is a good teacher and vice versa.

The high stakes nature of assessments has also led to reports of districts and teacher cheating, teacher anxiety, and a loss of professional identity. Popham (2001)

describes teachers providing students with more time, hints, and even reviewing the examination with students prior to or immediately after the test. Teachers are weary of reforms that de-skill or "teacher proof" the curricula. Teachers are also feeling the loss of professional autonomy as a result of conflicting reforms, pressures to cover large amounts of information, and a mobile leadership (Hargreaves, Giles, & Foote, 2004). For many, their psychic rewards of teaching (Lortie, 1975) are becoming clouded with anxiety and fears of performance, compliance, and competition with colleagues (Berry et al., 2003; Duke, Butin et al., 2003). Recently, 22.5% of the 708 elementary teachers surveyed in Florida reported feeling stress via new accountability measures (Jones & Egley, 2004). Not surprisingly, teachers are becoming more and more nostalgic of the neighborhood schools and professional autonomy of the past (Hargreaves et al., 2004).

Orfield (2004a) states, "If one were to add the cost of days lost to instruction for additional testing and test preparation unrelated to general learning, the true economic and educational costs would be enormous (p. 5)." Most assuredly, high-stakes assessments have weighed heavily on the teaching corps personally and professionally. And, if teachers are feeling the anxiety of this wave of assessment accountability, surely students must not be immune to the high anxiety of high-stakes testing.

Impact on Students

Tying together assessments and accountability has led to mounting pressures on students. The stakes associated with assessments vary from state to state- from states not attaching any repercussions, to states (e.g. Michigan) that dole out college scholarships based on proficiency levels on tests, to highest stakes states (e.g. Florida, New York, North Carolina, and Texas) that utilize assessment results for student placement,

promotion, and graduation. Sisken (2004) reports that high stakes states that require such testing are placing enormous pressures on high school students. While students may successfully matriculate through the required local curriculum, students in high stakes states must ultimately pass a one-time assessment to graduate. Charbran (2003) reports that students feel both pressures and inconsistent messages due to high stakes testing. Students must decipher the disparities between course content and standardized assessment (Chabran, 2003). While teachers want students to take tests more seriously, students want alignment between school, state and college expectations (Chabran, 2003). Today, students are asking which subjects "count" as they matriculate through school. Students understand the necessity to obtain a high school degree in today's society. The call for high stakes testing becomes a barrier for students seeking to complete high school.

Indications of student anxiety and stress are not only reserved to high school students. Recently, New York City announced that approximately 10,000 students were to be held back in 3rd grade as a result of failing the state's exam (Gootman, 2004). In a recent study of 3rd grade students in Texas, Merchant (2004) reports students exhibiting signs of high anxiety, deception, and depression. Students in the study asked older siblings about how they faired on the 3rd grade test, complained of stomach aches for days leading up to the test, and lied about test results to fellow students (Merchant, 2004). Similarly, Jones and Egley (2004) found that 25% of the 708 Florida elementary teachers surveyed reported student anxiety over assessments. The report also cites that 12% of the teachers indicated that assessments negatively impact the enjoyment of school for children (Jones & Egley, 2004). The only results that are conclusive about student

retention is student anxiety not student achievement (Roderick & Engel, 2001; Voke, 2002).

Aside from the pressures of placement, promotion, and graduation, diverse learners feel repercussions from assessment accountability. If assessment accountability is leading to one style of teaching pedagogy, as reported previously, those who learn in diverse ways may be left behind (Fusarelli, 2004). In addition, the metrics of the test places the burden on students to recall knowledge rather than demonstrate their ability to utilize knowledge (Kornhaber, 2004). Finally, reports that the bar of proficiency is too high in many states (Kornhaber, 2004; Linn et al., 2002) and the reliance on norm-referenced tests that rely on validation by asking questions students may not have been taught (Popham, 2001) can lead to additional stress on the children. Undoubtedly, these consequences of assessment accountability were unintended. However, high stakes states have yet to modify the policies in place that lead to these mounting pressures on students.

****

Performance-based accountability is unquestionably a work in progress (Elmore, 2004). The review of the extant literature highlighted a set of contentious assessment outcomes and dynamics. Specifically, the literature uncovered the impediments of assessments to guide teaching and learning practices including: increased drop out rate, teachers' and schools' cheating on exams, teachers' defection from the profession, increased mobility of administrators, behavioralist leadership approaches, narrowing of the curriculum, higher retention rates, teaching to the test rather than the students, questionable test metrics, a negative impact on students, and a didactic modality of teaching. Nonetheless, assessment accountability efforts march forward. Will

assessment accountability ultimately be what Shakespeare writes in *MacBeth*, "full of sound and fury, signifying nothing?" Or, will the most recent wave of assessment accountability have a true impact on teaching practices and learning outcomes, where reform most certainly needs to root itself? In an attempt to unravel the puzzle, a multi-focal framework is utilized to analyze the form and function of assessment accountability in schools today.

Analytical Framework

This section outlines the conceptual framework utilized in the analysis of the study's findings. To begin, the normative, prescriptive perspective, implicit in the No Child Left Behind legislation, is outlined. This includes a number of emerging normative models that are touted as guides for districts to effectively utilize evidence of student learning. Subsequently, an alternative perspective drawing on theories of the institution, organization, and information provided an additional set of lenses to view the reaction of assessment accountability efforts in schools.

The Normative Perspective

Early organizational work sought the "one best way" to improve productivity (March & Simon, 1958, p. 19). The rational view characterized humans as inert objects and focused on the influence of such variables as cost, capacity, speed, and durability (March & Simon, 1958). This scientific approach delineated organizational decision making as a process of problem identification, search for alternatives, evaluation of alternatives, and ultimate choice of perspective solutions (March, 1999b; Simon, 1957). A rational system is considered a self-correcting, interdependent organization that has the ability to build consensus between goals and means, and then coordinate the

dissemination of information and predict problems (Weick, 1979). Theories of action are often anchored in the ideals of rationality. The rational theory is a behavioralist model that is void of important considerations: cultural, social, psychological and political. Consequently, the theory of action logic presupposes that "participants are in a position to shift from naively performing actions to reflectively engaging in argumentation" (Habermas, 1984, p. 195).

The theory of action explicit in the No Child Left Behind legislation advances a behavioralist approach to changing existing structures and activities. At best this approach expects a game rationality where everyone knows the rules and has the same motivation (March, 1978). In reality, a theory of action is a utopia that serves "to reconstruct an undamaged inter-subjectivity that allows both for unconstrained mutual understanding among individuals and for the identities of individuals who come to an unconstrained understanding with themselves" (Habermas, 1984, p. 2).

The No Child Left Behind Theory of Action

No Child Left Behind expands the role of federal government, alters state and federal relationships, and seeks to reform the entire system. In simple terms, NCLB seeks to frequently test students[7], chart the schools' growth, publicly display the results, sanction schools' that fail, and defines proficiency in math and reading. The charge of the legislation is "to ensure that all children have a fair, equal, and significant opportunity to obtain a high-quality education and reach, at a minimum, proficiency on challenging state academic achievement standards and state academic assessments" (The White House, 2001). No Child Left Behind represents a complex apparatus of standards and

[7] NCLB expanded testing accountability to all students not just those receiving Title I funding in previous versions of ESEA.

sanctions that ratchets up the accountability of collection and use of student assessment data in schools. The exhortations and, at times, hyperbole, associated with the policy have led to claims of a new policy type, the hortatory (McDonnell, 2004). Hortatory policy combines the regulatory mandates and inducement typologies with the exhortations of the need to change our educational system.

The legislation demands that states meet all aspects of NCLB by 2014 and that they utilize "reliable research[8]" to chart their progress (The White House, 2001). Subsequently, each state was given a mandate to establish a plan of action and assessment measures for NCLB. Specifically, each state must assess Adequate Yearly Progress (AYP) for student learning, enforce staffing qualifications, and collect disciplinary statistics to name a few. The data must also be categorized by NCLB "sub-groups" and disseminated for consumption by the state department of education and the public at large.

States are required to develop annual measurable objectives and administer annual tests in math and reading at least once during grades 3 through 5, 6 through 9, and 10 through 12. By 2005-06, the annual reading and math test will be required every year in grades 3 through 8 and once in grades 10 through 12. In addition, a science component will be required in 2007-08. These assessments are then used to chart student growth. The growth is measured by the percentage of students proficient in each subject area. While the legislation allows for individual states to develop AYP goals and "proficiency" standards, all students in all states must be proficient by 2014. Finally, schools are required to test at minimum, 95% of the student body.

[8] Reliable research is defined by the following themes: "scientific method, replicated, generalized, meets rigorous standards, and convergent findings" (The White House, 2001). This is mentioned 110 times in the NCLB legislation.

States are ultimately required to publicize Adequate Yearly Progress by individual schools. In addition, districts must report the data by sub-groups including: economically disadvantaged, racial and ethnic groups, limited English proficiency and students with disabilities. Other data that must become public includes graduation rates, attendance, teacher qualifications, and student discipline in some states. Failure to meet the goals sets into motion a series of sanctions and supportive levers. These sanctions and levers are designed to achieve the goals set by NCLB. What sets NCLB apart is its implied theory to boost student achievement scores via sanctions and threats (Neill, Guisban, & Schaeffer, 2004). These stages of sanctions are outlined in table 1.

Table 1: Failure to Meet AYP

Failure to Meet AYP	*Designation*	*Support Levers*	*Sanctions*
2 Consecutive Years	In Need of Improvement	The school must develop a two-year school improvement plan, and will receive technical assistance from the state	The school must allow for in district school choice and pay for any transportation costs with set aside Federal funds. Priority is given to lowest achieving, lowest-income students
3 Consecutive Years	Continuing School Improvement	The school must supply supplemental services[9] to disadvantaged students outside regular school day paid with set aside federal funds.	Same as above. In addition, potential to lose block Title I funding.
4 Consecutive Years	Corrective Action-Restructuring	The school must supply supplemental services to disadvantaged students outside regular school day.	Implement corrective action: ➢ Appoint outside expert ➢ Extend school day or year ➢ Replace staff who are relevant to the failure to make AYP ➢ Decrease management authority at school level

[9] Supplemental services are defined as "additional academic instruction designed to increase the academic achievement of students in low performing schools." When a school is designated "continuing school improvement" or worse, 25% of Title I monies must be utilized for supplemental services. These services must be preformed outside of the regular school day (Sunderman & Kim, 2004b).

Failure to Meet AYP	*Designation*	*Support Levers*	*Sanctions*
5 Consecutive Years	Alternative Governance Plan- Reconstitute	All of the above.	Reconstitute School: ➢ Remove teachers and/or principal ➢ Re-open as charter school ➢ Hire company to run school ➢ Turnover school to state ➢ Other approved reforms

The designations of AYP status are a result of state assessment results. The state assessments must be approved by the federal Department of Education. States have begun to develop or have in place state-mandated assessment systems as a result of NCLB. However, only 11 states had approved plan by the Department of Education as of June 2003 (Sunderman & Kim, 2004a). Former, Secretary of Education Rod Paige asserted that high stakes testing[10] will force us to look at gaps in achievement and give consequences and motivate failing schools (Paige, 2001). Asp (2000) reported that the assessment accountability system forged out of NCLB could be characterized by four purposes:

1. Accountability for the educational system writ large.
2. Feedback to inform instruction and impact student achievement.
3. Classification and certification of students based on knowledge and skills.
4. Fodder for further reform efforts (Asp, 2000).

Amrein and Berliner (2002) also outlined the thrust of the high stakes testing theory:

- Students and teachers need to know what is important to know and what is important to teach.
- Teachers need motivation to teach better.
- Students will be motivated to perform better (especially poor performers).
- Tests provide a good measure of achievement regardless of the students' background.

[10] High stakes testing has become defined by the sanctions and incentives that have become associated with results (e.g. determining funding for schools and graduation for students).

- Teachers and students will use the test results in decision-making (e.g. pedagogy and design of professional development activities).
- Parents understand and know how to interpret test results.

In addition, NCLB purports the following: the definition of a set of key student achievement values, a premise that student performance can be accurately and authentically measured, a belief that consequences motivate student and teacher performance, and a faith that there will not be any unintended consequences (Fuhrman, 2004). The legislation asserts that, "the federal government will invest in educational practices that work- that research evidence has shown to be effective in improving student performance" (U.S. Department of Education, 2003). The U.S. Department of Education's document titled, "No Child Left Behind: Using Data to Influence Classroom Decision" states, "Research shows that teachers who use student test performance to guide and improve teaching are more effective than teachers who do not use such information" (U.S. Department of Education, 2003).

The over-reliance on one summative data typology (student learning) has led to states offering alternative, formative assessments. Massachusetts' Coalition for Authentic Reform in Education (CARE) proposed a plan to replace the Massachusetts Comprehensive Assessment System (MCAS) with a local assessment authority plan. This plan limited the testing reports for a school portfolio by sub-group and has annual reviews from teams of experts (multiple day visits): "Each school will report on progress or lack thereof toward its goals and the states' core standards or competencies" (Neill, Guisbond, & Schaeffer, 2004, p. 24). Similarly, Nebraska's STARS (Student-based Teacher-led Assessment and Reporting System) has created local portfolios submitted to the state based on six criteria for designing and reporting purposes. The hope, in both

accountability systems, was to utilize the amalgamation of data provided in portfolios to inform educators rather than to make decisions about students (Neill et al., 2004). While recent efforts have been made to legitimize alternative measures[11], single measure outcomes (student learning as indicated by NCLB approved assessments) are still the anchor of accountability today (Black & Wiliam, 1998).

The framers and supporters of NCLB believe the road to improved educational outcomes is paved with accountability (assessments and sanctions) and competition. Holding to the moniker of "One year growth for one year of instruction," NCLB is an a priori attempt to change the current system of education. That is, the legislation stipulates that the belief and support of the legislation will follow the action of utilizing student achievement data. As a result, models have begun to emerge for educational decision-making with the hope of realizing the theory of action implied by No Child Left Behind.

Emergence of Assessment Models

While educators are in the knowledge transmission business, they continue to be classified as a data impoverished, intuition-based profession. A U.S. Department of Education (2002) report states that education continues to operate, "largely on the basis of ideology and professional consensus. As such, it is subject to fads and is incapable of cumulative progress that follows from the application of the scientific method and from the systemic collection and use of objective information in policy making" (p. 48). Even the educational associations have joined the cause to utilize data in schools as a way to

[11] The state of New York, considered a high stakes state, recently began discussions of developing elementary committees to review student promotion of those who fail to pass state tests at the elementary level because of the high failure rates, especially as reported in New York City. In addition, the state of Massachusetts has allowed special education students to submit portfolios in lieu of taking the MCAS. However, not one student was considered proficient under the portfolio assessment in 2003-04.

take the "rancor out of decision-making process" (AASA, 2002). Additionally, data can be useful in targeting effective improvement efforts rather than " adaptation by trial and error or for experimentation with fads that inevitably lose their appeal" (Earl & Katz, 2002, p. 1009).

A number of normative models have emerged to address the use of student achievement data in schools. From workbooks (Bernhardt, 1998; Holcomb, 2004), to refresher statistical analysis models (Creighton, 2001b), to accountability models of utilizing data for school improvement (O'Day, 2002) to knowledge management strategies (Choo, 2001; Keeney, 1998; Petrides & Guiney, 2002; Petrides & Nodine, 2003; Thorn, 2001), to models of assessment enactment (Keeney, 1998; NCREL, 2004; Popham, 2001), to software programs (Wayman et al., 2004) models of use have become more relevant and prevalent. To help school organizations categorize the use of data, the Western States Benchmarking Consortium developed four stages of development for how well educators use data to personalize education for students and profile and monitor student performance (see Table 2).

Table 2: Categories of School Data Use

Level	Description
Emergent	The *emergent level* is defined as a pejorative "one size fits all" instructional approach. Organizationally, there is no systemic data mining taking place.
Island	The *island level* provides some signs of data examination by the individual and the organization. However, current data practices are still prevalent.
Integrated	At the *integrated level*, there is evidence of programmatic use of data. Specifically, teachers utilize student performance data for instructional preparation and the district provides profile of student performance for each school. At this level the school also begins to communicate student performance data with parents.

Level	Description
Exemplar	The most advanced stage is the *exemplary level.* At this level teachers routinely use performance data to truly personalize learning for all students. And, the district aggregates and disaggregates data to classroom/school/district levels to determine necessary improvements in instructional practice. At this level data are not contextualized as any one type; rather, multiple sources of data are utilized. Individual performance profiles are created for each student whose parents will receive informative communication in the form of progress reports based on the individual profiles.

(Western States Benchmarking Consortium, 1997)

How the data are characterized is also a function of the types of data utilized. Bernhardt (1998) describes four types of data in her typology (see Table 3) while Popham (2001) cites the need for affective student data including: attitude toward learning, interest in subjects, and confidence. In both cases the utility of the data lies in the interoperability of different data typologies; too often single sets of data alone are utilized.

Table 3: Data Typology

Type of Data	Descriptor
Demographic	Data that is disaggregated by race, socio-economic class or other classifications.
Perceptional	Data generated on the feelings, usually surveys, of students, teachers, and parents.
Student Learning	Data that are both criterion and norm reference assessments and other formative or summative assessments at the district to classroom levels.
School Process	Data includes information about the school programs, offerings, teacher evaluation instrument, etc.

(Bernhardt, 1998)

Beyond the scales and typologies of use, specific guides of student data have surfaced. The Annenberg Institution development of the "inquiry cycle" can be established as one such normative model. The inquiry cycle model deemed

"accountability-minded schools" as those that engaged in a model that includes: establishing desired outcomes, defining questions, collecting and organizing data, making meaning of the data, taking action, assessing and evaluating actions, and ultimately a return to the desired outcomes (Keeney, 1998). Furthermore, the model addresses the optimal conditions and resources of using data to improve one's will, leadership, and skills (Keeney, 1998). The North Central Regional Education Laboratory creates a recipe in which data is the main ingredient. The NCREL recipe has four steps:

1. Prepare for Data Exploration- Form a data team to promote a data culture.
2. Collect and Organize Data- The team identifies relevant data, gathers it, and organizes it for analysis.
3. Analyze the Data- The team reviews the data and looks for patterns. The team identifies special programs or student groups most in need of improvement.
4. Explain the Results- The team uses the data to generate hypotheses and explain the underlying causes of areas needing improvement. (NCREL, 2004)

Popham (2001) also introduced a model entitled "Assessment's Ideal Role within the Instructional Process" (p. 29). The model rooted assessment in the daily operations of teaching and learning where assessments: represent the curriculum; are utilized for instructional decision-making; modify curricular content; and making inferences about students (Popham, 2001).

Organizational models of data management have also emerged to shape how school organizations can use multiple sources of data in an interoperable dynamic. Knowledge management was born out of a need to manifest the growing pressures to use information to maximize productivity in industry. However, unlike previous models focusing on output, knowledge management seeks to transform data into individual knowledge. The knowledge management framework posits that: 1) knowledge repositories must be developed (data warehousing to contextualized information); 2)

access to information must improve (accessibility and usability of information); 3) knowledge of the environment must be enhanced (change organizational norms and values); and 4) knowledge must be managed as an asset (results) (Davenport, DeLong, & Beers, 1998; Davenport & Prusak, 1998). The framework asserts this process will transform the tacit knowledge of the members of the environment (Nonaka & Takeuchi, 1995). With the recent rise of data accountability in schools, the principles of knowledge management have been offered as a working model to implement data in schools (Petrides & Guiney, 2002; Petrides & Nodine, 2003; Thorn, 2001).

Petrides and Guiney (2002) utilize an ecological model of knowledge management in schools that does not seek to institutionalize data, but rather make data more attainable, usable and meaningful for practitioners. Petrides and Nodine (2003) state, "Knowledge management brings together three core organizational resources-people, processes and technologies" (p.10). Schools should collect data on the context of their mission; this would help them move away from reliance on one test to determine their merit and engage educators in a process of reflection and inquiry (Petrides & Guiney, 2002). Petrides and Guiney (2002) indicate that the knowledge management framework is anchored in four principles:

1. Evaluating the current available information,
2. Determining information needed to support decision-making,
3. Operating within the context and perspective of the school's organizational processes, and
4. Assessing the school's information culture and politics (p. 1711).

The authors contend that leaders can cultivate a knowledge learning environment in schools that will have the potential to "enhance the overall academic and fiscal performance of teachers, administrators, staff, and students by using knowledge-based

information systems as catalysts to redirect and balance organizational culture and performance that can support organizational learning, transform the school into an evolutionary and innovative learning environment, and meet global demands and issues" (Petrides & Guiney, 2002, p. 1715). In the end, "the culture of the school must communicate the value and importance of information" (Petrides & Guiney, 2002, p. 1714).

Similarly, Thorn (2001) states, "The progression from data to knowledge can be seen both as a temporal process in which data, imported into a system's architecture, aggregates individual facts into summaries and averages that are then presented in an appropriate context" (p. 4). Data has great promise in schools if it is understood beyond objective reasoning and is in turn analytical and part of an emotional process that contextualizes the available data for all people involved (Earl & Katz, 2002). In other words, data are benign without context. Contextualized, data becomes information that can be applied to one's existing knowledge. Data may very well become information via summaries and averages, but it will not become knowledge without a steady diet of experience, intuition, values, and judgments (Thorn, 2001). Exploring meaningful patterns, causes, and relationships of data that are contextualized within the community and school is a necessary step if information is to become knowledge in educational settings (Petrides & Nodine, 2003). In a study of knowledge management in the Milwaukee Public Schools, Mason (2002) cites six challenges of transforming data to knowledge:

1. Desire
2. Focus on information use in the decision-making process
3. Acquisition and creation of data
4. Organization of data management

5. Analytical capacity
6. Strategically aligned information and results

Choo (2001) provides an interesting example of a successful knowledge management framework in action. Utilizing data from the World Health Organization's smallpox program, Choo analyzed the smallpox eradication program of the late 1960s. He concluded that the program melded together the elements of sensemaking, knowledge creation, and decision making "into continuous cycles of interpretation, innovation, and adaptive action" (p. 202). Choo (2001) stated, "Organizational knowing emerges when the three modes of information use are connected to each other to constitute a larger network of process" (Choo, 2001, p. 197). Rather than focus on program input data (i.e. the number of vaccinations), "information gathering was comprehensive, involving participants at all levels of the program, including local villagers and community leaders" (Choo, 2001, p. 204). In order to ensure accurate and timely data, incentives were provided and information was used with sensitivity and expediency. Choo posits that the eradication of smallpox by 1977 in Somalia was more a triumph of effective information management than of the technology of medicine itself. In the end, the process-oriented and culturally centered approach provides promising leads to the utilization of student learning data in schools.

O'Day (2002) stated, "accountability systems will foster improvements to the extent that they generate and focus attention on information relevant to teaching and learning, motivate individuals and school to use that information and expend effort to improve practice, build the knowledge base necessary for interpreting and applying the new information in improve practice, and allocate resources for all of the above" (p. 294). However, the high accountability business model does not fit well in schools where

"teachers cannot predict the outcomes of their endeavors because students, not they, are primary architects of those outcomes" (O'Day, 1996, p. 3). Consequently, the high stakes and clear goals are no guarantee of meaningful use and effective outcomes (DeBray, Parson, & Avila, 2003). Moreover, the means to achieve the improved student achievement using assessment data has been problematic and tenuous (Linn, 2003). O'Day (2002) stipulates four significant factors that inhibit high-stakes accountability efforts:

1. Focus on outcomes and not on the process of instruction;
2. Accountability systems operate within a highly functional bureaucracy framework of interactive patterns and normative structures;
3. Reliance on negative incentives distracts from risk-taking and innovation and toward organizational survival; and
4. Resource allocation is often in the form of a transmission model of learning that is inadequate for individual knowledge building. (p. 314-15)

The normative perspective implies a theory of action that utilizing achievement data would result in (a) clear and salient goals, (b) modified behaviors of district actors, and (c) clear technologies of organizational support. However, the rationality of theories of action is often usurped at the school level. Consequently, a set of alternative theories can be utilized to analyze the utility of student assessment data.

An Alternative Perspective

Rational choice theory is grounded in decision-making based on maximum utility; yet, it does not explain failed implementation (Spillane, 2004). Mistakenly, failure is attributed to cognitive fallibility (Habermas, 1984) or lack of individual interest (Buchmann, 1986; Spillane, 2004). The axiomatic implementation of policies based of such theories has not proven wholly rational or scientific (M. D. Cohen, March, & Olsen, 1972). Consequently, other theories are offered in order to understand policy in practice.

Specifically, descriptive, behavioral analysis provides an alternative lens to view the utility of student assessment data in schools.

Institutional Theories

The neo-institutional theory takes a sociological view on how organizations actually function under complex and ambiguous circumstances. Weiss (1995) stated, "As important as values, interests, and knowledge are in the decision-making process, they take tangible shape within the context of the organizations in which individuals work and live" (p. 578-9). More specifically, schools are "molded by structures (e.g. of participation), the rules (e.g. of access to information), and the norms (e.g. of appropriate behavior) in the institution" (Weiss, 1995, p. 574). As a result of the recent trend to centralize educational reform efforts, analyzing social life patterns within the confines of deeply inscribed norms, rules, and routines has become increasingly important and appropriate.

Institutional isomorphism provides a framework to understand how schools react to fundamental shifts in policy. In a quest to gain legitimacy, and in an attempt to adhere to the mandates associated with the theory of action, "organizations tend to model themselves after similar organizations in their field they perceive to be more legitimate or successful" (DiMaggio & Powell, 1991b, p. 70). DiMaggio and Powell (1991b) purported that schools engage in coercive, mimetic, and normative processes that yield institutional isomorphism. Coercive isomorphism is based on politically sanctioned pressures. Mimetic isomorphism is construed on the need to bring harmony in light of the uncertainty that pervades an organizational setting. And, normative isomorphism is enacted via professionalization, especially in educational and credentialing institutions.

Organizations tend to respond to uncertainty and pressures by constructing structures and features that reflect the expectations of the institution. This mobilization of systemic resources to modify practices often creates additional ambiguity and uncertainty. As a result, the impact of institutional isomorphism can further bind rationality and fragment an organization.

Meyer and Rowan's (1991) isomorphic model can also be utilized to analyze the institutional impact on organizations. The model stipulated three consequences of institutional isomorphism on organizations: 1) The decoupling of structural subunits from each other and from activity; 2) The rituals of confidence and good faith; and 3) The avoidance of inspection and effective evaluation (Meyer & Rowan, 1991). This model posited that organizations are de-coupled as a result of creating goals that are ambiguous and vacuous, conducting professional activities out of the purvey of managers, and inherently avoiding both integration and evaluation (Meyer & Rowan, 1991). As a result, the structure of schooling is "*disconnected from technical (work) activity, and activity is disconnected from its effect*" (Meyer & Rowan, 1978, p. 79, italics in original). Similarly, Habermas (1994) believes that institutional pressure "exacts a determinant of legitimization that makes it possible to steer social action... [is reflected in the] uncoupling of system integration and social interaction" (p. 180-81).

Maintaining displays of confidence and satisfaction is perceived as legitimizing to the internal and external organization. Meyer and Rowan (1991) submit that this type of logic pervades an organization, creating stability and social efficiency while simultaneously halting efforts to reform. Subsequently, school administrators focus on the managerial aspects (hiring teachers, student scheduling, etc.) of a school without

attending to its instructional aspects (Bidwell, 1965; Meyer & Rowan, 1991; Sykes & Elmore, 1988; Weick, 1976). Consequently, the administrative apparatus focuses on the legitimacy of the organization while the teachers continue to become increasingly balkanized[12].

As a result, the external forces on school organizations create an " organization in a loosely coupled state" (Meyer & Rowan, 1991, p. 60). Simply put, school organizations have tightly linked institutional and managerial functions and loosely linked the functions of teaching and learning (Cuban, 1990; Hoy & Miskel, 1991; Rosenholtz, 1989). Historically, this loose state has provided long-term effectiveness by buffering outside influences (Thompson, 1967). However, daily practice suffers from the inefficiencies of developing goals and purposes specific to the needs of the organization. The de-coupling of structures and activities in public schools has made school organizations exemplars of highly institutionalized organizations (Meyer & Rowan, 1991).

Theories of Information

Information has powerful implications in decision-making (Allison, 1971) and in the development of cultural norms (Vaughan, 1996). Weiss (1995) stated, "Information helps people figure out where the problems are and which potential solutions hold promise for coping with them effectively" (p. 576). As a result, the process of collecting, analyzing (disaggregation), communicating (dissemination), and using data has garnered a great deal of attention. However, closing the gap of descriptive data (what is) and

[12] A number of metaphors have been used to articulate the decoupling of teachers in schools from the organization as well as from one another including: egg cartons (Lortie, 1975), fiefdoms (Tyack & Cuban, 1995), outposts (Hodas, 1996), islands (Dede, 1998), privacies (Little, 1990), microsystems (Rowan, 1990), and cliques (Brackeroff, 1991).

normative data (what should be) is proving problematic (Sproull & Zubrow, 1981). Efforts to collect, analyze, and utilize data are cognitive, social, and political processes (Weiss, 1995). Therefore, information in "organizations [is] not innocent" (March, 1988, p. 387).

Organizations are "consumers, managers, and purveyors of information" (Feldman & March, 1988, p. 409). However, the links between information and organizations are weak. Management has always been in charge of developing procedures and rules for gathering, disseminating and utilizing information (Feldman & March, 1988; Mintzberg, 1973). Feldman and March (1988) stated, "The value of information depends in a well-defined way on the information's relevance to the decision made, and on its precision, cost, and reliability" (p. 411). In other words, only information that can mitigate choices has staying power. As a result, investigating how the information is utilized for decision-making is an important organizational feature to understand. Feldman and March (1988) list six observations about information use by individuals and organizations:

1. Much of the information that is generated and communicated... has little relevance.
2. Much of the information that is used to justify a decision is collected and interpreted after the decision has been made.
3. Much of the information gathered in response to requests for information is not considered in the making of decisions for which it was requested.
4. Regardless of information available at the time a decision is first considered, more information is requested.
5. Complaints that an organization does not have enough information to make a decision occur while available information is ignored.
6. Most organizations and individuals often collect more information than they use or can reasonably expect to use in the making of decisions. At the same time, they appear to be constantly needing or requesting more information, or complaining about inadequacies in information. (p. 414)

Seemingly rational information streams are often ambiguous to the actors in an organization. As a result, information often takes form as signal, symbol, or gossip in organizations (Feldman & March, 1988; March & Sevon, 1988). Information often signals compliance to exogenous forces (Feldman & March, 1988). That is, information as signal is used strategically to minimize or expose failure[13]. Information as symbol represents competence and legitimacy for organizations. Symbolic information is on stand-by or in the ready to justify previous decisions. Information as symbol supports the belief that the more information collected, the greater the belief that having any information, good or bad, is of great value and importance. Finally, information as gossip or "idle talk" signifies the personal nature and use of information. Schools are organizations that rely on and practice "idle talk;" in contrast, there is very little gossip between an air traffic controllers and a pilots (March & Sevon, 1988). Teachers in the Weiss' (1995) study of shared decision-making models in schools "most often turned for information to the opinions of their fellow teachers" (p. 583). Weiss (1995) confirmed that teachers have a general suspicion of outside information. Consequently, the legitimacy of information is often incestuous in organizations (Weiss, 1995).

March and Sevon (1988) observed that, "Relevant information is information of contemporaneous events or past experience rather than forecasts on the future" (p. 433). When optimized, information is "conspicuously consumed" in organizations (Feldman & March, 1988, p. 420). Because organizations often ignore information that is contextualized in the events and experiences of its actors, information is ultimately a tool

[13] Feldman and March (1988) call this strategy the information economic perspective where the strength of the information signals is based on the economic utility.

of surveillance. Information as surveillance is nothing more than a "thermostatic linkage between observations and actions" (Feldman & March, 1988, p. 416).

Organizational Capacity and Coherence

The notion of a theory of action --in this case assessment policy-- implies a direct and appropriate response for local actors. The unit that is the target of reform is treated as a static, malleable system. However, decision-making is a process of sensemaking and interpretation of policy signals, prior knowledge, and social context (Spillane, Reiser, & Reimer, 2002; Weick, 1995). Local educators' interests and ideological beliefs are powerful forces in their practice (Weiss, 1995). Buchmann (1986) purported that individual habits leads to conservatism that is "void of reason, centering on one's habits, interests, and opinions" (p. 530). Often, street-level actors are seeking autonomy from the goal oriented bureaucratic dependency theories (Goldring, 1995). As a result, educators' interpretation often undermines the notion that a fixed theory can have the intended impact on practice. Highlighting the actions of policy with the reactions of local educators helps address the collision of an action theory with local enactment.

When institutional or environmental pressures are low, the organizations are placed in an exploration mode (Brown & Duguid, 2000; March, 1999a). March (1999) described exploration with the vernacular of: slack, discovery, novelty, innovation, variation, risk taking and experimentation (March, 1999a). Similarly, Brown and Duguid (2000) characterized organizational exploration as emphasis on an inductive process of enacting a constant flow of intuitive actions, improvisation, reflections, and re-actions. Such enacting environments "gather information by trying new behaviors and seeing what happens. They experiment, test, and stimulate, and they ignore precedent, rules,

and traditional expectations" (Daft & Weick, 1984, p. 288). Enactment delineates the importance of individual practice and solicits satisfaction through gratifying efforts, not from inducement alone (M. D. Cohen & March, 1974). This exploratory search process allows actors in the organization to develop meaning and a deeper commitment to the enactment process (Rowan, 1990).

On the other hand, when the environmental pressures are high (e.g. federal mandates), organizations often engage in a process of exploitation. March (1999a) describes this organizational exploitation with words such as: refinement, routinization, production, implementation, efficiency, and reliability. As a result, the process focus imposed on organizations that have been inducement driven has dehumanizing effects (Brown & Duguid, 2000; M. D. Cohen & March, 1974). For example, these exploitive processes are manifested in what Daft and Weick (1984) called the discovery mode where an organization uses carefully devised probes to get the results they originally intended. Brown and Duguid (2000) posited that this type of overpowering process leads to organizational tunnel vision. Many organizations have no clear goals or feedback mechanisms; yet, they overly plan and are goal driven (M. D. Cohen et al., 1972). Consequently, the tighter the mandates are to couple performance and goals, the quicker the organization is to respond in a command style (Rowan, 1990).

Cyert & March (1963) posited that when performance fails to meet goals, a search process for a solution intensifies. This is a sequential process whereby the first best option is chosen in a wholly satiety oriented fashion. March (1997) stipulated, "Search is stimulated by a failure to achieve a goal and continues until it reveals an alternative that is good enough to satisfy existing evoked goals" (p. 12). The chosen alternatives are

most often sought in the neighborhood of old ones. As a result, the systemic factors may inhibit risk-taking, inquiry, and conversation while the individual factors may inhibit meaningful goal attainment. In psychological terms, the search process evokes the process of the *assimilation* of new knowledge into one's current schema as opposed to adjusting one's current schema to *accommodate* the new knowledge (Piaget, 1975).

The failure to take into account the local context presents an additional tension--organizational coherence and capacity. Policy failures are often focused on failed implementation based on organizational incapacity and incoherence (Chubb & Moe, 1990; Elmore, 2003). As a result, building internal capacity is seen by many as the key to coherence, clarity, and enactment of any external demand (Elmore, 2003; O'Day, Goertz, & Floden, 1995). However, imposing policy prior to building such internal capacity is problematic. Plank (1987) provides an illustration based on the 1980s press for graduation policies:

> High school graduation requirements have been increased in almost every state; stiffer teacher certification requirements and pre-service competency tests for teachers have been widely adopted. The distinguishing characteristics of these reforms are that they require virtually no changes in the present structure or operation of schools. Instead they seek to change the behavior of students and prospective teachers, in the hope that this will be sufficient to improve the performance of the educational system. (p. 146)

Similarly, Elmore (2003) identifies the problem of the delivery of new reform policies on top of existing organizational features:

> It is absolutely essential to understand that when policies lay down stakes in incoherent organizations, the stakes themselves do not cause the organizations to become more coherent and effective. The stakes are mediated and refracted by the organizations on which they fall. Stakes, if they work at all, do so by mobilizing resources, capacities, knowledge, and competencies that, by definition, are not present in the organizations and individuals whom they are intended to affect. (p. 288)

As a result, policy is often subsumed by organizational dynamics that mediate the policy. The constancy of policy reforms indicates competing explanations: either policy architects believe coherence and capacity can be mandated or additional policies are constructed to mediate the flaws in previous policies.

Conclusion

Formal resistance efforts have emerged to counter the demands of assessment accountability. There are emerging reports of parent boycotts of tests, proposals of alternative performance indicators, and some states considering opting out of NCLB[14]. However, as of March 2004, all fifty states were reportedly on track to meet at least half of the 40 requirements of NCLB (ECS, 2004). Additionally, more and more states (28 by 2008) are utilizing high school assessments as a requirement for graduation (Gayler, Chudowsky, Kober, & Hamilton, 2003; Goertz, Duffy, & with LeFloch, 2001). Regardless of the readiness of schools, the accountability efforts of NCLB press forward, forcing educators to keep one eye on assessments at all times.

Assessment accountability has been widely promoted and criticized; however, it can be stated that research has been unable to conclusively link high stakes testing to the transference of knowledge. Assessment accountability pundits submit that the system will hold teachers accountable to teaching curricular standards, and the disaggregated assessments will highlight underachieving sub-groups of the population that need special attention. In addition, high-stakes assessments promise to motivate data-driven administrative decision-making, changes in teacher practice, and improved student

[14] The cost of compliance may be more that the benefits some states receive from Title I funds. However, only a few states have the luxury of not relying on Title I monies.

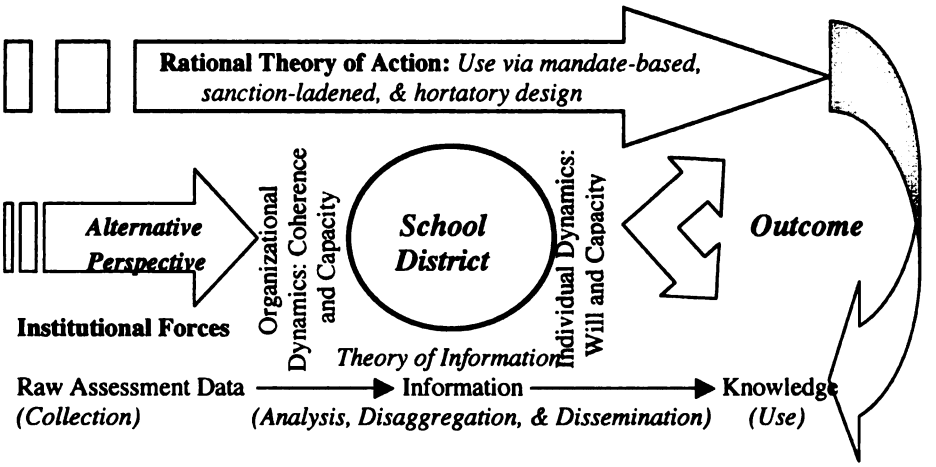
learning outcomes. Test based accountability has set goals to tie assessments to state standards, made measures to evaluate growth, created targets of progress, and introduced incentives by way of sanctions; however, schools have found it difficult to combine information with users (educators) and uses (pedagogy and student achievement).

Assessment policy in schools will either have a direct impact on decision-making—as expected by the theory of action-- or the alternative theory's predictions will be realized (see Figure 1). If the theory of action is realized, then it can be said that the establishment of a complex apparatus of rewards and sanctions and a set of normative models can affect the nature of schooling. One could then trace a sequential line from raw data to meaningful and effective use in practice. Specifically, it can be hypothesized that if data is normalized in a culture of knowledge management then student assessment data may positively impact decision-making and instructional practice, and, thus, student learning. In other words, the normative hypothesis states: *Student assessment data can guide decision-making and pedagogical practice with the proper mechanisms of pressure and support.*

If the actions of schools and the outcomes of achievement do not correlate to the theory of action, the alternative perspectives may prove worthy of prognostication. The alternative institutional theory hypothesizes that schools will begin a legitimizing process to enact a solution based on the actions of like organizations. Moreover, the solution will further de-couple sub-units of the organization, be void of inspection or evaluation, institute rituals of good faith, and impose command oriented leadership characteristics. The theories of information foresee that the school organization will symbolically collect more information than is needed, complain that the information available is not useful,

and gather information ex post facto of decision-making. Finally, the perspective of organizational theory posited that reforms begin and end with organizational coherence and clarity. This theory predicted that any policy, theory, or practice mounted onto a dysfunctional organization would fail. In sum, the descriptive, behavioral hypothesis stipulates: *The institutional, informational, and professional pressures mounted on an organization will mediate change practice, regardless of the stakes.*

Figure 1: Conceptual Framework



The purpose of this work is to understand the utility of student assessments to guide decision-making and instructional improvement in a school district. This case study aimed to analyze the use of student learning through normative and behavioral perspectives. Exposing each perspective will provide explanatory indicators of what fosters and inhibits to use of assessments in schools.

CHAPTER THREE: DISTRICT CONTEXT AND SETTING

Because no matter how hard the schools in Reo probably try, we are every year going to fail AYP. ~Mr. Caulfield, Elementary School Principal

Introduction

The purpose of this chapter is to characterize and contextualize the setting of the district that was studied[15]. To begin, the characteristics of the state and its accreditation system are summarized. Specifically, the state's response to the No Child Left Behind legislation is examined. Next, a brief description of the city in which the district resides is provided. This section is intended to contextualize the socio economic dynamics and educational needs of the city and surrounding area. Subsequently, the district itself is explored in regard to its demographics, organizational structure and performance on student achievement indicators. Finally, a description of the study's focus schools is offered.

The State of Michigan

This study is set in the state of Michigan. Michigan relies on heavy industry and tourism as major sources of economic income. There are a total of 554 school districts and 3,574 public schools, as well as 189 charter schools in Michigan. $17.2 billion are expended on the 1.6 million students in the state each year. Teachers in the state averaged the second highest salaries in the nation in 2003.

The state's Department of Education went through a number of changes in the past decade. The changes are closely associated with the political structure of the governor's office and the economic shortfalls faced by the state. In the late 1990s, the

[15] All of the data reported in this chapter was gathered from the 2000 United State Census Report, the district's 2003 annual report, 2002 Standards and Poor's dataset, and individual schools' annual reports.

governor began to decentralize state departments causing a significant reduction of staff in the state's Department of Education[16]. In addition, the school funding mechanism shifted from local property taxes to a sales tax system. In an attempt to bring about funding equity and stability for schools, the state created funding formulas for school districts. Districts are now provided a per-pupil funding grant. As a result, the pressure for districts to retain and expand their student population is now an important aspect of fiscal survival for districts in Michigan.

Currently, the new governor is re-establishing legitimacy in the Department of Education by expanding resources and programs to support the state's accreditation program. In particular, the state is ratcheting up efforts to train individuals and teams to assist schools in the state that fail to meet the state accreditation requirements. In addition, control of the state testing mechanism (MEAP) moved back to the Department of Education. While many of these moves are lauded, the lack of financial and personnel resources continue to strain the effectiveness of the Department of Education.

While No Child Left Behind marked a new era of federal demands on schools, the founding dynamics of state's rights continue to trump efforts to create one universal system of measurement accountability. As a result, each state was left to generate an accountability system within the federal guidelines. NCLB mandated states to develop a single-record database for 140 student variables. The Federal Financial Reporting System is not only mandated, it also determines federal funding. The reporting system includes multiple streams of assessment scores. As a result, states were required to develop a state assessment system and reporting guidelines to detail achievement gains.

[16] This included moving all of the state's assessment responsibilities to the Department of Treasury

To meet the federal demands of NCLB, the state designed an accreditation system dubbed "Education Yes!" The accreditation plan relied on the state's assessment program previously in place. The Michigan Education Assessment Program (MEAP) (developed in 1969) was designed to better understand what students know and are able to do in regard to the state standards of education. The MEAP is a criterion-referenced test that plays a major role in the state's accreditation plan.

Under Education Yes!, schools receive grades of "A", "B", "C", "D" (Alert), or Unaccredited. Each school building receives a letter grade based on a formula that includes three major components: achievement status on the MEAP, MEAP achievement change, and school performance indicators. Each component comprises 33.33% of the school's overall score. Most controversial to educators is the reliance on one test to determine the performance of the school. In fact, the accreditation system was delayed by more than a year by the state superintendent because of this concern. To appease the educational community, the performance indicators were added to the formula. However, this self-graded component has also taken a great deal of criticism. The performance indicators include: parent communication, building condition, curriculum alignment, student attendance, dropout rates, advanced placement offerings, and teacher preparation. Schools were charged to form committees in order to grade themselves in these areas.

As a result, the overall composite grade assessed for each school is under scrutiny from multiple angles. Charges of an over-reliance on one assessment methodology and concerns of grade inflation by schools on the performance indicators continue to dominate conversations about Education Yes! Additionally, the "stakes" associated with

the high school MEAP continue to be debated. Currently, students can receive a high school diploma without passing the 11th grade MEAP. However, scholarships and public perceptions are linked to these test results. The "stakes" associated with the test continue to change based on the economic dynamics of the state. For example, recent cuts in the state's budget caused a decrease in the amount granted for MEAP merit scholarships. The governor was able to retain the scholarships, but ultimately reduced the number by adding a community service component to the scholarships. The MEAP "Merit Award" provides students who meet or exceed the state's high school assessments with $2500 for in-state tuition and $1000 for out of state tuition. The budget for the Merit Award has grown fourfold over regular educational spending in Michigan.

In 2002 the U.S. Department of Education placed 1513 of the state schools on the list of failing schools as a result of not meeting Adequate Yearly Progress (AYP) goals. Originally AYP was measured based on the 1994 state school performance guideline in which 75% of the students needed to meet or exceed proficiency on the MEAP in four subject areas. The state took a publicity hit because of the number of schools that were considered failing. This was a result of the performance standards in the state as well as the number of other states that did not have a system of accreditation in place. By late 2002, new standards for the state were issued in accordance with NCLB. The state adopted a new accountability system that no longer took into account a science or social studies component on the MEAP and drastically decreased the benchmarks of proficiency in reading and math. Currently, the targets for meeting or exceeding state standards are set at 47% for elementary math, 38% for elementary reading, and 31% for middle school

reading and math. Not surprisingly, the number of schools labeled as failing diminished from 1513 in 2002 to 216 in 2003.

Reo City[17]

The metropolitan population of the city is close to a half a million residents. The city itself has a population of over 119,000. According to census data from 2000, the median age of its residents is 31.4. Ethnically, 65% of the city's residents are white; 22%, African American; 10%, Hispanic; and 3%, Asian. 58% of the residents own homes while 43% rent. The 2002 Standards and Poor's report noted that 15% of the children in the city reside in a single parent household (9.8%, state average). The household median income is just under $35, 000 and the unemployment rate is at 6.2% (state average, 6.7%). The average property value in the city in 2002 was near $14,000 (state average, $150,885). The 2001 FBI Uniform Crime Report indicated that the city reported over 900 violent crimes, over 4,000 property crimes, 11 murders, more than 90 rapes, over 1,500 robberies and aggravated assaults, and over 3,000 larcenies per 100,000 people.

There are three major categories of employers in the city: heavy industry, service industry, and a nearby major university. Of the three, the heavy industry, which is made up of a large manufacturing company, is the major employer. The service industry includes the schools, hospitals, and the headquarters of major retail companies.

The city is surrounded by a number of affluent communities. As a result of the intermediate school district's school choice plan and the 1994 state legislation supporting charter schools with public monies, competition for students has increased. Recently,

[17] All names and programs are pseudonyms. In addition, the city, district, and schools percentages and totals have been rounded.

there has been a proliferation of construction of new schools in the neighboring districts. Consequently, districts in the metropolitan area, including Reo, have engaged in an advertising campaign to lure students to their district. Like other urban centers, the city has problems retaining a mixed socioeconomic population. The exodus to the suburbs has caused both the city and the school to wage a campaign to retain students.

The Reo School District

The Reo School District has twenty-nine elementary schools, six middle schools, and three high schools. The district has three streams of thematic magnet schools. Each theme has an elementary, middle, and high school magnet. The three magnet themes are visual and performing arts; math, science, and technology; and language, culture, and communication arts. These magnet schools were developed through a three year 6.4 million dollar federal grant received in the fall of 2001. In addition, the district has two specialty schools including a regional school for severely handicapped students and a technology based school for high school and adult learners.

Demographics

The Reo School District currently enrolls just over 17,000 students. The district has begun to slow the exodus of students out of the district in the past decade (enrollment in 1987-88 was 22,000). The district employs seventy-eight administrators, more than 1,300 teachers and professionals and 1,100 support staff. General fund revenues totaled just over $184 million for the 2002-03 school year. The ethnic demographics of the student population includes: 41% African American, 37% Caucasian, 15% Hispanic, 5% Asian, 1% Native American, and .06% Native Hawaiian. 4.4% of the students in the district are considered to have limited English proficiency. Finally, the 19% of the

students qualifying for special education (up from 7.4% in 1994-95) is above the state average of 13%.

The percentages of students in the Reo School District who are considered economically disadvantaged is nearly double that of the state average (57% locally and 30.7% for the state). As a result, a high number of students in the district qualify for free or reduced lunch, 61% (up from 50% in 1994-95). The mobility rate is high in the district. Since 1995, 36% of families with children have lived in a different house within the city. The state funding allowance to the district per student in 2002 was $6900 (state average, $6701), and federal source revenue per student was $1,400 (state average, $520). The Reo school district had a $10.1 thousand per student expenditure compared to the state average of $7,733 in 2002-03. The per student operating expenses are broken down by instruction, $5.400 (state average, $4603); administration, $1.100 (state average, $956); operations and maintenance, $1100 (state average, $775); transportation, $400 (state average, $343); and food service, $380 (state average, $293). The average teacher in the district makes just over $56,000 (state average $49,379). Finally, the dropout and graduation rates continue to fall below the state averages. In 2003, the graduation rate in Reo was 71% (state average, 84.8%) and the drop out rate was 9% (state average, 4.1%).

While there have been improvements in the Reo School District, fifteen schools in the district did not make Adequate Yearly Progress (AYP) in 2003 (the schools included the three high schools, six middle schools, and six elementary schools). However, 2004 data placed only eleven of the schools on the failing AYP list (the schools included the three high schools, three middles schools, and five elementary schools). It

should also be noted that an elementary school that failed AYP in 2003 was restructured, and, thus began a new accreditation cycle in 2004.

Parent involvement, as measured by parent attendance at parent/teacher conferences, is high at the elementary level, averaging more that 90% in attendance. The middle school and high school's average is 40%. In regard to student discipline, the percentage of physical assaults by students was 43% (32%, state average); weapons possession, 14% (3%, state average); and drug interventions, 36% (8%, state average).

Student Performance Indicators

Student performance measures provide the context for the current achievement of students in the Reo School District. The chief assessments that are utilized for public consumption and federal and state mandates are provided.

PSAT/ACT/SAT

The Reo District students are at the national averages for both the ACT and the SAT tests. The district's ACT scores averaged 20.2 in 2003 with 48% of the students participating (national average, 20.8). Local students averaged 531 on the math SAT and 511 in the verbal in 2003 with 12% of the students participating (national average, 519 and 517 respectively). In addition, 10th graders have the opportunity to take the PSAT in order to qualify for merit scholarships. Students in Reo averaged a score of 141 on the PSAT with 27% students participating in 2002. The state average was 145 with 32.6% of the students participating that same year. The three district high schools receive feedback on the assessments from the testing services. Post secondary institutions utilize these tests for admissions.

Elementary MEAP Results

The state categorizes assessment results by four levels: exceeds standards, meets standards, basic proficiency, and apprentice. The overall pass rate for the MEAP in the district was 40% (state average, 52.8%). The MEAP assesses students in English and language arts (ELA) and math in grade four and science and social studies in grade five. The state provides schools with disaggregated scores by building, sub-group, and individual student. Currently, the disaggregated reports do not provide teachers with item-analysis reports. Table 4 provides sub-group percentages of students meeting or exceeding standards for the 4th and 5th grade MEAP along with the state average for all students in parentheses.

Table 4: 4th and 5th Grade MEAP Results 2003

	English Language Arts	Math	Science	Social Studies
All Students	50% (75%)	54% (65%)	69% (77%)	16% (28%)
Boys	43%	56%	73%	18%
Girls	58%	52%	64%	13%
Native American	57%	71%	87%	30%
Asian	39%	58%	67%	10%
African American	45%	43%	59%	9%
Hispanic	43%	55%	65%	12%
Caucasian	61%	65%	80%	25%

Middle School MEAP Results

The MEAP requires students to be assessed in the 7th and 8th grades. The 7th grade students are administered an English Language Arts assessment while 8th graders take assessments in math, science, and social studies. Table 5 provides a summary for 7th and 8th grade MEAP in 2003. The percentages of students meeting or exceeding standards for by sub-groups are also provided. State averages for all students in provided in parentheses.

Table 5: 7th Grade and 8th Grade MEAP Results 2003

	7 ELA	8 Math	8 Science	8 Social Studies
All Students	36% (60%)	24% (52%)	44% (65%)	16% (33%)
Boys	31%	27%	45%	16%
Girls	42%	23%	43%	15%
Native American	31%	25%	41%	12%
Asian	42%	28%	38%	9%
African American	31%	14%	31%	9%
Hispanic	36%	22%	40%	12%
Caucasian	44%	38%	62%	27%

High School HST Results

The MEAP assessment at the high school level is also called the High School Test (HST). In 2003, 101,282 students took the 11th grade HST. Taken in the spring of each year, the assessment is designed to test student knowledge through 10th grade. The test is aligned to the Michigan curricular framework. Students may re-take tests in the fall and spring of their senior year. The state has four levels of score categorization for the high school assessments: exceeds standards, meets standards, endorsed at basic level, and not endorsed. Table 6 provides HST results for the Reo high schools meeting or exceeding standards with state averages for all students in parentheses.

Table 6: High School HST 2003 Results

	Reading	Science	Math	Social Studies	Writing
All Students	57% (67%)	37% (61%)	47% (60%)	13% (26%)	43% (61%)
Boys	26%	18%	23%	7%	17%
Girls	32%	19%	24%	6%	27%
Native American	<1%	<1%	<1%	0%	.<1%
Asian	3%	3%	4%	14%	4%
African American	14%	7%	10%	1%	11%
Hispanic	6%	3%	4%	1%	4%
Caucasian	29%	21%	24%	9%	21%

Advanced Placement Tests

The district offers sixteen advanced placement courses in each of the three high schools. Of the 333 year-end exams taken in May 2002, 137 (41%) were granted college credit. 19% of students who qualified to take the assessment participated. The state average for earning college credit was 54% with 10% of eligible students participating.

The Focus Schools and the Study Respondents

This study focused on two individual schools in the district. River High School and Wood Street Elementary School were investigated to better understand the role student achievement assessments played in the decision-making process to guide instructional improvement. Each school is briefly described.

River High School

River High School was built in the 1940s. The school houses 1189 students in grades 9-12. The school's population is ethnically distributed as follows: 59% African American, 27% Caucasian, 8% Hispanic, 5% Asian, and 1% Native American. The school's 2001-02 dropout rate was 6.5% with a student retention rate of 93.5%. The district reports that 43% of the River H.S. students received free or reduced lunch while the Standards and Poor's report categorizes 34% of the school's students as being economically disadvantaged. River High School's overall MEAP pass rate was 39% in 2003. Table 7 provides a percentage of the River High School students (by sub-group) who are categorized as meeting or exceeding the MEAP assessment standards.

River High School has recently completed year two of a three year Comprehensive School Reform Grant (CSRG). The CSRG required a selection of a

model provider to administer professional development, and an external evaluative instrument. River High School selected High Schools That Work (HSTW) as a model provider, and contracted with the Southern Regional Educational Board (SREB) as consultants for HSTW. The grant mandates a norm-referenced evaluative assessment. The Iowa Test of Basic Skills fulfills this requirement.

Table 7: River High School HST 2003 Summary

	Math	Reading	Science	Writing	Social Studies
All Students	49%	62%	38%	43%	14%
Native American	<1%	<1%	<1%	1%	0%
Asian	2%	2%	1%	<1%	<1%
African American	19%	26%	11%	20%	3%
Hispanic	1%	3%	<1%	<1%	<1%
Caucasian	23%	27%	21%	19%	9%

Wood Street Elementary School

Wood Street Elementary is the district's oldest elementary. The building opened in the early 1900s as a ten-room school. The school's population includes 271 students and 16.5 certified staff members. Over 90% of the students receive free or reduced lunch. A 2002 Standards and Poor's report identified more than 85% of Wood Street's population as economically disadvantaged. The school has a student mobility rate of 40%. The overall pass rate for the MEAP was 40% at Wood Street. Table 8 provides a percentage of the Wood Street students (by sub-group) who are categorized as meeting or exceeding the MEAP assessment standards.

Wood Street recently completed a federal 21st Century Learning Grant. The grant provided for extended day instructional programming. Currently, the school is in year one of a three year Reading First Grant. The grant targets reading in grades K-3. The

DIBELS assessment is utilized to evaluate student and program performance. In addition, the district has contracted over fifty hours of professional development from a textbook publishing company. The grant provides the district with a full time literacy coach who works exclusively with early elementary students and teachers. Finally, the school is partnered with the local university in a program called "Young Greeks." The program includes job shadowing for 5th graders, pen pals, and one-on-one tutoring.

Table 8: Wood Street 2003 MEAP Summary

	4 ELA	4 Math	5 Science	5 Social Studies
All Students	46%	70%	45%	13%
Native American	**	**	*	*
Asian	**	*	**	**
African American	45%	55%	30%	5%
Hispanic	*	*	*	*
Caucasian	*	*	*	*

* 9 or fewer students tested
** No students tested

Study Respondents

This study was informed by a number of district participants. In sum, twenty-seven district employees were interviewed for this study. Additionally, a district contracted consultant participated in the study. The district participants represented central office administrators (n=6), River High School administrators and teachers (n=9), Wood Street Elementary principal and teachers (n=6), and additional teachers (n=2) and building administrators (n=4) from buildings other than the focus schools. A detailed descriptive narrative is provided for each participant in the appendix (see Appendix B).

Conclusion

The purpose of this chapter was to describe the contextual setting of the Reo School district, its employees, and its surroundings. Understanding how the district is

using evidence of student learning to guide instructional improvement is fostered or inhibited by the organizational and individual dynamics of the district. As a result, a descriptive account of the people and places of the district was warranted.

CHAPTER FOUR: THE REO SCHOOL DISTRICT RESPONDS

When content instruction and local and state assessments are aligned, they become powerful forces that contribute to the success of student achievement. ~ Michigan Department of Education

Introduction

The purpose of this chapter is to explore the Reo district's response to the current wave of assessment accountability. It is the aim of this chapter to make transparent the district leaderships' beliefs about and use of evidence of student learning. As a result, this chapter unveils a concrete account of the district's policies and mechanisms that seek to promote and stimulate the use of student learning evidence to guide decision-making and instructional improvement.

This chapter provides a brief summary of the district's use of assessments and mechanisms to improve teacher practice prior to the 2000-2001 school year. This year marked an important rupture in external and internal policy levers for the Reo district. Specifically, Dr. Pinkerton began her tenure as superintendent in 2000 that began a chain of internal reorganization and accountability efforts. Moreover, in 2001 the federal No Child Left Behind legislation was enacted that mandated a new set of sanctions and incentives that catapulted the use of student assessment data to the fore. Next, a detailed account of the district's beliefs and policies to utilize student achievement data is described. These beliefs led to the creation of an assessment accountability plan of action. A summary of the plan's purpose, design characteristics, scope and sequence, and mechanisms of supervision and professional support are subsequently outlined. The use of direct quotes from participants, the analysis of artifacts collected, and the researcher's

field notes and memos are embedded in this chapter to provide a rich accounting of the story.

Reo Before Dr. Pinkerton and NCLB

The Reo school district has historically faced a number of challenges endemic in many urban settings. Before 2000, the district faced a number of obstacles including poor performance on the MEAP tests, budgetary distress, and the mass exodus of students from the district. The public reports of the underachieving academic standards in the district coincided with the emergence of new schools (new physical structures) in adjacent districts and the intermediate school district's approval of school choice options. Like many urban centers, property values have declined in the city center along with rising rates of unemployment and those living in poor socio-economic conditions. Even so, educators and community members take great pride in a school district that has a rich history of community involvement and academic success.

The assessment landscape in the Reo district prior to 2000, like many schools in Michigan at the time, was rather sparse. The district participated in the required state MEAP examinations as well as a norm-referenced test (Metropolitan Achievement Test) that became board policy in 1975. The 1990s saw Reo test scores, graduation rates, and enrollment decline. In addition, there was an emerging perception that Reo was yet another urban school district in peril. In 1996, a district and community ad hoc committee was formed to make recommendations to the district. The committee's recommendations included aligning the district's curriculum with the state standards. The committee also had discussions about the utility of student assessments to track student growth. Consequently, the Reo's Board of Education search for a new

superintendent in 2000 was centered on these goals. What the Reo community knew was that the next superintendent would have a number of specific board goals to achieve. What was yet to be introduced was the federal reauthorization of ESEA. Together, these two forces would set the stage for a radical departure from previous policies to impact decision-making and pedagogical practice in the Reo schools.

The Arrival of Dr. Pinkerton and NCLB

Upon arriving in the district, Dr. Pinkerton was surprised that the district was just beginning the process of aligning the district's curriculum with the state benchmarks. To her surprise, she found that the elementary schools were "employing five different reading programs and that there was no consistent textbook adoption system in place... [Teachers] adopted what they wanted instead of what was best for the students" (Interview Transcript, Paragraph 30). In addition, the district lacked of an articulated, standards-based textbook adoption policy. As a result, there were curricular fractures from classroom to classroom and school to school.

Dr. Pinkerton's first order of business included the reorganization of the central office leadership. This new organizational structure was based on the model utilized in Pinkerton's previous district. Pinkerton created a new position, Area Director, in an attempt to bring central office administrators closer to the school buildings[18]. Three Area Directors were established and housed in schools throughout the district. Each director had oversight of one of the district's three high schools as well as all of the feeder schools. In addition, Pinkerton created a new position of Chief Academic Officer (CAO) who had a targeted mandate; improve MEAP scores. The CAO utilizes all assessments

[18] Pinkerton served as an Area Director in the district she was hired from.

that go through the district's office of student assessments. This data set includes the MEAP and the Iowa Test. As a result of the position's job description, the MEAP scores are clearly in the crosshairs at all times.

The Assistant Superintendent for Curriculum (ASC) responsibilities includes oversight and planning of all state and federal grants; professional development for teachers and administrators; gifted and talented programs; career pathways; English as a second language programs; Title I funding; coordination of school improvement plans; and the management of the district Quarterly Assessment and the district's curricular Pacing Guides. The ASC coordinates the administration, collection, analysis, and dissemination of the district's Quarterly Assessments. As a result of grant or professional development compliance, the ASC utilizes data that is disaggregated by outside agencies including the DIBELS, Gates-MacGinitie, Houghtin-Mifflin, and the PROM/SE. Both the CAO and ASC work very closely with student achievement data, but each utilizes a different database. To complicate matters further, the only district wide assessment that was not administered, collected, and analyzed through the district's Office of Research Evaluation and Program Assessment (REPA)—which reports to the curriculum office--were the Quarterly Assessments. Mrs. Redding, Director of REPA, believed the Quarterlies had not been tested as a valid instrument, and, thus, they had little use for them. Dr. Pinkerton stated:

> Mrs. Redding has no respect for the assessment process because they have never been validated... But nobody will bite the bullet and say that she needs to validate these if that is what the teachers feel. Or just take them for what they are, and let her not worry about it, and just use it as a mark up of, 'Is this a good question and should a 10th grader know it?' (Interview Transcript, Paragraph 37, 39)

As a result, the multiple streams of student assessment data collected by the district are not warehoused in one location.

In the fall of 2004 the district announced another central office re-organization. The district's Chief Academic Officer was moved into an Area Director position that included a decrease in salary (three positions still remain based on moving another director to a different position in central office). The vacated position was not filled and the Assistant Superintendent for Curriculum was given additional duties as well as a salary increase.

When Dr. Pinkerton began, she accelerated the previous efforts to align the core curriculum with the state benchmarks. The superintendent also established a number of new educational structures including: all day kindergarten and an elementary Montessori program to name a few[19]. Additionally, external funding sources were utilized to establish new programs in schools including: magnet schools, an elementary Reading First grant and a high school Comprehensive School Reform grant[20]. Finally, Pinkerton initiated an aggressive publicity campaign in the local media[21] to highlight the new learning opportunities Reo had to offer in an attempt to stem the exodus of students.

Pinkerton arrived at Reo in the midst of an educational organization in dire straits. The Reo Board of Education hired Dr. Pinkerton not to maintain the current operations, but to restructure the system in order to improve education for students in Reo. In fact,

---

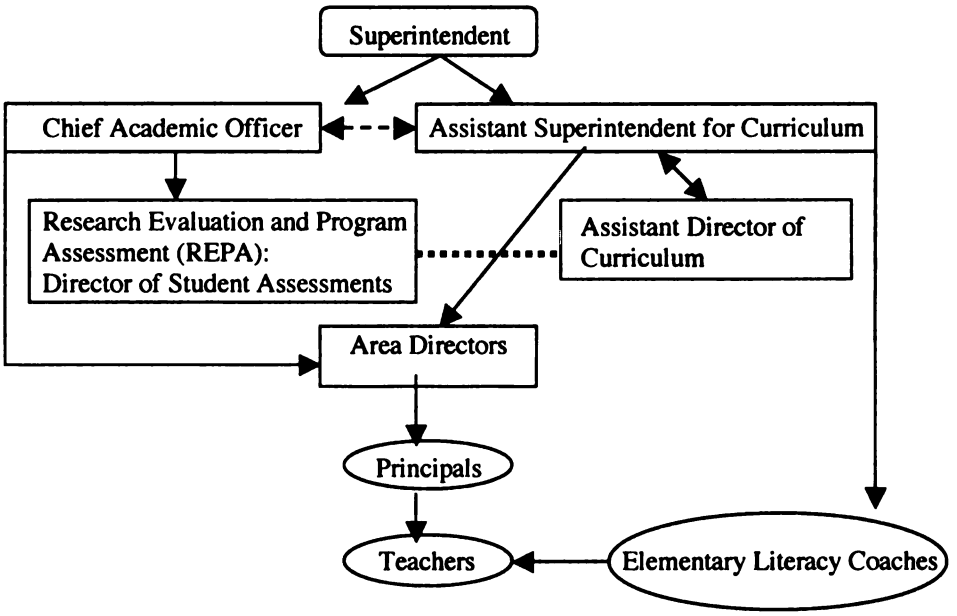
[19] The most recent effort in creating new educational structures involves the implementation of an International Baccalaureate program in one of the district's high schools.

[20] Pinkerton's pursuit of external funding continues. In 2004, Reo was awarded a federal Smaller Learning Communities grant and a science PI-CRUST (Professional Inquiry Science Communities for the Reform of Urban Science Teaching) sponsored by Athens University. To date, the superintendent has supplemented the Reo budget by $40 million.

[21] Pinkerton created a public relations position in the district and embarked on utilizing the media in the form of billboard, newspaper, and radio advertisements. Additionally, the district began to create professional annual reports that have won publication awards in both 2003 and 2004.

the Board attached performance pay in Dr. Pinkerton's contract to student retention rates and student test scores on the MEAP. Pinkerton accepted the challenge of turning the district around: "Challenges make me tick; there is nothing I am afraid of" (Interview Transcript, Paragraph 4). As a result, upon Dr. Pinkerton's arrival in the Reo School District a complete administrative reorganization took place (see Figure 2). Pinkerton's next phase of restructuring would be the introduction of a new accountability plan. The accountability plan for student improvement would be rooted in the use of student assessment data.

Figure 2: District Organizational Structure



Reo's Assessment Accountability Plan

The emergence of the assessment accountability system in the Reo School District can be described as a confluence of external and internal pressures. The No Child Left Behind legislation (2001) and the subsequent enactment of the state's accountability system Education Yes! (2002), characterize the external demands placed on the district. The external, legislative pressures previously detailed, provided a mandate to

reconceptualize how Reo utilized student-learning data. The internal assessment pressures coincided with the arrival of a new superintendent in Reo. Dr. Pinkerton's tenure, which began in 2000, marked a significant change in the way the district administration viewed and utilized evidence of student learning. This portion of the chapter summarizes the ideological beliefs of Dr. Pinkerton and her central office administrators in regard to assessment data. Subsequently, the assessment plan of action enacted in 2001 is outlined in regard to the purposes, sequence, and feedback. Finally, the mechanisms of communication, supervision, and professional support are examined.

Central Office Beliefs about Student Assessment Data

For a number of central office administrators who had been in the district for years, the arrival of Dr. Pinkerton and her vision of creating a strong assessment system were welcomed. Central office administrators believed they needed additional resources to hold teachers accountable to the curriculum and to improve instructional practices. That is, the assessments were the missing tool needed to "prove" whether or not teachers in the district were teaching *what* they were mandated to teach. As a result, the architects of the assessment strategic vision fully supported its implementation.

The district leadership consistently spoke of the conceptual utility of student assessment data. Dr. Pinkerton believed that assessment data should be an essential element in decision-making in schools:

> One of my frustrations has always been teachers always claiming, 'I don't want to pre-judge the students, I don't want to know anything about them.' But that's like a physician saying, 'you know you're new, you have been to doctors but I want to start over with you, I don't want to see any previous test results.'" (Interview Transcript, Paragraph 74)

Pinkerton believed that difficulties of teaching and learning in urban districts could be affected by first establishing common curricular standards:

> When you are in a mobile urban district, kids can't help that their parents move over the weekend. But you can make sure the curriculum is the same as they move from one school to another. It would be a travesty to have them enter a new school in the district and have to begin a different curriculum... that is what we had here when I arrived. (Interview Transcript, Paragraph 45)

Dr. Pinkerton also believed that common curricular standards needed measures of growth. According to Pinkerton, "assessments [would] heighten the interest in what kids were learning and how teachers were teaching... Assessments will always be seen as a negative creature; they aren't. They are help aids, they are tools" (Interview Transcript, Paragraph 53, 37).

Mrs. Redding, Director of Student Assessments, was elated to have a superintendent who supported the utilization of data. Mrs. Redding stated:

> The teachers will identify where their own weaknesses might be and address the needs of the students. So, if there are any triggers of pattern that seem to emerge in terms of addressing particular groups of students that they know: I have a problem with kids with limited English proficiency; I have a problem with kids from an economically disadvantaged background, or whatever, they can seek to find professional development or resources to help them address those specific needs. (Interview Transcripts, Paragraph 49)

Redding added:

> If you have a one year old and that child hasn't started to walk and you know that you have to do some kind of physical thing, because usually kids will walk by the time they are one year old. Then you need to retool, and you need to do something, because the expectation is that that will happen and unless there is something else going on with this kid, which is less than a certain percentage of the kids with a physical disability or something else going on, kids will be walking. You are not going to give up. You are going to do something different. You are not just going to let them not succeed. (Interview Transcripts, Paragraph 93)

Dr. Whitehurst, Assistant Superintendent for Curriculum, also embraced the new move to added accountability in the Reo schools:

> I think we have to be a little more accountable as educators when we know that children aren't learning and something is happening. It would happen anywhere else. A doctor would be told, 'Good-bye.' If they find something is happening with a lawyer, they can lose their license. But, you know what I mean, and we seem to be different. We have the lives of children in our hands, the most important products, but we aren't held to the same standards. And that doesn't make sense to me... We need to look at assessment analysis and find out what pieces students are not doing well in and address these pieces (Interview Transcript, Paragraph 107, 33)

The district's Chief Academic Officer, Mrs. Greenly, supported this assertion that data can be utilized for pedagogical decisions:

> The data specifically should tell the teacher how well the material was taught. I have always said that children don't fail a test, we should look at the test scores as an indicator of how well we presented the subject matter. Of course, we have to have something to record on the report card for how well the students are learning, but the teachers really need to use that data to drive the decisions they are making for why I am spending time on this particular area. If our children are doing well in fractions, why would I continue to do concentrated lessons over and above and sending fractions for homework if we know they understand fractions? Let's hone in on where the weak areas are, and let's try to change our instruction so we address those areas (Interview Transcript, Paragraph 22).

Administrators cited the utility of the assessment data for a comparative, motivational force in the district. The Director of Student Assessments added that data serves a comparative utility:

> It has been the district's philosophy in my time of being here that we would assess our local curriculum, our state requirement which is our MEAP, and that we would try to stay in step with national norms and expectations because we are a global society, and we want our students to be competitive on a national and international level. So we test them on other assessments so that we are keeping in pace with the national norm, so to speak. (Mrs. Redding, Interview Transcript, Paragraph 23)

An Area Director also commented on the utility of comparing schools and classrooms:

> I like to compare the teachers; really, that is what I am doing. And I will bring it up to a principal and say, 'I know we should have two third-grade teachers, one seems to be quite successful in multiplications here and the other is not. Have you done any teaming with them, and have you observed any difference? Her manners were not good or so on.' So those are the kind of things that I look at. I am really comparing teachers' ability to improve test scores. (Mrs. Grey, Interview Transcript, Paragraph 75)

The competitive nature is considered an incentive to evoke conversation. Dr. Whitehurst explained, "There is a purpose for this, and it is not to intimidate or find out who is the worst teacher. It should be to provide assistance and help for those teachers" (Interview Transcript, Paragraph 97).

The utility of assessment data was also intended to impact the budgetary process for building principals. Mrs. Greenly, described how decisions about the budget have become localized by means of the utilization of student assessment data:

> Dr. Pinkerton's vision on data, that is how we address doing our school improvement plans that we not simply determine a monetary figure. Give it to the schools and tell them the plan around that figure. She has now asked the schools to look at your present data and support the reasons for your requests through what you have learned through data. (Interview Transcript, Paragraph 28)

As a result, building principals had to focus attention on data to plan and justify district expenditures.

Finally, it was explicitly stated that the assessments were to stimulate conversations about achievement and pedagogy among teachers, administrators, and parents. The central office administrators believe that teacher conversations based on specific achievement results will generate meaningful and necessary pedagogical changes. Dr. Whitehurst described how effective dialogue about assessment results could be:

> Schools should be pulling together to make sure they are looking at the data and talking about their specific situations.... something as simple as collaborating and

> talking about student learning and talking about teaching will improve instruction. (Interview Transcript, Paragraph 41)

Dr. Pinkerton also cited how achievement data should drive curricular and pedagogic discussions as she recalled a meeting with elementary school teachers in her previous district:

> Our English people got together and said, 'What did your kids do on question two? Well, how did you teach that?' So we as area directors had to go and sit down in those meetings. So I remember being in a second grade teacher meeting at a very poor performing elementary school, and I remember these second grade teachers were sitting there, there was four of them. And they said, 'Why did your kids all get the regrouping of the numbers correct?' And the little young teacher, like a second or a third year teacher, looked up and said, 'I took them out to the buses. I made a song out of it. You ride a bus, 352. And said the '3' is in the hundreds; the '5' is in the tens, and the '2' is in the ones.' I got it. I am on bus 542, 'the '5' is in the hundreds; the '4' is in the tens. And she said all of them did that and all of the kids got the regrouping right the next time. (Interview Transcript, Paragraph 26-27)

Mrs. Greenly, Chief Academic Officer, stated that assessments "give us opportunities for conversations between grade level teachers to seriously talk about what they have been doing and coming to an understanding for what we need to do more of and what practices do we need to eliminate" (Interview Transcript, Paragraph, 16). In addition, Dr. Pinkerton believes the conversations must expand beyond the school walls and welcome the parents into the dialogue. Pinkerton described how the Quarterly data were shared with parents in her previous district:

> When [parents] came to our back-to-school night, we handed them the assessment piece and told them what the kids' weaknesses and strengths were. We went over it with the parents to discuss what they could work with their kids on at home as well. (Interview Transcript, Paragraph 37)

In the end, the Reo central administration believed that the confluence of external demands to utilize data (NCLB and Education Yes!) and the new internal assessment accountability philosophy of Dr. Pinkerton would lead to marked improvements in

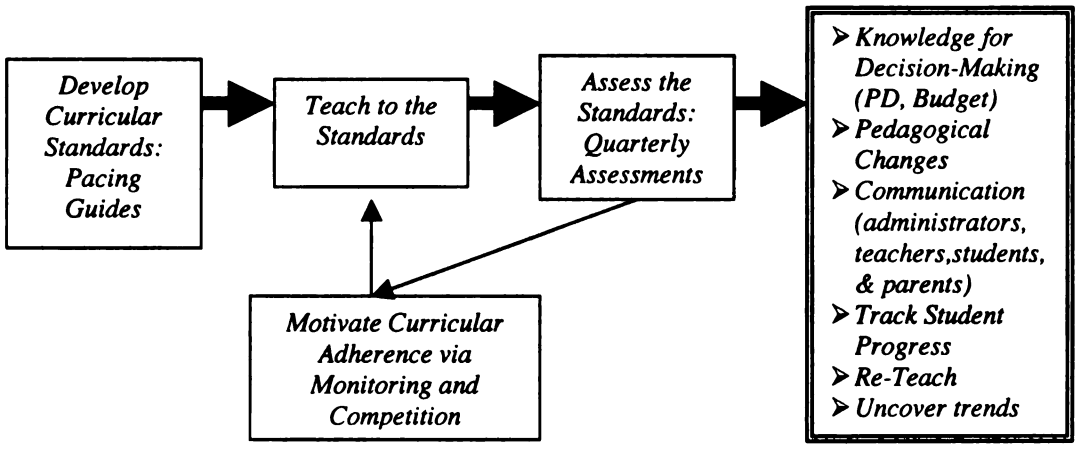
decision-making, curricular alignment, and student learning. Central office administrators believed that these pressures provided the district with important layers of accountability that were perceived to be lacking previously. The Reo central office administration saw assessments as a motivation, at times competitive, to inform decision-making about the budget and professional development needs, to monitor curriculum, and to begin communication about how teachers teach. Dr. Whitehurst, Assistant Superintendent for Curriculum, summarized the utility of assessments: "[They] monitor the curriculum if it is being taught, and the children are actually learning those standards and meeting those standards and benchmarks" (Interview Transcript, Paragraph 33). This set of beliefs provided both the catalyst for assessment accountability and the tensions of implementation. How these strong beliefs in assessment accountability were formulated into a district policy is the focus of the next section.

The Formulation of an Assessment Accountability Plan

While the external demands provided the impetus to use student achievement data, the central administration created an internal mechanism for the collection, analysis and use of student data in the district. That is, the Reo School District created its policy based on its own interpretation of the federal and local pressures to improve student achievement by making use of assessments (see Figure 3). This new system of assessment accountability was based on the belief that assessment data can inform pedagogy and monitor curricular adherence through mechanisms of communication and competition. Similar to Pinkerton's structural reorganization at central office, the new assessment plan would be modeled after the plan in her previous district where each core

content area had a specific set of curricular guides based on state benchmarks that were tested each quarter in grades 2-12.

Figure 3: The Reo School District's Theory of Action



The work to develop the accountability plan was developed and implemented in three phases. In 1998 the initial work entailed understanding the state's curricular framework compared to the district's scope and sequence of course content. By 1999 the second phase of work began to formally coordinate the district's curricular content. District writing teams were created to coordinate specific course content (by grade level) with the state standards. The documents created were called the District Pacing Guides. The most controversial and final phase of the district's accountability plan was the development of district specific assessments to evaluate student growth and teacher compliance in regard to the Pacing Guides.

District Pacing Guides

Pinkerton's work in the district began by accelerating the work of the writing teams to complete the Pacing Guides. Each Pacing Guide is accompanied with five principles:

1. The Pacing guides can provide students with an equitable learning process regardless of mobility within the district.
2. The Pacing Guides provide alignment and sequencing of curriculum, instruction, and assessments synchronized with MEAP and national standards.
3. The Pacing Guides provide a solid plan for instruction, reinforcement, and mastery of essential skills in addition to year long concepts, in a timely manner.
4. The Pacing Guides create a realistic time frame for instruction which is compatible with student growth and achievement and measurable using an assessment instrument.
5. The measurement of proficiency is continuous throughout the year and is assessed by the district at the end of each 9-week quarter.

(District Pacing Guides Introduction)

The district specifically outlined the utility of the guides for teachers, principals, area directors, and central office. For teachers, the guides were to be used to create learning goals, work in collaboration with colleagues to plan instruction and share ideas, and to re-teach, accelerate, and intensify instruction. For principals, they were to expect, encourage, and support their use and to monitor instruction to ensure sequencing based on the Pacing Guides. Area Directors also assumed the role of monitoring evidence in each classroom of the Pacing Guide's influence on instructional sequencing. Finally, central office was to provide training and support to district personnel, revise and update the guides as recommended by principals, teachers, and area directors, and to develop assessments that correlate with the guides.

The guides break each core academic subject area (math, English, social studies, and science) into nine-week units. Subsequently, each unit has a detailed account of essential skills to be addressed in each quarter. The skills are accompanied by a set of vocabulary words that should be integrated into the units. Finally, each guide has a content specific list of "continual process skills and year long concepts." The following chapter provides greater detail of the design and utility of the Pacing Guides.

Quarterly Assessments

Pinkerton instructed the writing teams to create internal quarterly assessments for the four core subject areas. Pinkerton stated, "If you are really doing what you are supposed to be doing in the Pacing Guides, you should do really well on your assessments" (Interview Transcript, Paragraph 124). As a result, the Quarterly Assessments were developed and implemented in less than a year. The Quarterly Assessments were officially instituted in 2000 in grades 2-12[22]. The Quarterlies assess students in math, science, English, and social studies every nine weeks[23]. The Quarterly Assessments were designed to assess the most important components of units taught. The tests were created to look like the state's MEAP. The content of the assessments is based on the district's Pacing Guides. Each test contains a multiple-choice section and an extended response. The test is administered in one class period (or less than an hour for elementary students). Teachers utilize a rubric, created by the writing teams, to score the extended responses. Teachers then send the assessments and extended response scores to assistant director of curriculum, Mrs. Rose.

Mrs. Rose's office scores and generates reports that are sent to department chairs (secondary) or grade level representatives (elementary schools). The feedback shows a school's comparison to other schools, classrooms, and students. Teachers receive a section item frequency report that details the percentage of responses to each item on the test. Individual student scores and responses to each item are also provided for teachers. Central office also internally publicizes a building comparison chart. The results and

[22] The math and English Quarterlies were implemented in 2000-01 and the science and social studies assessments were implemented the following school year.

[23] There is no Quarterly testing in certain quarters for grades 4, 5, 7, and 8 because of MEAP testing.

reports take about two weeks to be completed. Teachers are encouraged to utilize the assessments in a number of ways including their formal evaluation of students (card reporting purposes). In addition, teachers and department chairs are encouraged to utilize a district document that was created to initiate collegial discourse over the Quarterly results. These guides ask building educators to probe patterns in the data (including trends based on ethnic data); evaluate their use of the Pacing Guides; and reflect on lesson planning (including the use of multiple learning style strategies). However, schools and teachers are not mandated to share the results with students or parents.

The strategic initiative included a number of publicized documents including the district's Quarterly Assessment Purpose and Design document that states, "The assessments are designed to drive instruction." The document details the marriage of data and instruction, stating that the Quarterly Assessments are designed to:

- Provide student achievement data around rich conversations regarding curriculum delivery and student achievement
- Be an extension of classroom instruction
- Provide a model for classroom assessments that mirrors other assessments (e.g. MEAP, SAT, ACT)
- Monitor the curriculum

In addition, the document stipulates that the assessments are an inherent tool of accountability:

> We hold ourselves accountable to design, deliver, and assess a strong curriculum aligned with state standards and benchmarks. Teachers and administrators are held accountable to design the curriculum and deliver the instruction; students are held accountable to learn what is taught. Central administrators are held accountable to provide a tool for measuring student achievement, an analysis of achievement data, and professional development opportunities in areas of need as indicated by data analysis. (Quarterly Assessment Purpose and Design)

The use of assessments to fulfill the district's goals was not limited to the use of the Quarterly Assessments. A number of additional student assessments were targeted

for utilization for instructional improvement. The additional assessment strands included those that fulfilled federal and state mandates (MEAP and ITBS), bi-products of grant (DIBELS and PROM/SE), textbook publisher recommendations (Houghton-Mifflin), and additional district mandated assessments (Gates-MacGinitie). Each is described in regard to its design and purpose.

MEAP

The MEAP assessments were first administered in the state during the 1969-70 school year. The assessment did not weigh heavily on the function of schooling until it became tied to school improvement efforts (Public Act 25 in 1990). The purpose of the MEAP continues to be based on the measure of what the state deems all students should know in five content areas: math, reading, science, social studies, and writing. Currently, the state cites three uses of the MEAP: for school improvement purposes, for improved instructional practices, and as one piece of the state accountability model. The stakes associated with the MEAP were ratcheted up with the addition of a Merit Award for students who performed at the proficient level. Table 9 provides a summary of the scope and sequence of the MEAP assessment. The state provides schools with disaggregated scores by building, sub-group, and individual students. Currently, the disaggregated reports do not provide teachers with item-analysis reports.

While the state's accreditation system is relatively new, it has been in constant flux. Each year has brought new criteria for the accreditation system or the dynamics of the MEAP test itself. What subjects are tested and when they are tested continually

change, forcing districts to respond accordingly[24]. Recently, the Michigan Senate and House of Representatives voted to discontinue the 11th grade MEAP (HST) and replace it with a version of the American College Testing (ACT) as the student achievement measure for the state report card. All 11th graders would start taking the ACT in the 2006-07 school year. The new test includes college entrance exams in English, math, reading, science, social studies, and a "wraparound" exam with a work-skills component. The legislation would not affect elementary and middle school students who take the MEAP. Supporters of the new test, called the Michigan Merit Exam, argue it would better prepare students for college and the workplace, create a more timely and meaningful analysis, take less time to administer, and motivate students to perform well. The new assessment would be utilized to determine the state Merit Award. The state is also creating Grade Level Equivalency (GLE) standards that would replace the current state benchmarks.

Iowa Test of Basic Skills Test (ITBS)

The Iowa Basic Skills assessment is given to students once a year in grades K-9 in reading, math, and language arts. This nationally norm-referenced assessment is utilized in the district to comply with the Board of Education Policy #5121 adopted in 1975 that stipulates the district must use a norm-referenced assessment. In addition, the ITBS fulfills the NCLB mandate that stipulates all students in grade clusters 3 through 5 and 6 through 9 must be assessed in reading and math each year. The ITBS replaced the Metropolitan Achievement Test in 2002. In addition, the assessment fulfills the evaluation of student learning component for many federal grants. Specifically, the River

[24] Schools are constantly revising their curricular scope and sequence to match their curriculum with the state assessments.

High School Comprehensive School Reform (CSR) grant mandates a norm-referenced assessment. The assessment company provides disaggregated feedback in June of each year, too late for teachers to review the data. Moreover, the student data are not provided to following school year teachers. Currently, the ITBS data is utilized to place students. Students scoring below the 29th percentile in reading work are targeted to work with building literacy teachers. In addition, middle school honors programs and a special high school algebra course that provides students with an additional semester to complete the course use the ITBS for student selection.

Gates-MacGinitie Reading Test

The Gates-MacGinitie test is utilized in grades 4 and 5 and for selected students in grades 6 through 9. The test is given three times each year. The Gates was instituted to provide a district standardized reading assessment for the upper elementary level[25]. The test has two components: vocabulary and comprehension. The Gates is a norm-referenced reading assessment used to diagnose and monitor student progress. In addition, the instrument is utilized to place students in special reading programs in grades 8 and 9. Teachers in grades four and five receive timely feedback from the Gates assessment. The Gates provides scores with percentile ranks in reading comprehension based on national benchmarks. Feedback from the 6 through 9 assessments is provided to school counselors for placement purposes.

[25] Recently, the district has replaced the Gate-MacGinitie and DIBELS assessments with a new reading assessment, the DRA (Developmental Reading Assessment). Reading First schools must continue utilizing the DIBELS until the completion of the grant. Teachers will be provided with professional development activities for the new assessment.

DIBELS

The Dynamic Indicators of Basic Early Literacy Skills (DIBELS) is administered to all Reo District students in grades K-3. The DIBELS assessment is a component of the Reading First Grant in which thirteen of the district elementary schools participate. The assessment is administered three times per year. This skill-based assessment focuses on phonics and word recognition (fluency and comprehension). The first assessment (given in the fall of each year) sets a benchmark for each student. Consequently, individual grow is measured with each subsequent test. Teachers administer the test to each child in a one-on-one format. A teacher assistant or the literacy coach monitors the class during testing. Teachers can see the immediate results of the assessment because of its administration. The assessment is designed to measure and chart student reading knowledge and growth. In addition, scores are disaggregated by an outside agency. While only the thirteen participating Reading First elementary schools receive training and a literacy coach, all elementary schools are required to administer the DIBELS assessment.

Houghton-Mifflin

As part of the Reading First Grant requirement to establish a core curriculum, the district chose a Houghton-Mifflin textbook. Houghton-Mifflin was also contracted to provide fifty hours of professional development work. As a result, the publisher provides and asks the teachers (K-3) to administer a textbook-based theme test six times per year. Each test takes approximately forty minutes to administer. Because the Houghton-Mifflin assessment is administered and evaluated by classroom elementary teachers, it provides immediate comprehension feedback to teachers.

PROM/SE

In the spring of 2004, teachers were mandated to administer the PROM/SE (Promoting Rigorous Outcomes in Math and Science Education) assessment to students in grades 3-12. The assessment focused on science and math. Each test took forty-five minutes to administer. Athens University, the local university, sponsored the assessment and the district received funds for administering the assessments. In addition to the two subject area tests that teachers are required to administer, teachers must also complete a detailed survey of their math and science preparation, strategies, and background[26]. The district educators have yet to see the results of the student assessments or the teacher surveys. Phase II of PROM/SE will incorporate professional development activities.

Table 9 below summarizes the district testing in 2003-04 while figure 4, also below, provides a chronology of the federal, state, and local events related to assessment accountability.

[26] Teachers are asked to provide a detailed record of their pre-service teaching education including the names of course taken.

Table 9: Summary of Reo District Testing 2003-04

TEST	DATES	GRADE	Tests/ Year	DESCRIP-TION	Type of Data Collected	Feedback
Elementary MEAP	January & February	4 and 5	1 per subject	4- English and math 5- Science and social studies	Criterion Referenced Test-Student Learning Data	District, Building, Classroom, Student May & Oct.
Middle School MEAP	January & February	7 and 8	1 per subject	7- English 8- Math, science, and social studies	Criterion Referenced Test-Student Learning Data	District, Building, Classroom, Student May & Oct
High School MEAP	April-May	11	1 per subject	Reading, writing, science, and social studies	Criterion Referenced Test-Student Learning Data	District, Building, Classroom, Student May & Oct
Iowa Test of Basic Skills (ITBS)	April	K-9	1 per subject	Reading, math, and language arts	Norm Referenced Test-Student Learning Data	Building and Student June
DIBELS	Sept., Dec., & May	K-3	3	Phonics and word recognition	Criterion Referenced Test-Student Learning Data	Building, Classroom, Student Item Analysis/ Immediate
Gates-MacGinitie	Sept., Dec., & May	4 and 5; 6-8	3	Vocabulary and Reading Comprehension	Norm Referenced Test-Student Learning Data	Building, Classroom, Student Item Analysis/ Immediate
Quarterly Assessment	End of each quarter marking	2 –12	4 per subject	Math, science, English, and social studies	Criterion Referenced Test-Student Learning Data	Building, Classroom, and Student Item-Analysis/Two Weeks after Test
PSAT	October	10	1	Verbal, math problem solving, and writing	Norm Referenced Test-Student Learning Data	Aggregate Student Score
ACT/SAT	April	11 and 12	1	College entrance exam	Norm Referenced Test-Student Learning Data	Aggregate Student Score
Advanced Placement	May		1 per subject	Subject specific	Criterion Referenced Test-Student Learning Data	Aggregate Student Score/ June
PROM/SE	May	3-12	1 per subject	Science and math	NA	N/A
Houghton Mifflin	Sept.-May	K-3	6	English language arts	Criterion Referenced Test-Student Learning Data	Item-Analysis by Student/ Immediate

Figure 4: Chronology of Federal/State and Local Events Related to Assessments

Federal/State Event	Year	Local Event
MEAP Assessment was legislated	1969/ 70	
	1975	BOE Policy #5121 requires norm-referenced district assessment
	1976	District institutes the Metropolitan Achievement Test
Reauthorization of ESEA, *America 2000*, calls for the creation of a national and state report card utilizing assessment data.	1989	
State Public Act 25 requires detailed reporting of assessment data	1990	
Reauthorization of ESEA, Goals 2000, ties Title I funds to standards and assessments.	1994	
	1996	Reo forms an ad hoc committee to address problems including achievement scores
	1998/ 99	Work begins to align the district curriculum with state benchmarks
	2000	➢ Dr. Pinkerton is hired by the Reo BOE ➢ Leadership structure is reorganized ➢ Pacing Guides are implemented throughout the district ➢ Quarterly Assessments are implemented for English and math
Reauthorization of ESEA, *No Child Left Behind*, ties sanctions and incentives into a 100% proficiency target on state approved measures by 2014. In addition, Adequate Yearly Progress (AYP) must be reported by sub-groups	2001	➢ Quarterly Assessments are implemented for social studies and science. ➢ District receives a $6.4 grant to establish magnet schools.
Michigan adopts Education Yes! Accreditation to fulfill NCLB mandates. New standards and benchmarks are also instituted.	2002	➢ Thirteen elementary schools receive a Reading First Grant. Implementation of the DIBELS assessment begins K-3 ➢ River H.S. receives Comprehensive School Reform Grant and must utilize ITBS data to chart improvements.
State reports fifteen Reo District schools are failing to meet AYP	2003	➢ One failing AYP district elementary school is restructured and begins a new cycle. ➢ District participates in the PROM/SE grant programs which utilizes a set of assessments in science and math ➢ Pinkerton reorganizes district leadership for a second time
State reports eleven Reo District schools are failing to meet AYP State Senate approves adopting ACT to replace high school MEAP.	2004	Gates-MacGinitie and DIBELS (for non-Reading First schools) are replaced with DRA
State plans to unveil Grade Level Equivalency (GLE) benchmarks to replace current standards. The state is also considering replacing the MEAP a course content test at the conclusion of each grade based on the GLEs.	2005	Reading First Grant and CSR grants will be completed

Mechanisms of Assessment Support

The Reo district has created a number of mechanisms to support the utilization of student achievement data in its schools. The student achievement information mechanisms of communication, access and feedback, professional support, and supervision are detailed in this section.

Assessment Communication

The district plan explicitly calls for an increase in dialogue. The district central office administrators asked department chairs and building principals to lead the discussions. During a district wide school improvement meeting, Dr. Whitehurst encouraged school improvement teams (SIT) teams to give teachers a chance to talk about instruction:

> Dr. Whitehurst told SIT members that research indicates scores will improve if conversations about achievement data are allowed to take place. She encouraged teachers to use planning time wisely and department chairs to lead discussions about test scores and data. She stressed the role of department chairs to look at data and talk to teachers about it. (Field Notes, March 17, 2004)

To assist building principals and department chairs, the office of the Assistant Superintendent for Curriculum, Assessment issued a document titled "Guidelines for Assessment Analysis." These documents provide principals and department heads with guidance on beginning discussions on the disaggregated Quarterly Assessment results. The document states, "It is the hope that this exercise will raise awareness of your staff regarding instructional needs and delivery, thereby raising the achievement levels of students. *The key is to ask questions, explore, and investigate; seek answers*" (ASC Office Memo, 2002, italics in original). As previously mentioned, the discussion

document solicits discussion about the patterns from the data (including sub-group trends), about what was actually taught and omitted, and how the content was taught.

In addition, the school improvement plan each building annually submits to the district includes an entire section devoted to data review such as: standardized tests, alternative tests, student outcomes (e.g. graduation rates, attendance, suspensions, discipline, mobility), and authentic assessments (e.g. running record, student portfolio, teacher observation). The plan also requires the data to be disaggregated by socio-economic status, ethnicity, grade level, special education, and limited English proficiency. Dr. Whitehurst, Assistant Superintendent for Curriculum, indicated that the forms require buildings to look at their data:

> So the message has been to look at the data and you talk about it, analyze it, and decide how you can use it to reform instruction. That was one piece. So when they walk in as they do their school improvement plans, we get with the Evaluation Services and have them to come up with a chart for each school to show their data to present it. There is a work sheet that Evaluation Services has provided to help them disaggregate the information, and they are to use that information to determine their weaknesses and build their school improvement plan. So this has been happening about three years now. So many have grasped that idea and they are doing very well. Some are still struggling with it because they don't quite know what to do. We have had in-services for the last three years on data analysis and collecting data for principals. So, in starting with the principals and hoping that it flows down to the teachers is kind of what we have been doing. (Interview Transcript, Paragraph 35)

What began as an intensive and formal effort to stimulate pedagogical discussions about student achievement data soon changed as Mrs. Rose, Assistant Director of Curriculum recalled:

> We would generate the data; we would structure the conversation about the data; and then we would go back to our buildings with your principals and the liaison: a teacher liaison at the elementary level, one person at the middle school and high school (the department chairs were the liaisons unless for whatever reason somebody else was selected). So we had these conversations together, as a big group, on how to analyze our data and what we could look for. And then we went

> back to the buildings, and on a half-day we had a relatively scripted discussion on using the data and finding trends and that kind of thing. So for the first year, for 2001 and continuing on a little less frequent basis in 2001-02, we actually sort of designed the opening of the conversation, and then the buildings took it from there. But then in 2002-03, we had a much lower amount of money to be able to work with. And we had new initiatives, and we had... we sort of took the focus that the principals should really understand by now how to use this data. (Interview Transcript, Paragraph 7)

Mrs. Rose's experience at building staff meetings supports this notion that conversations about assessments by principals are overly simplistic and void of learning:

> So a principal may come to a staff meeting and say, 'Look at these bar graphs.' And they'll say, 'Look how we did on sixth grade social studies this quarter. We're here, and here's the district average, and here's the building with the top score. We need to be this building.' End of discussion. (Interview Transcript, Paragraph 75)

Assessment Access and Feedback

The multiple assessments that have been instituted in the district have yielded a large warehouse of data. However, central office respondents were concerned about a number of issues including the inability of the district to create a single data collection portal; the lack of multiple types of data to inform student achievement; issues of data accessibility; and the timeliness of the feedback.

To begin, there is no one single warehouse of data in the district. In fact, assessment analysis is housed in various offices within the district as well as in warehouses outside of the district. The Office of Research Evaluation and Program Assessment (REPA), who reports to the Chief Academic Officer, housed a number of assessments including the MEAP and Iowa results. The Office of Curriculum warehoused the Quarterly data, and outside agencies control the Gate and DIBELS data. As a result, the data are not relational or interoperable. That is, one cannot view all of the

assessment results for any given student at the same time in order to draw conclusions based on a number of variables:

> We don't even have at this point any place that you can go in by student number. I can go in our data by student number and tell you what tests the kids have taken and then look to see what their scores were and then click on that test and find out what the questions were, but we can't do that for MEAP, of course, and there is no place in our system where you can just call up a report based on including certain students and then finding all of their data. (Mrs. Rose, Interview Transcript, Paragraph 25)

Mrs. Rose also complained about what data were not collected:

> The data doesn't tell that half of my kids don't come to class, and they didn't do their homework. All the data tells is that this is an area where we aren't doing what we hoped as well as we hoped we would be, and we need to look more closely at what might be the problem. (Interview Transcript, Paragraph 69)

Mrs. Rose added that the friction between REPA and the Office of Curriculum might have contributed to the inability of the district to create a relational data warehouse:

> Evaluation Services have always been involved in any kind of outside testing, using any kind of outside test that is adopted or determined that we are going to use in the district and collecting data. But they have no connection to making decisions about curriculum. So they generate data, but the people are not instructional people, and they are not curriculum based. Their purposes they serve now don't include anything to do with our data. But, if we could come up with some kind of collection system that allowed us to put it all together, then they would have a connection to our data. (Interview Transcript, Paragraph 59)

Mrs. Lysander, district data consultant, believes that building user-friendly and accessible data reports goes beyond a single warehouse and paperless reporting. She believes that the philosophy of data being used as interoperable, as between various data types, and as the creation of individual goals for students is the route the district should embark upon. Only then can one track the progress and growth on a value-added, individual basis. However, the district continues to rely on student learning data. While some demographic data are utilized for analysis, there are other types of data that are

largely ignored, including perceptional data (from students, teachers, administrators, and parents) and process and program data.

Dr. Pinkerton stated that she understood the need to make the data more accessible. In fact, she hoped the system would be paperless:

> I guess, in retrospect, I am not sorry for the assessments. I am disappointed that we haven't technologically, due to financial constraints and the wiring of old buildings, put them in a way that teachers would consider them, within four years, a help. And I think technology would have helped us there, like they are doing in [another urban city in the state] where the kids can take it on line and get his own responses there. It is more of an instructional tool to use with kids. But here we have kept it in such a managed state that it is still seen as an extra layer of work, rather than help. And I don't think we will get out of that. (Interview Transcript, Paragraph 110)

The reaction to such claims at central office were mixed. Mrs. Rose who works most closely with the Quarterly Assessments believes the testing streams should remain intact so long as the state MEAP scores remain low:

> Our board members have asked at different times, is it really necessary that we do this? And my response has always been, if our MEAP scores, if our schools were running at 80 and 85 percent satisfactory we wouldn't need to look closely to make sure that we are doing what we say we do, but we are nowhere near that. So until we get closer, I think it is a good idea. (Interview Transcript, Paragraph 47)

However, Dr. Pinkerton has already begun to think about scaling assessments back in the district, "when you look at the testing, when you look at the Iowa and the MEAP and the DIBELS and the Quarterlies, it is too much" (Interview Transcript, Paragraph 24). Dr. Pinkerton stated that she is considering changing the Quarterly Assessments into semester tests (two times a year)[27]. This proposition would be welcome news to the building educators who all suggested this as a needed change in the district's assessments.

---

[27] The Quarterly Assessments beginning in the 2004-05 school have been moved to a semester assessment; however, the content of the assessments have not been modified to reflect this change.

Finally, the feedback mechanisms for assessments come from multiple offices as stated previously. As a result, buildings and teachers are receiving various forms of disaggregated data. Additionally, while the REPA office has a large staff to disseminate assessments, the Office of the Assistant Superintendent for Curriculum only has one person to administrator, disaggregate, and disseminate the Quarterly Assessments. Mrs. Rose, who is in charge of the Quarterlies stated, "I am just overwhelmed with getting the tests out, bugging principals and department chairs to get them back to me, and then to score them and get buildings the results" (Interview Transcript, Paragraph 213).

Assessment Professional Development

Three professional development practices anchored the district's strategy to use student achievement data. First, the implementation of the district's Quarterly Assessments was supported by the development of assessment writing teams. These teams of teachers were provided release time to assist in the creation of the assessments and to periodically meet to address issues and make modifications to the assessments. The teams consisted of two representatives from each grade level (elementary) or each subject (secondary). Second, teachers at grade levels (elementary) and subject areas (departments in the secondary schools) were originally allocated time to dialogue about the achievement analysis they received from the district office. Finally, the district supplied external resources (grants) and training to facilitate the utilization of student achievement data.

Both REPA and the Office of Curriculum Development provided building in-services on the utilization of assessment data. District administrators were privy to developmental activities that had moved beyond analysis to use. The administrative

training sessions that were offered by the district were externally based. The district contracted Mrs. Lysander from Athens University to work with a number of building administrators specifically on the use of student achievement data. This day-long training provided administrators with an overview of the various types of data schools gather as well as practical skills for analyzing their own building data. Building principals (n=7) responded that the training had been effective, but admitted that they were doing very little to analyze their school data or engage teachers in conversations about achievement data. Lysander described the administrators as being under-skilled and terrified of data in general. Additionally, because the turnover of building principals remains high in the district, a number of principals have not attended the training. As a result, this lack of consistent and "just-in-time" professional development was an inhibiting factor to principals' feeling comfortable to talk about achievement data:

> I think principals are uncomfortable with it. You can look at it and say I am in the 40th percentile, and I need to get to 49th percentile. I know I need nine points. But I don't feel they really digest it to be able to tell teachers other than, 'We have got to do better.' We have some principals that are very good. You need to know what your subgroups are. If we were just to answer two questions correctly, we would have been there. So it is really the principal feeling comfortable with that. (Mrs. Lysander, Interview Transcript, Paragraph 35)

At a recent administrative in-service titled, "The Effective Use of Data," the district's data consultant, Mrs. Lysander, told principals they should be asking teachers a series of questions:

- Do you know where your students are in relation to the benchmarks by the end of September?
- Which students need intervention?
- What evidence do you have of student learning?

(Field Notes, August 26, 2004)

In addition, she stipulated that principals should make inquiries about teacher capacity to utilize assessment data:

- Do teachers understand how to assess content and skills?
- Do they know how to teach what will be tested?

(Field Notes, August 26, 2004)

Mrs. Lysander asserted that principals should be using assessment results to drive professional development efforts and to understand student issues they come into contact with: "You should be using data to better understand issues like skipping school. Seek out perceptional data to better understand and uncover patterns of student behaviors" (Field Notes, August 26, 2004).

The district has also provided personnel support. In 2002, the district added two science specialists, in 2004 a math specialist. The purpose of these positions is to assist teachers in looking at student achievement data (MEAP and Quarterlies) and to help teachers devise new pedagogical strategies. In reading, each school in the district has a reading teacher. The elementary reading teachers work with classroom teachers utilizing the Gates assessment. In addition, the thirteen schools that are covered by the Reading First grant received LETRS (Language Essentials for Teachers of Reading and Spelling) training that is directly correlated to the curriculum. The LETRS curriculum utilized the DIBELS assessment to establish a baseline and to measure student growth. The LETRS curriculum states, "Instruction in assessment and evaluation of student performance will be embedded in the topical modules" (LETRS Document). Reading First schools are also provided fifty hours of training between 2002-05 including supplemental materials on how to utilize the DIBELS assessment in their practice. In addition, Reading First schools are allocated a full time literacy coach who meets with teachers weekly. The

training modules, literacy coach, and grade level teacher work in concert to understand the DIBELS feedback and administer necessary interventions (e.g. new strategies to re-teach and literacy coach group work with selected students).

River High School is also in the second year of a grant. The high school received a Comprehensive School Reform (CSR) grant in 2002. In order to comply with the grant, the high school had to implement a norm-referenced test (ITBS) and contract professional development services as well. River High School utilized the new funds to implement a block schedule. The school contracted with the Southern Regional Educational Board and its *High Schools That Work* (HSTW) program to provide training sessions that focuses on strategies to teach in the block.

Assessment Supervision

The supervision of the use of student assessments begins at different levels of the central office structure. As previously mentioned, the Chief Academic Officer focuses on the MEAP and ITBS assessments. As a result, she works with building principals and school improvement teams to best utilize the data. The Assistant Superintendent for Curriculum supervises the use of the Quarterly Assessments and the assessments required by grants (this includes the supervision of the building literacy coaches). In addition, Area Directors are the direct supervisors of building principals. These three central office forces have been mounted to help building principals and teacher-leaders (department chairs and school improvement chairs) utilize different sets of student-learning data. However, there is no formal evaluation of principals and teachers in regard to their effective use of data. For Dr. Pinkerton, this was by design. She had to be very explicit to the teacher's union in Reo that Quarterly data would not be an evaluative measure for

teacher performance. Pinkerton hoped that teachers would want to see and interpret this important data about their students as a means to improve their instruction. In fact, the superintendent has taken issue with many administrators who want to use the assessments to evaluate teacher performance and provide a competitive incentive for improvement.

Central office's interest in holding teachers accountable (albeit without evaluative measures) to the curricular framework placed building leaders in a highly pivotal position. Dr. Pinkerton made it very clear that the effective use of achievement data is a result of instructional leadership. She specifically mentioned three key leadership positions in the district: the area directors, building principals, and department heads. As a result, Pinkerton has a strong interest in helping leaders to become instructional leaders. The district's Chief Academic Officer specifically earmarked the building principal as the key player, "As goes the principal, so does the school. As we ask the teachers to learn all of the idiosyncrasies of their students, the administrator has to know all of the idiosyncrasies about the staff" (Mrs. Greenly, Interview Transcript, Paragraph 54).

The central office administration believed a limited number of building principals had the requisite capacity to become instructional leaders. Dr. Whitehurst considered only seven of the district's thirty-nine building principals as instructional leaders. Mrs. Grey, Area Director, also believed a small number, one-third, of the twenty building principals she supervises have instructional leadership qualities. In addition, Dr. Pinkerton was concerned about the ability of her central office staff (including area directors) to advance instructional leadership in building principals. Pinkerton provided an interesting anecdote about her practice of highlighting examples of instructional leadership practices at administrative meetings:

> Principals say, 'Please don't point me out anymore because [fellow principals] will tease me.' And I'll say, 'What do you mean?' A principal said to me, 'Dr. Pinkerton, every time you ever call me out, we're doing a great job, everybody is teasing me after the principal is leaving... They say [singing] 'I want to be like you... I want to be like you.' It is sad when you are in a school culture when you are doing good, it is negative. (Interview Transcript, Paragraph 75)

Dr. Pinkerton described that this type of "dysfunctional cultural ethos" became even more troublesome when combined with a lack of instructional leadership in the district:

> I think the assessments have never been taken to heart by the leadership of middle schools and high schools. Because, if a principal called in all sixth grade teachers by subject matter at the end of the assessments and said, 'Let's go over this and see how we did,' it would heighten the interest in what kids were learning, and what people were teaching, and, most importantly the methodologies that were used on kids who were achieving well. We have administrators for whom the instructional leadership part is weak. So these are seen as an extra layer of extra work rather than a way for staff to communicate at a staff meeting. (Interview Transcript, Paragraph 113)

Pinkerton also cited barriers of the principals' union and the inability of the Area Directors to effectively evaluate principals' as inhibiting the utility of assessment data to drive instructional decision. Pinkerton stated:

> All of my lieutenants are from inside. They all have twenty-five to thirty years. They are a part of the original system. If you ever plan to be a superintendent, if you plan to make major change and sustain the change, you have to have people with you who are new and believe in the change. Too often friendships are getting in the way of supervision. (Interview Transcript, Paragraph 59, 61)

Dr. Whitehurst suggested the district's teachers union as the most significant factor inhibiting the use of data:

> It is basically the union. Period. That is the biggest one because the union influences the teachers, and their attitudes, and what they choose to do, and what they choose not to do. So it becomes a challenge. We know the schools that are the challenges because they are big union schools. So we will walk in and say we would like, you to do this, this, and this. And we actually have some staff say, 'We will not!' So we have to find a way to work around that to make it work. That is the challenge. (Interview Transcript, Paragraph 113)

Conclusion

Dr. Pinkerton sold the assessment accountability plan on the premise and promise that student-learning data are a valuable learning tool for decision-making and improving pedagogical practices. The plan was not intended to be an evaluative tool but rather a necessary component for teachers to diagnose and dialogue about *how* they were teaching. Pinkerton moved the internal system of accountability from fulfillment of state standards toward in-house measures of achievement. The central office architects of the assessment theory of action assumed that the internal accountability system would first bring curricular coherence and consistency to the district and then steps to assess and evaluate would provide a mechanism that would trigger educational decision-making and pedagogical improvement. The Chief Academic Officer, Mrs. Greenly best summarized the belief and intent of assessment accountability in the Reo district:

> Without data you are just another person with an opinion. We must ask teachers to 'show me the data!' We are now moving in a direction to use data to support decisions and drive instruction. We are now making it a major piece of everything we do in our business about being schools. (Interview Transcript, Paragraph 48)

Simply put, the district's theory of action posited that teachers must understand the curriculum standards, teach to these standards, and assess student learning according to such standards. At an administrative in-service, principals were told that there was an 80% correlation of student success on assessments and teaching to the standards. The message to principals was simple, "Tell teachers what to teach; make sure they are teaching it, and make sure they are utilizing assessments to chart growth and uncover needs" (Field Notes, August 26, 2004). The district's theory of action evolved from curricular alignment to the creation of assessments as a measure to hold teachers

accountable. However, the interpretation of assessment accountability varied within the central office. As a result, three major tensions within central office have emerged. First, the definition of accountability was inconsistent throughout the district. Specifically, different district actors defined accountability as evaluation, supervision, or competition. Second, there was contention on the metrics of the assessments. The source and type of data gathered and disseminated determined its utility for administrators and teachers. Finally, the interoffice dissension around the collection, warehousing, disaggregation, and dissemination of the various types of student learning data impacted its utility.

The district's theory of assessment accountability had clear expectations for building principals and teachers. While the expectations were rooted in the implicit pressures to utilize data and not in an explicit evaluative mandate, the saliency of the student data issue has emerged. Dr. Pinkerton's belief that the data will cause teachers to reconceptualize pedagogical practices has been the center of great frustration:

> Teachers didn't like them [Quarterlies]. Because what are you doing when you walk in? We are all supposed to be on section four, chapter nine of the U.S. social studies book, and they are like, 'I didn't have to do that.' Whereas, now those guides are driving the instruction. And then we brought that assessment behind it and made it look like two whipping boys... And they are not getting it, I mean, they are matched. So, if you love the Pacing Guides, why do you hate assessments? (Interview Transcript, Paragraph 49, 124)

Prior to Pinkerton's arrival there was no articulates plan to use student assessment data. Specifically, little data was collected, and data that was collected was not utilized for decision-making or instructional practice. In totality, the multiple streams of assessment data make up a significant potential resource for instructional improvement; however, the question remains are they working in concert to provide administrators and teachers with aids to improve their practice? And, are building educators in Reo now

paying attention and utilizing data? How the mechanisms of saliency, communication, and competition played out in the schools themselves is the focus of the next chapter.

CHAPTER FIVE: THE SCHOOL RESPONSE

It's just a hell of a time to be alive, is all-- just this goddamn messy business of people having to get used to new ideas. And people just don't, that's all. I wish this were a hundred years from now, with everybody used to the change.
~Vonnegut, *Player Piano* (1952)

Introduction

The Reo district response to both the external (NCLB and Education Yes!) and internal pressures (arrival of Dr. Pinkerton and an underperforming school district) to improve student learning outcomes led to the development of a new assessment-based accountability plan. The previous chapter outlined the central office response to the pressures by first discussing the beliefs about assessments, the development and implementation of a plan of action, and the mechanisms of support, communication, professional development, and supervision. While the central office beliefs, plan, and support are vital for new programmatic changes in schools, the real transformation is rooted in the hearts and minds of school principals and teachers. Specifically, front-line administrators and teachers play a pivotal role in the implementation of policy. Consequently, the focus of this chapter is to understand the street-level perspective of the Reo response.

This chapter presents two mini-cases to illustrate the operational response to the district's assessment accountability plan. First, an elementary case provides a glimpse into the response and impact of the plan in the district's elementary schools[28]. Similarly,

[28] This case relied on the interviews conducted with participants at Wood Street Elementary School (principal, three teachers, literacy coach, and the literacy teacher). In addition, the interviews conducted with two additional district elementary principals and an additional elementary teacher informed this mini-case.

a high school mini-case will be presented[29]. In combination with the central office response, the mini-cases will provide an up-close view into the landscape of assessment accountability, its form and function, in the Reo School District.

The Case of Elementary Schools

Elementary education has been the focal point of many recent reform efforts. The belief that skill development begins at an early age has led to heightened demands of accountability at the lower levels of education. In Reo, the lack of consistent curriculum and low achievement scores led to the enactment of a number of interventions. Specifically, the Reo elementary schools have seen the addition of new programmatic infusions of curriculum and professional development. Many of the new programs have created new assessment demands.

Currently, the elementary schools in Reo have a number of assessments that are to be utilized in decision-making and practice. The assessment battery in Reo elementary schools includes: the district Quarterly Assessments (grades 2-5 administered four times per year in the four core content areas), the Gates-MacGinitie (grades 4-5 administered three times per year to assess reading levels), the DIBELS (grades K-3 administered three times per year to assess reading levels), the MEAP (grades 4 and 5 administered once a year to assess the four core content subjects), the Iowa Test of Basic Skills (grades K-5 administered once a year to assess reading, math, and language arts), the PROM/SE (grades 3-5 administered once a year to assess science and math), and the Houghton-Mifflin Theme Tests (grades 3-5 administered six times a year to assess English skills). This section begins with the case of Mrs. Templeton as a means to describe the impact of

[29] This case relied upon interviews conducted with participants at River High School (principal, assistant principal, counselor, and six teachers). In addition, the interviews conducted with two additional secondary principals and an additional high school teacher informed this mini-case.

assessment accountability on teachers at the elementary level. Subsequently, descriptions from other elementary teachers and principals will fortify this case.

The Case of Mrs. Templeton

Mrs. Templeton teaches a fourth and fifth grade split classroom with twenty-eight students. According to Templeton, split classrooms are becoming more prevalent as a cost cutting benefit to the district[30]. Because Templeton teaches a split class she must administer all tests scheduled for each grade level. In 2002-03, Templeton assessed the Gate-MacGinitie three times for each grade (six total tests), the ITBS once at each grade (two total tests), the PROM/SE once at each grade (two total tests), the MEAP for both grades (four total tests), and Quarterly Assessments three times per core content area[31] in two grades (twenty-four tests). In all, Templeton administered thirty-eight assessments last year.

Mrs. Templeton must cover all of the content outlined in the fourth and fifth grade Pacing Guides for English, science, math, and social studies. Templeton tried to follow the guides, but says that it just is not feasible to complete all of the objectives. As a result, the week preceding the Quarterly Assessment is spent teaching the objectives that were not previously covered. Templeton estimated that she spends approximately six hours on the preparation, administration, and evaluation of the Quarterlies each nine weeks.

In order to better understand the curricular requirements of Templeton, an overview of the fourth grade Language Arts Pacing Guide and Quarterly Assessment is

---

[30] Mrs. Rose at central office confirmed this claim. Budget cuts have created teacher lay offs in buildings. Consequently, an increase in split elementary classrooms has occurred. To date, thirty-six elementary classes are split.

[31] The grades 4 and 5 Quarterlies are not administered because of the grade level MEAP.

described for one nine week unit. The fourth grade Language Arts Pacing Guide has three units for the third nine weeks of study. The core theme and concepts for this period included the coverage of narrative text (fables, fantasy, and realistic fiction), informational text (course textbook), and key ideas, values, and beliefs (cultural diversity, self-reliance, flexibility, and acceptance). The essential skills for this unit included:

- *Narrative Text Characteristics:* Knowing how dialogue provides insights into characters actions and thoughts; mood; personification; simple flashforward and flashback; and lesson and theme.
- *Comprehension Strategies/Skill Development*: Make inferences about theme, characters, and plot; summarize main events; use multiple graphic organizers to analyze and respond to text; synthesize ideas and concepts from story.
- *Features and Pattern of Informational Text*: Know cause and effect; use legends; identify persuasion; and identify organization and structure.
- *Comprehension Strategies/Skill Development*: Summarize; know and use analogies; use graphic organizers to analyze and respond to text; generate questions before, during, and after reading; compare and contrast text within similar themes; synthesize ideas to gain understanding; and identify fact, opinion with reasons to support.
- *Written Communication*: For narrative text- Write a story from the first person point of view with protagonist/antagonist using plot, sequence of events, setting, and problem resolution; write a simple fable using simple dialogue and simple personification; use similes; and identify a lesson supported by details. For informational text- Take notes; write a report with topic, main idea, and supporting details; and write a persuasive essay support by facts.
- *Oral Communication*: Tell a story to a large group using expression, humor, appropriate pace and volume, eye contact, transitional words, and body language.
- *Mechanics and Grammar*: Spell words from 3rd quarter 4th grade list; know prefixes and root words- tele, graph, sign, spect, attract, and rupt; use commas to set off words and phrases; use apostrophes in possessives; and identify prepositions and prepositional phrases.

Finally, the "continual process skills and year long concepts" for fourth grade language arts includes a detailed list of strategies and suggestions for literature and understanding, skills and strategies, and performance tasks.

The third nine week Quarterly Assessment has four components. To begin, the teacher reads a story (in this assessment "Paul Bunyan") to the students. Student then answer the first five questions based on their understanding of the story. Second, students read a story from their text (in this assessment "The Algonquins of Maryland") and answer the next five questions. Third, students answer the final three multiple-choice questions that are comparative questions for the two stories. Next, the constructed response question asks students to complete a graphic organizer that required two examples from each story of the theme "Self-reliance." Finally, students are to write a paragraph about both stories. Students must agree or disagree that the stories share the theme presented and provide specific examples or details to support their thinking. Templeton must score the constructed response using a rubric. The completed tests and scored responses are sent to Mrs. Rose's office for scoring and report generation. Templeton described spending a great deal of time teaching students how to complete the Quarterly Assessments, specifically the constructed response assessment items. Mrs. Templeton found little utility from the Quarterlies.[32] She stated, "there is no time for remediation when I get the assessment results back because of all the objectives that need to be covered for the following nine weeks" (Interview Transcript, Paragraph 34).

While Templeton found the Quarterlies to be time intensive, the MEAP assessment devoured a greater amount of her instructional time. The MEAP has multiple-choice and a constructed response section on all four content tests that Templeton must administer (4th grade ELA and math and 5th grade science and social

[32] Because this portion of the Quarterlies is grade by teachers, Templeton finds little comparative validity in the responses.

studies). Templeton described the fourth grade English Language Arts MEAP tests as time consuming and difficult to administer and prepare for:

> The writing test provides me with a text section two weeks before the test. Students have to do a pre-write assignment then read the text. The MEAP then asks them a set of multiple-choice questions on the text. Then, there is a listening component where they listen to a tape and answer questions. We then have to have the students read stories that are similar to the text the state provides and students write about the common themes between stories. The next day the students read yet a different story and they have to write an essay. The essay has a number of parts: use evidence from both stories to uncover the common themes, what does the story mean to them personally, draw conclusions, and answer the questions 'how you could use the story in your life?' (Interview Transcript, Paragraph 132)

A review of the released MEAP 4th Grade Reading test supports Templeton's notion that the assessments are in-depth and complex. The MEAP scoring rubric stipulates a top score must exhibit the following:

> The student clearly and effectively chooses key or important ideas from each reading selection to support a position on the question and to make a clear connection between the reading selections. The position and connection are thoroughly developed with appropriate examples and details. There are no misconceptions about the reading selections. There are strong relationships among ideas. Mastery of language use and writing conventions contributes to the effect of the response. (4th Grade Reading MEAP Holistic Scorepoint Descriptions)

Templeton stated that preparing students and administering four of these MEAP tests (two in fourth grade and two in fifth grade) takes a substantial amount of time. Templeton estimated that two weeks of instruction are interrupted during MEAP testing. In addition, Templeton believed that these types of assessments are antithetical to the writing process that she teaches in her class. She stated that the writing process she was taught (professional development and pre-service teaching) utilized a different approach than is required by the MEAP. As a result, she believed that she spent more time teaching about the testing process than content itself.

Templeton stated that MEAP data are the only student learning data that are discussed on the school level. While the district provided grade level professional development. Templeton was concerned that there was a lack of cross grade level interactions. She deemed these crucial since a grade level assessment measures more than the knowledge that was obtained in the current year. Specifically, Templeton believed communication with third grade teachers would be beneficial for the students and teachers.

Templeton also administered the Gates-MacGinitie assessment three times last year for each grade. She found utility in the Gates for a number of reasons. First, since she administered the Gates herself, she received immediate feedback of results. Second, conversations about the assessments took place on a regular basis with the school's literacy teacher.

Finally, Templeton administered the ITBS assessment at both grade levels. While the assessment itself took less than two hours to administer, a week was spent in April teaching with an Iowa practice book. Templeton stated, "There was no correlation between the practice book and the Iowa test. But it did not matter anyway; by that time [April] the kids are just damn sick of tests" (Interview Transcripts, Paragraph 143). Templeton did not receive any feedback on the ITBS, but understood its utility to place students in reading programs the following year.

In the case of Mrs. Templeton, the 38 tests that are administered each year in her class account for over 100 hours or 15 days of instructional time each year, not accounting for the formal and informal classroom assessments she utilized.

Assessment Impact on Elementary Schools

At the elementary level, MEAP results drove schools to create artifacts in an attempt to raise scores. For example, the reading teacher and principal of Wood Street Elementary created an after-school MEAP Camp for 4th graders who scored below the proficiency level on the MEAP. Similarly, the principal of Paradise Elementary School, Mrs. Moriarty, created an after-school Science Club for students failing to meet the proficiency standards on the science portion of the MEAP. She described the need for the program based on consistently low MEAP scores:

> Science specifically was one area that our scores had not been out of the teens [percentile proficient]. And we had just brought on a new program to try to address that. So we structured that program so the teachers were still teaching the curriculum. The program was a hands-on experiential way of interacting with science. And we called it a Science Club. In that year our scores went up to 88.6 percent. I think it was. And we have not fallen below, I think, 75 percent since then. (Interview Transcript, Paragraph 25)

Mr. Lombardi, principal of Packer Elementary, also created special programs for students based on the achievement data:

> We added a lunchtime literacy program in our school. We had math club and a Writing Club also. We looked at that information and gave kids more experience with those things. Some of the studies tell us the more reading kids do on informational reading and recreational reading that the higher the score is going to be. (Interview Transcript, Paragraph 99)

However, many of the principals and instructors interviewed reported that the curricular design of the specialty courses had shifted to a test preparation mode. That is, the content of the special programs focused on test skill development rather than learning itself. The principal at Paradise Elementary School, Mrs. Moriarty, described how creating superficial test preparation programs also distracted educators from the real work of schools, a focus on student learning:

> I specifically remember the third year I was a principal. We buckled down, and we said we are doing everything, everything. Small group testing, bottled water, gum, extra pencils, packet of work for when they finish. We just strategized to the ends of the earth and our scores went down. And I remember the day I got them. I got the call, 'your scores are in.' I drove to pick them up. And I thought, 'This is it. Finally, we have nailed this.' And I opened it. And I remember thinking, 'What else can we possibly do?' (Interview Transcript, Paragraph 112)

As a result, Moriarty has begun to focus on discussions of improving student achievement in the classroom rather than creating environments that specifically target assessments.

Other elementary assessments were utilized to place students. The ITBS placed students in special reading programs at the elementary. The Gates and DIBELS assessment also are used to identify students who would work with the elementary literacy teacher and the literacy coach if enrolled in a Reading First elementary school.

Student leaning data in the Reo Schools were not confined to student placement. Two specific data streams indicate a promising utility of student assessment data in guiding instructional improvement for teachers. In particular, the elementary DIBELS and Gates assessments emerged as providing meaningful data for educators in the district. Mrs. Moriarty, principal of Paradise Elementary, stated, "While the fervor comes from the MEAP, I think the stuff the teachers really use are the DIBELS and the Gates-MacGinitie" (Mrs. Moriarty, Interview Transcript, Paragraph 100). Moriarty indicated that the utility of the Gates assessment was embedded in administration and feedback of the assessment, "I think that the teachers, just from giving it, seem to get a better read because it is a one-on-one time with the child. And I think they make some subjective decisions based on what they see the child scoring. And you know, with Gates-MacGinitie, I hear people talking a little bit more confidently about students' reading

levels" (Interview Transcript, Paragraph 77). Mrs. Pinecrest, reading coach, added, "The DIBELS gets right down into what exact skills the kids are missing. That is why they are useful" (Mrs. Pinecrest, Interview Transcript, Paragraph 96).

Not only did the elementary teachers administer the tests in a one-on-one fashion, they also received reading instructional professional development. LETRS training is directly correlated with the DIBELS assessment. As a result, there are numerous opportunities to discuss the assessment results and instructional strategies. At Wood Street Elementary, the teachers worked closely with the Reading First Coach, Mrs. Pinecrest, or the school's reading teacher, Mrs. Balsa, to better understand what the students needed to improve. Mrs. Pinecrest met weekly with grade level teachers (grades K-3) in her building. She described the meetings as being deeply analytical and cooperative. The teachers shared specific strategies as a result of understanding differences in the data reports from question to question and classroom to classroom. Mrs. Balsa, literacy teacher at Wood Street Elementary, summarized the utility of the DIBELS over other assessments:

> They [central office] need to be more careful about being choosy as to which assessments they give and what information they want to get out of these assessments and how they are going to use it. Just giving the test without making sure that the data is going to be used for something is pointless. And giving a test that is not going to produce data that can be used is pointless. Right now the DIBELS is useful because we are using it. They are looking at it. Mrs. Pinecrest produces the printouts; teachers look at it; Mrs. Pinecrest looks at it; I look at it, and we all try to work with that data and bring the kid up and give the kid those skills that he is missing. (Interview Transcript, Paragraph 109)

While teachers cited work with the reading specialists as useful, other communication strands were problematic. Field notes from a Wood Street Elementary

School Improvement Meeting indicated that this portion of the plan was given a cursory review:

> When discussing what numbers to put on the school's improvement plan someone arbitrarily said, 'How about a 10% gain.' After a moment of uneasy silence, a colleague blurted out, 'Sounds good—10% it is... let's move on.' There was no talk about specific strategies to address the goals. Additionally, there was no talk of ascertaining the utility of the baseline data. Specifically, how did the data compare from years past (trends); how did the data compare to the same students in years past (longitudinal data); and are there other types of data that would mitigate the results (mobility rates, etc.)? Rather, data are a necessary evil to fulfill district mandates. The notion that the data would stimulate rich discussion about student achievement, teaching pedagogy, or the necessary resources and developmental needs was absent. The use of data in this manner reminded me of Marlan Brando's famous line in *The Wild One* where he is questioned about what he was rebelling against- Brando responds, 'What have you got?' (Field Notes, Wood Street Elementary School Improvement Meeting, March 8, 2004)

A number of explanations accounted for this lack of concern about the multiple streams of assessment scores. To begin, teachers felt the assessments were imposed upon them with little input in regard to their current practices. That is, the pressures to fit the curriculum into their current practice became uncomfortable. Mr. Burch, elementary teacher, alleged:

> The way they keep pushing more and more stuff on us to take time and to diagnose the problem, it is kind of like being a doctor putting everybody in a doctor's office, 20 people, and have them say, 'Ahh,' at the same time. That is kind of what we feel like anymore because you can't take your time and say, 'Okay, this child needs this, this, and this.' But I am pressed by time to get so many things done. So my first priority is becoming to get this stuff done; it is not to diagnose these kids. (Interview Transcript, Paragraph 47)

More specifically, reports indicated the inordinate amount of time assessments took. A third grade teacher stated:

> Next week I am supposed to start giving the Quarterlies, and like I said, I have been busting my butt, and I am not even to the middle of fractions. But I have got two weeks to give this exam. So I have got to get all of my fractions in next week, so I can give the exam the second week and pray no one is absent. And then I still have a week and a half of school left, what do I do? What the hell do

> you do the next two and a half weeks when you are supposed to have it done in time for the Quarterlies? (Mrs. Oakley, Interview Transcript, Paragraph 39)

Mr. Burch at Wood Street Elementary was forced to make decisions about foregoing certain assessment data based on the number of tests that are required by the district: "I will be very honest. I have not used them [referring to the Houghtin-Mifflin elementary theme tests]. I used them two times this year. I don't have time for that. There is just no time" (Interview Transcript, Paragraph 143). Burch saw the merit in assessments, but believed the district must re-evaluate the current state of testing:

> I am going to be very brutally honest with you. When you are doing umpteen assessments, they all begin not to mean so much anymore. So it is kind of like, the more assessment, the better. You get more information about your kid. But most of your assessments are informal, and people forget that. Most of your assessment of a child is what you're watching them do and what they are doing. You give them two problems, and have them do it, and you look at it; are they handling it, or they are not? If they are not, do another lesson. So I think people overdo what assessment really is and what needs to be done in assessments. (Interview Transcript, Paragraph 35)

Time was also an issue with the elementary administrators. The elementary teachers, including the literacy coach and literacy teacher, have had to rely on informal meetings before and after school to address needs uncovered by the assessments. Mrs. Moriarty described meeting with elementary teachers during the day as being difficult:

> I set up grade-level meetings and memoed every teacher, please see me along your partner teacher to choose a time. And it was nearly impossible. One would come and say, 'I have gym class at 9:00, but the other one can't come at 9:00.' So, for lack of a cooperative meeting, I would meet with one teacher at a time. But it lost the whole element of everybody in a room together saying, 'Hey, I have an insight to somebody.' All of that was lost. And, truthfully I quit meeting with them one on one. (Interview Transcript, Paragraph 48)

Additionally, elementary administrators stated that the accessibility to good, concise data was inhibiting their ability to use data with teachers:

> Either I am going to have to say I am going to use MEAP data, and I am going to use "X" data; I am going to focus on these two things because I can't focus on everything, or else I am going to need some quick snapshots, three or four pages that really cut me to the chase of what I need to know (Mrs. Moriarty, Interview Transcript, Paragraph 49).

Consequently, elementary principals tended to rely on the reading teacher or literacy coach to interpret data. They delegated much of the responsibility to disseminate and articulate data results at staff meeting to the reading specialists.

Issues of test validity surfaced on a number of levels. Teachers had a growing concern that the assessments were not assessing students on the skills and knowledge they were taught in class. Mrs. Pinecrest describes the problem with the Quarterlies in this regard: "The Quarterlies only give you comprehension levels, not the specific needs and skill students are lacking. As a result, teachers must decipher the results. They are not trained to do this" (Interview Transcript, Paragraph 96). There is also the perception among teachers that utilizing assessments as a tool to measure the cognitive ability of students was an inappropriate:

> The whole point is you shouldn't test anything unless you know it. Otherwise, there is no point in testing. If you are trying to test in something that no one knows, then that is unfair. I believe that you should only test exactly what in the format and the way that you test if you teach. You cannot change that. That is not right. Testing is not about, 'Let's hide what is going to be on the test; I am not telling you,' and then surprise them. I am totally opposite. I believe that kids should know every single thing that they are expected to do and how to do it. If they can't, then I shouldn't test them. This is teaching 101- you test what you teach. And you don't go out to learn how to drive a car, and by the time you get your driver's license, they put you in a big truck. You don't do that. (Mr. Burch, Interview Transcript, Paragraph 73)

Mrs. Oakley, elementary teacher, added that how the questions were constructed was problematic as well:

> Sometimes it is a stupid question; sometimes it is worded stupid. I have been known to change the wording of a test myself while they were doing it. Because I

> thought it was stupid. Not because I didn't teach them the correct vocabulary, but because I said, 'This is what they really mean, guys.' (Interview Transcript, Paragraph 87)

Mr. Lombardi added that focusing only on basic skills can be problematic, "There are other skills one needs, like life skills. Those are the things that you need for your soul. That will make the quality of life better" (Interview Transcript, Paragraph 107). Finally, Mrs. Rosewood was concerned about assessments being culturally biased:

> Some of those tests are so biased. That bothers me. It bothers me that we have to give a test that we know a child can't perform good on. And because they are low, and we know that they are probably going to be retained or whatever, but we give them the test anyway. (Interview Transcript, Paragraph 161)

District teachers and administrators had also reported signs of student anxiety as a result of the multiplication of district assessments. Mrs. Moriarty, Principal at Paradise Elementary School, described the common practice of hyping assessments, especially the MEAP tests, to get students excited and comfortable enough to take the test. She reported that she had stopped this practice because it was creating more stress than relaxation and less focus on content understanding than on test taking instruction:

> I don't find any benefit in a pep rally or decorating the hallways or having the little kids come and tell the big kids, 'Hope you pass that test.' And I did that the first few years. We did that at nauseam. How many different ways can we get to celebrate the MEAP and pump up the kids? You are not pumping up the kids, you are stressing out the kids. (Interview Transcript, Paragraph 95)

Second grade teacher, Mrs. Rosewood, provided a sad account of a student's throwing up before taking the DIBELS assessment in her class last year. Mr. Lombardi, principal at Packer Elementary, believed that the competitive nature of our society and the competitive use of assessments in the district are perpetuating a dangerous behavioral pattern:

> Unfortunately we are a society based on winning, and unfortunately it is winning at all costs or any costs. And it is simple, as if you are in the Olympics and you get a bronze medal, you are nothing. I mean, by just the accomplishment of being there, that should be celebrated, but we don't.... I have talked to kids after the soccer game, 'How did you play?' They say, 'We lost.' I would say, 'I didn't ask you that. I asked you how did you play? Did you do your best?' (Interview Transcript, Paragraph 37, 125)

Teachers also felt anxiety based on the assessment. A number of teachers reported that the assessments focused too much on teacher practice and not enough on a number of other variables that affect assessment outcomes. Mr. Burch, teacher at Wood Street Elementary believed the high mobility rate of students must be taken into account:

> The year before last we scored 82% on reading scores. But what somebody doesn't realize about Wood Street School is that 51% of our population is transient. It has been more in the past. In the 5th grade, twenty-one kids are new this year. So how can I be held accountable for their scores? (Interview Transcript, Paragraph 109)

Burch went on to point to the lack of parental involvement as a major factor in the district:

> I think if you want to make kids' learning better, we have to understand why don't parents care about their kids' homework? Why do kids come in the second grade, and they can't even pronounce even one word in a book? And they can't read a lick, but nobody cares in their family. Hello, where is the break down? (Interview Transcript, Paragraph 147)

Mrs. Oakley agreed that the current focus on assessment data did not account for the current state of the students' abilities or the progress they make, "When the child comes to school at three years old ability level... If I get him to four by the end of the grade, I have succeeded. But yet the scores still are nasty" (Interview Transcript, Paragraph 101). As a result, teachers remarked feeling de-valued by the assessments:

> When you are talking about devaluing, yeah, I feel very devalued by tests, because it is all boiled down to a test, and that does not at all explain the picture whatsoever. (Mr. Burch, Interview Transcript, Paragraph 39)

Elementary teachers also pointed to the constant addition of new programs, curriculum, and tests that further complicate the ability to utilize data. Stand-alone assessments have also entered the district as a funding mechanism. In the spring of 2004, the district required funds to administer the PROM/SE. Neither teachers nor building administrators had any idea of the purpose of the assessment. Neither knew of the feedback they would ultimately receive:

> Not that I have seen the groaning, and I have asked people what it is, and nobody knows what the PROM/SE Test is, besides the fact that we are giving it to the kids for reading and math, and that Athens University is somehow going to provide a stipend to the district. Is that common practice for someone to say that we are going to have a test on this day, but you don't know what it is? (Mrs. Tiber, Interview Transcript, Paragraph 117)

Teachers have come to understand that new programs and mandates will come and go. An elementary teacher stated:

> We are so used to getting new programs for a while; and you know how they come and they go. So everything was kind of stand back, and wait and see how this thing went. (Mrs. Pinecrest, Interview Transcript, Paragraph 69)

Teachers have often returned to a new school year with a new curriculum. One elementary teacher described this:

> We know that we are getting new math next year. Why the hell can't you get us the teacher manual now, so we can take it home and look at it over the summer? Do they think that we are all that stupid that we're going to lose the manuals over the summer? But I feel nothing but anger when I walk in in the fall, and I have something brand new like a math series, and I am supposed to start using it, teaching it, and be familiar with it all at the beginning of school. That is just stressful as all hell to me. And I felt very stressed this fall when we had the new reading from Houghton-Mifflin. I dragged it home. I tried to get familiar. It took me the month of September to make myself familiar with all of the different parts. (Interview Transcript, Paragraph 117)

The theory behind the district's assessment plan was creating the ability to diagnose specific achievement results and then to take corrective action. Corrective

action was most frequently characterized as re-teaching; however, teachers cited a paradox between the expansive curriculum outlined by the Pacing Guides and the opportunities to re-teach. According to teachers, this has led to the didactic teaching practices, at best, and teaching to the test, at worst. One teacher stated, "There is no time to re-teach it. We are lucky to get the teach part in" (Mrs. Tiber, Interview Transcript, Paragraph 106). Teachers also discussed reviewing specific test questions with students. One elementary teacher stated, "Oh, hell, yeah. I pulled a test. In fact, it is sitting here. Now, what did I not cover that is on the test" (Mrs. Oakley, Interview Transcript, Paragraph 93)? Mr. Burch added, "I am going to make sure that if there is something that I didn't cover, I am going to give you the answer on that. I have no problem with that" (Interview Transcript, Paragraph 71).

Finally, teachers felt disenfranchised from the district's policy of assessments. The writing teams that were created for the district's Pacing Guides and Quarterly Assessments were never to be disbanded; however, they have rarely met. A number of the teacher participants in the study were original members of writing teams in the district. While these teachers had a better understanding of the purposes of the district's accountability plan, they became disengaged with the plan as their roles diminished over time:

> That is part of the reason why our attitudes are so poor about these tests, because we had a big meeting the end of last year about the Quarterly tests. Dr. Whitehurst, Mrs. Rose, and representatives from all three high schools were present, but nobody listened to our concerns. The bottom-line: no moratorium on the tests. You are going to be giving them without it and the general consensus was, 'You are a bunch of greedy teachers who want more money to rewrite these tests.' (Mrs. Rosewood, Interview Transcripts, Paragraph 155)

In the end, Elementary teachers in the Reo district utilized student assessment to guide instructional improvement when they found value in their practice and received support. Compliance to administer assessments alone did not lead to use. The elementary teachers were not shirking their responsibilities when it comes to test compliance; however, the intended meaningful use of student achievement data was reserved for assessments that made sense to their teaching, their time, and their students.

The Case of High Schools

While elementary schools have been the focus of many programmatic reform efforts, secondary schools, especially high schools, have been at the center of the public debate about accountability stakes. More specifically, the stakes associated with graduation tests have brought high school assessment results to the fore. While the state of Michigan does not currently have an assessment requirement for graduation, high schools have felt the pressure to perform well on assessments due to their weight on the state's Adequate Yearly Progress formula. In addition, financial pressures associated with the MEAP Merit Award may promote anxiety with students and parents.

The assessment battery in the Reo high schools includes: the district Quarterly Assessments (grades 9-12 administered four times a year in the four core content areas), the PSAT, ACT, and SAT (grades 10 and 11 administered once a year to assess college entrance levels), Gates-MacGinitie (grade 9 administered three times per year to assess reading levels for identified groups), the MEAP (grade 11 administered once a year to assess the four core content subjects), the Iowa Test of Basic Skills (grade 9 administered once a year to assess reading, math, and language arts), and the PROM/SE (grades 9-12

administered once a year to assess science and math). This section begins with the case of Mrs. Pilgrim to better understand assessment accountability in the high school setting.

The Case of Mrs. Pilgrim

Mrs. Pilgrim teaches basic algebra and a special algebra course that covers the algebra curriculum over a year and a half (students are identified for this class by the 8th grade ITBS). The size of her classes range from a high of 37 and a low of 17 (just under 30 average enrollment) and is comprised of students in grades 9 through 12. Pilgrim stated that many assessments are not administered in her class, but they impact her instruction by virtue of students leaving to take tests: "every test in school affects my classes" (Interview Transcript, Paragraph 12). Specifically, the Gates-MacGinitie, the PROM/SE, and the PSAT are example of tests that remove students from her classes. She stated that the 9th grade students who miss class three times a year for the Gates are students that "cannot afford to miss class" (Interview Transcript, Paragraph 45).

Mrs. Pilgrim determined that the ITBS test takes an entire week of instruction out of the school year. She, along with other math teachers, prepared students for the ITBS days preceding the assessment. It is the case with the Gates, ITBS, and PROM/SE that Pilgrim does not receive any feedback. While Pilgrim finds these assessments problematic in regard to losing instructional time with students, their effect on her teaching pales in comparison to the Quarterly Assessment and the MEAP tests.

Mrs. Pilgrim stated that there was constant attention paid to the Pacing Guides because of the looming Quarterly Assessments. Pilgrim described spending time before each Quarterly reviewing for the assessment. She also had difficulty motivating students to take the test seriously. Pilgrim finally devised an extra-credit plan to motivate

students. Because Pilgrim uses the assessment for credit, she must also take the time to grade the assessments. Finally, Pilgrim does not utilize the data that is analyzed by the district. This is a result of already knowing the results of the assessments and also the absence of a structure that fosters informal or formal conversations about the data with colleagues, her department chair, or building administrators.

Based on the Algebra Pacing Guide, Pilgrim must cover twelve units of study each year. The Algebra Pacing Guide has three units of study per nine weeks. Looking specifically at the first nine weeks, the guide's units include: basic operations with variables, basic operations with real numbers, and linear equations. These units have fifty-seven vocabulary words that must be covered. The guide also delineated the skills to be assessed in the Quarterly Assessment as covered in the first nine-week period. The skills included:

- Evaluate expressions containing variables and exponents.
- Evaluate real number and algebraic expressions using correct order of operations.
- Solve equations and inequalities using mental math.
- Develop algebraic models given real-world scenarios.
- Interpret tables and graphs.
- Construct and interpret appropriate tables and graphs given various data.

Finally, the Pacing Guide listed a set of "continual process skills and year long concepts" which included: communication, reasoning, connections, representations, and problem solving.

The Algebra Quarterly Assessment for this same quarter had ten multiple-choice questions and three constructed response questions. The content of the questions aligned with the previously described Pacing Guide for this nine-week period. The constructed response items require students to show their work in problem solving which includes the

creation of a double bar graph utilizing information from a data table provided. Pilgrim must evaluate the responses and report the scores to the district office. Like the elementary Quarterly Assessments, the high school algebra Quarterly is modeled after the state's High School Test (HST) for math which includes multiple choice and constructed response items.

Pilgrim cited the MEAP as the only assessment that is discussed in department and staff meetings. She stated that because of the stakes placed on the results, specifically Adequate Yearly Progress, everyone in the school paid attention to the MEAP. The high school math departments in the district utilized the "Buckle Down with MEAP" review guide that was developed by the state. The review occurs for five consecutive Wednesdays leading up to the MEAP. During the Wednesdays, teachers practice multiple-choice tests and constructed response items based on the content they believe will be covered on the MEAP. Pilgrim stated:

> While much of the content that we think will be on the test we have covered, a number of concepts we have not. So, we have to spend time teaching new content that we think might be on the test. But, most of our time is teaching the kids how to take the test. (Interview Transcript, Paragraph 111)

Because Pilgrim has students in grades 9 through 12, many of the students in the class must sit engaged in the 11th grade preparation for the MEAP. While she does see some value in this (preparing students for taking the MEAP style of assessment), because the content was at the 11th grade standards, many of the students were unable to participate in a meaningful manner.

In sum, Pilgrim cited that nearly twenty-seven hours of instruction were utilized for district assessments: ITBS, five hours; Quarterly Assessments, twelve hours; and MEAP, ten hours. In addition, if you take into account her own formal class evaluations

(about three per semester), students are actually being assessed or prepared for assessment in Pilgrim's mathematics classroom for thirty-nine hours. By Pilgrim's account, 17.6% of her classroom instructional time (39 of 182 school days) is devoted to assessment administration, evaluation, or preparation. These numbers do not include the assessments which take students out of class: PROM/SE, two hours; Gates, three hours; and PSAT, two hours (this increased the assessment time to over 25% of total instructional time).

Assessment Impact on High Schools

At the high school, achievement data was utilized to place students in special classes. The MEAP results helped place students in special programs including the extended algebra class taught by Mrs. Pilgrim. River High School created a similar program for English language arts. The high school counseling staff scheduled students into a supplemental 9th grade reading course based on scores from the 8th grade Gates-MacGinitie reading test. Mrs. Charles, River High counselor, stated, "If they had low scores on their reading, we double dipped their reading class. They have an English and a reading class to help improve their skills" (Interview Transcript, Paragraph 53).

Teachers welcomed the curricular alignment associated with the first phase of the accountability plan, the Pacing Guides. A high school teacher stated:

> A lot of people complained when we first did them. I mean, because I was in it from day one, when we started writing the Pacing Guides. It doesn't tell you how, but it says you need to teach causes and effects or you need to cover this vocabulary. So it keeps us aligned with state standards... That is what the Pacing Guide is to me, a map. And I just frame it out. And I know what has to be covered because of the Pacing Guide. (Mrs. Tiber, Interview Transcript, Paragraph 240, 158)

One of Mrs. Tiber's teaching colleagues agreed with the utility of the Pacing Guides by stating, "So really it has held us accountable... that is a good thing" (Mrs. Rhine, Interview Transcript, Paragraph 185). Teachers utilized the Pacing Guides as a guidepost to create their own planning of curricular scopes and sequence for the academic year. In fact, when teachers were asked about which artifacts they used for lesson planning, a majority of teachers indicated the effectiveness and convenience of the Pacing Guides. However, teachers were concerned with the breadth of content that needed to be covered in each class. The mandate to teacher a number of curricular objects ultimately affected their pedagogical practices.

While the Pacing Guides did not fully intrude on the teachers' interpretation of what should be taught, the second phase of implementing the Quarterly Assessments had a profoundly different impact. The Reo teachers felt the Quarterlies made explicit curricular demands. That is, choices had to be made between what they wanted to teach and items that would be tested. As a result, the teachers' felt that the constraints of the Quarterlies infringed on *when* they were teaching items and *what* items they had to teach. Teachers cited concerns about exchanging their curricular aims of interpretation and inquiry-based learning with a fact-based approach because of the nature of the assessments. River High English teacher Mrs. Rhine described the pressures to adjust the curriculum based on assessments:

> Like right now, for example, we have not finished the last major unit *A Raisin in the Sun.* And we just started that today. We will finish *Raisin in the Sun* by Wednesday of next week. Friday is the last day of the marking period. So I then have to throw at them this last Quarterly test, review for a final exam, and give them a final exam the following Monday. It really is tight. And I think that is partly my fault because there are some things that I like [to teach]. [For example] *Of Mice and Men;* they just love it. I have four different versions of the *Mice and Men* movie, analyzing and critiquing the different characters and how the

> directors and the producers depicted the characters in four different ways. The kids adore it. (Interview Transcript, Paragraph 87)

The content specificity of the assessments was also troublesome to teachers. For the most part, teachers in the district question the utility of assessment questions that are overly objective. Mrs. Rhine, River High School teacher, stated that she did not utilize Quarterly achievement analysis because the objective questions are too content specific. She stated, "They are way too driven by the text itself, not general enough, or not thematic enough. We don't want details anymore... I don't care about detail, I want you to know theme and mood and tone and those kinds of things. I don't care if you don't know the last name of one of the characters; that is not a big deal" (Interview Transcript, Paragraph 93). Another River High teacher, Mr. San Juan, provided a specific example, "On the recent Quarterly Assessment on Eastern Europe, there were an inordinate amount of questions on the Balkans. I thought that was some people in the district's personal interest, and I think it ended up being reflected a little too much in the assessment" (Interview Transcript, Paragraph 47). Mr. San Juan was describing the third quarter assessment that was to cover two units: Europe (3-4 weeks) and Africa (6 weeks). Upon reviewing the Pacing Guide for San Juan's class, only one of the ten European objectives mentioned Eastern Europe; however, 25% of the Quarterly Assessment questions focused on Eastern Europe.

Similar to the elementary case, high school teachers spend time covering assessment content prior to the Quarterly test. Mrs. Rhine stated the Quarterlies forced her to change how she taught days leading to the test:

> When it comes two days before the Quarterlies, and we know that we haven't spent any time on this, you tend to teach in a didactic way, sit-and-get to just that,

> which is clearly teaching to the test for a couple of days. (Interview Transcript, Paragraph 70)

This caused the high school teachers to believe that the district assessments were in direct contradiction to their pedagogical belief system of how the curriculum should be taught. Teachers had begun to characterize the movement of their pedagogy from one based on higher order thinking to the uncomfortable approach of didactic teaching:

> You get to the point where you say, 'I just need to do more direct instruction just to make sure that I cover these things so they are ready for the test'... I would like to do more higher order thinking skills. The tests sort of bring it back in the other direction. Because, if you want to do well on the test, the best way to do that is through rote instruction, memory, repeating drills. (Mr. Thames, Interview Transcript, Paragraph 140, 145)

Teachers became puzzled in regard to utilizing multiple pedagogical strategies and the organizational pressures of teaching to the content:

> What is funny, too, is that all of the educational circles and the universities are teaching all of this project-based learning. And I am a major advocate; at the end of each of our units, there are all of these options that kids have. I have kids burning CD's and videos and making these cool videos; it is wonderful. But then, they are responsible for objective tests. So, on the one hand, the powers that be, the educational professionals, are telling us that we need to be more eclectic, and the district is saying something different. (Mrs. Rhine, Interview Transcript, Paragraph 79)

Interviews with department heads indicated that conversations about Quarterly Assessment results were originally to be incorporated into department meetings after the implementation of the tests. However, these conversations have not occurred in the past few years. In fact, teachers interviewed could not recall any recent conversations that strategically discussed how achievement data could affect teaching. And, when conversations did occur, they were programmatic in nature. That is, more discussion is held about how to change *what* is taught in certain classes. Absent are the pedagogical

dialogues about *how* to teach. Mr. Nile, principal of River High School agreed that communication about student assessments was the key to making changes:

> This is the type of discussion that we have once we create that culture in a non-threatening way. But if I am prepared, if I did well, or my students did well, and yours didn't, then I am saying look at Mr. Nile, 'He did great.' Well, I could do great, too" (Interview Transcript, Paragraph 83).

However, Mr. Nile admitted that he has not met with department heads on a regular basis, "I don't meet with all of them. But I speak with department heads regularly, not always. I need to meet with them regularly on the Quarterly Assessments; I don't" (Interview Transcript, Paragraph 43). Mr. Trout, Principal at Kilgore High School, replicated the central office strategy of producing charts to motivate teachers. Trout described posting charts and telling teachers, "'Here's what the city averaged, here's what our kids averaged; here's how we compare to the other two schools.' Put a little competition in there. That pisses them off" (Interview Transcript, Paragraph 100). All of the secondary principals interviewed delegated the responsibility of student data to an assistant principal. Mr. Caulfield, Principal of Salinger Middle School, exemplified the problematic dissemination of achievement data:

> Basically, what I do is I give it [the data] back to the teachers and ask them to tell me what they think it means. But basically it breaks down the percentage of kids that got A, B, C, D, E. And you can kind of make your own conclusions from that, at least have dialogue with it. (Interview Transcript, Paragraph 35)

This description of limited use of data by building principals was explained by the lack of user-friendly data reports. Mr. Trout, principal of Kilgore High School, described the various places and people he needs to contact to receive data:

> It depends on what I need. For the Quarterlies, you go to one office. If you want MEAP, then you have to go to another office, and if you want any baseline data, especially data teachers grade etcetera, then you have to go to someone else. So

> you have got at least three different places that you have got to draw from. (Interview Transcript, Paragraph 79)

During a recent administrative in-service on assessment data, principals were unable to accomplish the task of data analysis because of complaints of "bad" data. That is, the administrators were not provided the necessary achievement data to fully understand student needs. As a result, "the strategy session turned into a griping session about how the data itself was problematic, not about the teaching practices or student learning" (Field Notes, August 26, 2004).

Many high school teachers admitted the practice of not administering the Quarterly test at times. In addition to their concerns about the content and timing of the assessments, teachers cited the feedback mechanisms as an inhibiting factor to the use of data. River High social studies teacher Mr. San Juan, social studies teacher at River High School, stated:

> I believe that if data are brought back to the participants in a timely and organized manner, that it can help to clarify objectives that possibly were not accomplished or procedures that possibly could be fine-tuned to increase efficiency and knowledge. (Interview Transcript, Paragraph 25)

Mrs. Hudson, teacher at River High School, also addressed the timeliness of the assessment results:

> The Quarterlies... I don't get them back in time in order to re-teach. That was the original design, to say, 'Okay, the majority of my students do not understand semicolon usage, so, therefore, I am going to go back and re-teach this.' Well, that is in an ideal world, and we don't live there. (Interview Transcript, Paragraph 99)

Finally, Mr. San Juan, added that if re-teaching were even a possibility, additional assessments would be needed to ascertain the effectiveness:

> But even if we do go back and cover the area that appeared to be weak, there is no reevaluation to determine whether the second approach worked better than the

> first. But that gets into more testing and less time, so I don't know where the limits are on that stuff. (Interview Transcript, Paragraph 75)

As was the case for Mrs. Pilgrim, some teachers incorporated the Quarterlies as a student grade; however, this was the exception, and those that did had a number of qualifiers. Mr. Thames, River High School teacher, acknowledged that he had to throw out a number of questions on the test if he felt he had not appropriately covered the material. Mrs. Rhine, high school English teacher used only the essay portion of the assessments, "The essays I must say are pretty good. We really worked those out so the essay parts always count. I grade the essays myself and those go in my book" (Interview Transcript, Paragraph 93). Mr. Seine, high school science teacher, used the Quarterlies in his grade book, but stated, "I built it in as a certain percent. So I would round it up dramatically" (Interview Transcript, Paragraph 133). Many teachers also stipulated that the students would not take the test seriously unless the threat of use in the teacher's grade book was established. Mrs. Tiber, high school social studies teacher, stated that, "Students don't find them relevant unless the teacher pushes the relevancy in their grade book" (Interview Transcript, Paragraph 49).

High school teachers, like elementary teachers, questioned the validity of the Quarterlies. Mrs. Rhine felt that the practice of teachers grading their own tests was problematic:

> We are grading the writing of the essays completely ourselves and we have heard all of this scuttlebutt that teachers are saying things like, 'Why would you give these kids anything but the perfect score on their writing? You are doing yourself a major injustice.' And I would never do that. But, if there are other teachers who, in fact, are doing that, what validity do these things have? (Interview Transcript, Paragraph 69)

Mrs. Hudson added:

> This is a test [Quarterlies], mind you, that is not field-tested. This is an invalid assessment. It is a measure of what one teacher or a small group of teachers think that the average student should be able to do and respond to it accurately. And even if they at least field-tested in the district, at least it would be a little more valid. (Interview Transcript, Paragraph 45)

Mr. Seine, high school science teacher, wondered how a written assessment could be utilized to measure and mimic important concepts such as scientific understanding.

Both principals and teachers reported implicit pressures in regard to the data. Teachers claimed to have heard of a colleague who was questioned about assessment results. However, beyond informal comments like, "Got to get those scores up" (Mr. Thames, Interview Transcript, Paragraph 121) from passing administrators, the teachers did not report any formal repercussions from poor assessment results. Nonetheless, teachers felt a looming anxiety that formal evaluation based on results was on the horizon. The only tangible form of accountability for teachers came from the certification issues of the "highly qualified teacher" clause stipulated in the No Child Left Behind legislation. That is, there was a growing concern that teachers might have to gain additional certification or risk being displaced because they might not be teaching in their certified content area. One teacher explained:

> We haven't really had a chance to review No Child Left Behind except for the threats that were put onto your expert certification, etcetera, etcetera. So the staff is more concerned with the staffing and the new criteria and special certification as it involves job security. And that has been the focus. (Mr. San Juan, Interview Transcript, Paragraph 99)

Teachers reported that assessments might be utilized to weed out teachers[33]: "So the assessment might be used as a whipping board for a staff person who actually would be covering more essential issues for their subject matter" (Mr. San Juan, Interview

[33] It should be noted that a number of teachers considered this a positive use of assessment data. That is, competent teachers did not want ineffective teachers to continue as their colleagues.

Transcript, Paragraph 95). One teacher was quick to point out if the assessments were utilized for evaluation, the dynamics of use would change, "If I had an inkling that they [administration] were going to use this for my evaluation, I would cheat. I can't deny that" (Mr. Seine, Interview Transcript, Paragraph 78).

Building principals felt their own range of anxiety associated with the assessments. Principals mentioned the possibility of being relocated. During the data collection of this study, a district elementary school that had failed to meet Adequate Yearly Progress for the third straight year was restructured. Restructuring must entail the movement of staff members. This is most frequently associated with the removal of the building principal, as was the case in this Reo elementary school. Mr. Caulfield, Principal of Salinger Middle School, expressed his concern about being relocated if his building fails to meet AYP:

> Eventually we will start feeling pressure because of the AYP issue as we start getting to be warned, because what is going to happen is, no matter how hard the schools in Reo probably try, we are every year going to fail the AYP! (Interview Transcript, Paragraph 312)

While accountability for the elementary administrators seems to be embedded in the assessment results, the assistant principal at River High described the accountability system in the district as being deeply steeped in issues of management and damage control:

> Do you know what the accountability is? 'Mrs. Platte, I just got a call from Matt's mother. And she said that Mrs. Campbell is not passing Matt because he didn't turn in his paper on global warming, and that when he came into your office today, you made a face at him. Now what is going on? You need to get this taken care of. Call me back when you get it straightened up.' But you don't get that same kind of phone call saying, 'Five of your English teachers didn't give the Quarterly Assessment and the scores are horrible.' Never. Never. (Interview Transcript, Paragraph 179-181)

One high school teacher offered the following on the difficulties of building leaders using data:

> Schools do not operate on a long-term basis. The school operates on the simple principle, 'is there a kid in the hallway? Do I have to take care of a discipline issue? Now, do I have a parent waiting for me in the office?' Everything is pure reaction or five minutes ahead of time. Anything that requires data means actually spending a time period to examine data to affect a decision more than a few months in the future. There is no administrator that I have met at the building level who is willing to do anything like that. (Mr. Seine, Interview Transcript, Paragraph 211)

Mrs. Hudson, high school English teacher added:

> The rhetoric is that it is a very instructionally based district. Is the leadership instructionally-based? Because on one hand, from an outsider like me looking at all of the tests, I say, 'Like, wow, they got their finger on the pulse; they are really looking at instruction.' Or I could look at it from an insider point of view and say, 'This is just a bunch of rhetoric; this is just a bunch of tests. They are putting lip service to it, but what actually happens is just the management of the school.' (Mrs. Hudson, Interview Transcripts, Paragraph 166)

However, administrators squarely placed the blame for not infusing achievement data into teacher planning and use on the teachers' union. The River High School Principal stated:

> I have never ever seen a union issue that involves students and instruction. Never. Give me one union issue that is based on students' needs and instructional needs. It is always based around the teacher. I am more concerned about the students. (Mr. Nile, Interview Transcript, Paragraph 115)

Again, like the sentiments expressed by elementary teachers, high school teachers were concerned about the impact assessments were having on students. Mrs. Charles, high school counselor, believes the constant testing is having effects on student efficacy:

> I think we test way too much. I think the kids get burned out. And a lot of the times we are not communicating to the kids why are they doing this. Like with the Gates the kids think every time we give it, 'We're the dumb kids.' And it is hard to explain to them that is not the case; we are just trying to help [them]. But they take it three times [in the 9th grade]. Well, I see where they would think, 'Why am I sitting in this classroom with the same people, and we are all taking

> this test three times!' I don't think it is going to help them improve their scores, and it is damaging their self-esteem, what little they have already. I don't want to sound like a counselor and focus just on their self-esteem. But it does affect their performance. How they feel about themselves and how they feel about their potential affects their performance. (Interview Transcript, Paragraph 95)

Mr. Seine, an original member of the secondary science writing team, best summarized the discontent with the assessment plan at the secondary level:

> Unfortunately, despite the vision that a teacher or, even say, many of our administrators may have had for the assessments, which was simply to guide instruction, is gone, those of us who had a vision for it and thus really dedicated huge amounts of time for this; we are the ones that are very disappointed and disillusioned. (Interview Transcript, Paragraph 8, 288)

Seine believed that the district did not conduct a needs assessment prior to understanding what types of assessments were needed and how they can be best utilized:

> Any time you read the literature on the school improvement team, you were supposed to first know what is going on, and then being able to pick the top five goals based on what is not matching your objectives. And then, based on that, you figure out what kind of assessment will let you know whether or not you accomplished that goal etc., etc. Like North Central[34] does. But what happened here is that we inherited committees; we inherited goals from our previous school years, and we're just keeping with those same goals, no matter what the data says. (Mr. Seine, Interview Transcript, Paragraph, 123)

Like a number of teachers, Seine, who was on-board with the assessment accountability plan at its inception, has become so distraught that he did not submit his classes' Quarterly data to the district last year.

Conclusion

The cases presented capture insights of the impediments and the promises of utilizing evidence of student learning for instructional improvement. A number of barriers to fully implementing achievement data as a catalyst for decision-making in the

[34] North Central Accreditation (NCA) is a popular school accreditation system utilized by schools in Michigan.

district emerged. The themes are similar between the elementary and secondary schools: lack of time, questions of assessment metrics, constancy of change, intrusions of pedagogical practice, de-valuing the profession, concerns of student efficacy, anxiety, and stress. Unlike the elementary schools, high schools did not report any assessments that were both valued by teachers. The impact of the assessment plan on schools was one of compliance. Teachers were paying more attention to the standards-based curriculum, but the changes in instructional improvement were not realized. The pedagogical shift that did occur was based on a more didactic practice focused on teaching test skills and covering the breadth of curricular content rather than student learning.

The movement of raw student scores to disaggregated information alone did not lead to knowledgeable use by district actors. The individual (curricular and instructional intrusions and anxiety and stress) and organizational (professional development, evaluation, instructional leadership, and assessment metrics) issues addressed in the cases provide a glimpse of the numerous dilemmas faced when implementing an achievement driven, decision-making model in a school district. Using data as a trigger mechanism to change practice was problematic at the operational levels of the organization. All of the participants in the study were asked to rate their ability to utilize student achievement data in their practice. On a rating scale of one to four (four being advanced understanding and use), teachers and administrators averaged a rating of 2.5[35]. This self-report demonstrated the need for additional support and training to best utilize evidence of student learning.

[35] Based on the administrators (n=13) and teachers (n=14) who participated in the study.

The theory of action that was created to inform and alter pedagogical practices; however, absent were opportunities for learning about the utility of assessments. As the pressure to utilize achievement data increased, so did adherence to the tested curriculum, at the expense of alternative forms of pedagogy and time taken from instruction. As the screws tightened around the curricular controls, pedagogical resistance increased. Undoubtedly, the investment of time to prepare, administer, and score the multiple assessments in the district did not produce an equivalent quid pro quo for teachers. The reported positive elements of assessments (e.g. a curricular roadmap and the promising utility of the Gates and DIBELS) were overshadowed by the inhibiting factors to use. DeBray, Parson, and Woodworth (2001) posit, "The mechanisms of [a] theory of action vary and look different once they are *in use* at the school level" (p. 187). This was certainly the case in the Reo School District.

CHAPTER SIX: DISCUSSION

In God we trust. All others bring data.
~ Brad Duggan, CEO National Center for Educational Accountability

Introduction

The previous two chapters provided a glimpse into the Reo School District as an assessment plan was created and implemented. Specifically, the district's reaction to an underperforming system and the demands of the new federal and state assessment policies were outlined through anecdotes and reports of the central office personnel of the district. Subsequently, an account of how assessments were perceived and utilized at the school level was presented. The purpose of this chapter is to address the research questions posed at the beginning of the study. To begin, how the district is utilizing student achievement is explored. Second, the tensions and impediments to the use of data are uncovered. Next, the conceptual framework is revisited in order to understand the predictive nature of the alternative frameworks. Finally, the original theory of action is used to answer the question, "did it work?"

The Use of Student Learning Data

The assessment plan of action that emerged from the district's assessment accountability plan was anchored in the premise that evidence of student achievement would inform and modify educational practice. The goal was to improve student achievement via mechanisms of information. In turn, information was the catalyst on a number of fronts including: monitoring teacher adherence to the curriculum; making budgetary decisions; decision-making for program and professional development; competitive spurs to improve performance; placing students in special programs; advancing professional communication about curriculum and pedagogy; opportunities to

re-teach content; and propelling teachers' pedagogical practice. The cases illustrate that only a limited number of theses goals were actualized in the district: data for student placement, data to monitor curricular pacing, and a limited set of assessments that are utilized for pedagogical improvement.

Student Data for Student Placement

Student learning data was most frequently utilized to place students in existing or new educational programs in the district. The district's utilization of a norm-referenced assessment, the Iowa Test of Basic Skills, placed students who scored in the 29th percentile on the ITBS reading assessment with the schools reading teacher. The test determined student placement in a remedial high school algebra class and a middle school honors program. Because the test was administered in the spring of each year, building administrators and teachers were not provided with useful feedback for instructional purposes. Rather, a list of students that qualified for support or special programs were sent to the school counselor (for scheduling) and reading teacher (for the student client list for the following year). Nonetheless, teachers throughout the district spent a great deal of time preparing students for the assessment. Mrs. Templeton claimed she spent a few hours a day the week preceding the ITBS helping students prepare. In addition, the test itself took two hours to administer.

The Gates-MacGinite assessment, at the secondary level, was also utilized to determine student placement. Students took the Gates at the end of 8th grade to determine placement in a high school remedial reading course. These students were subsequently tested with the Gates three additional times during the ninth grade to measure growth[36].

[36] All of the students were taken out of class and administered the together.

While the feedback for this assessment was provided to the reading teachers at the high school, reading teachers reported not looking at the assessment results.

Finally, MEAP results also initiated a number of special programs at the building level. Wood Street Elementary created an after-school MEAP Camp, Paradise Elementary launched an after-school science programs, and Packer Elementary instituted a lunch literacy programs as well as Math and Writing Clubs. These programs targeted students achieving below the state expectations. However, the principals reported targeting students that scored just under the passing percentile rather than all students not meeting the standards. Principals thought trying to propel a small number of students into a higher achievement level on the MEAP would raise the overall grade level score. These special programs were created weeks prior to the MEAP test. Mr. Caulfield, Principal at Salinger Elementary School, echoed this practice:

> You have to focus on the small population of kids that are just below the proficient mark. You can siphon off 10-20% of the kids that will pass regardless of the teaching and understand that a number of kids just will not pass the test. To be realistic, you have to focus on the small population that may pass the test. (Interview Transcript, Paragraph 79)

Similarly, Mrs. Moriarty, Principal at Paradise Elementary, described the creation of her school's after school science program:

> I figured out just how many students I needed to pass the MEAP to meet AYP. Let's say I have fifty 5th grade students and this is the percentage that needs to be proficient. So I need thirty-three of those students to pass. I then sat down with the teachers and I said, 'list all of the names of the kids that you have no qualms about passing the test.' I counted them up and there were not thirty-three. So we focused on seven kids that teachers' thought could pass with additional help. (Interview Transcript, Paragraph 35)

In both cases, the principals claimed that much of the teaching and learning focused on the preparation of how to take the MEAP. The content of new learning was not the focal point, the test itself was.

Undoubtedly, this type of information is necessary for student selection into special programs that may help them. However, Mrs. Pilgrim reported that special programs often equated to losing additional instructional time that the students and teachers needed with one another. Mrs. Charles, high school counselor, previously described the constant testing and placement of underachieving students as "damaging their self-esteem" (Interview Transcript, Paragraph 95). The practice of focusing and tracking the borderline students further marginalized the students who "had no chance of passing" (Interview Transcript, Mr. Caulfield, Paragraph 79).

Student Data for Curricular Adherence and Pacing

The district utilized assessments as a major component to monitor the curriculum. In fact, adherence to the curriculum was the main objective of the Quarterly Assessments. Of all the assessments, the Quarterlies had the most impact on teachers' practice to adhere to the curriculum. Prior to the creation of the Quarterly Assessments, the district's Pacing Guides established a set curriculum for each core subject area at each grade level. Teachers utilized the guides in their lesson planning and the scope and sequence of their class. The Quarterly tests were designed based on the Pacing Guides. Because the Quarterlies tested specific content that teachers were to cover each semester, teachers paid specific attention to the content. This was especially true at the end of each nine weeks when the Quarterly tests were sent to the teachers. Teachers reported making

curricular and pedagogical adjustments in the weeks preceding the assessments. How this assessment and others changed their practice is the focus of the next section.

Student Data to Change Pedagogy

The district's assessment plan espoused the belief that student-learning data informs and changes pedagogical practice. The cases illustrate that the assessments did in fact change practice. However, the meaning, effect, and utility on the effect on practice are open for interpretation.

More than the other assessments, the MEAP test triggered a significant test preparation approach for teachers and administrators. Decisions about pedagogy were based on teaching students how to take the MEAP test. In both mini-cases, Mrs. Templeton and Mrs. Pilgrim described the classroom instructional time in the weeks preceding the MEAP as preparation mode. Teachers who administered the MEAP diverted their practice in an attempt to prepare students for the test. Templeton claimed that two weeks of class time were interrupted during MEAP testing each year. Both Pilgrim and Templeton utilized district artifacts (sample tests and the "Buckle Down with MEAP" program) to prepare students for the test. During this time, it was difficult to introduce new content into the curriculum. And, teachers found it difficult to teach in different modalities. Because of the specificity of the MEAP, teachers focused on the aspects of the test design. For example, because of the difficulties associated with the MEAP constructed response items on the 4th and 5th grade MEAP, Mrs. Templeton practiced the elements of what the MEAP was looking for. Mrs. Templeton taught a process of writing that was antithetical to her prior understanding of elements of good

writing. As a result, Templeton modified her pedagogical practice away from what she believed to be useful toward a direct assault on the test.

While teachers were busy teaching test taking skills and sample assessment queries, administrators were establishing an appropriate testing environment. Mrs. Moriarty, Principal at Paradise Elementary School, articulated the demands to motivate students and teachers to do well on the assessment, "I felt you needed to put a lot of 'rah rah' behind the test. At this year's principal's meeting we were asked [by central office] 'what special events are you doing for the MEAP' (Interview Transcript, Paragraph 121). Consequently, in the weeks preceding the test, "MEAP Week" became a slogan that signaled a departure from regular teaching practices in all of the Reo schools.

If the MEAP triggered a "teach how to take a test" mode for teachers, the district's Quarterly Assessments propelled a "teach to the test" pedagogical strategy. Teachers in Reo commonly stated that upon review of the assessments, they adjusted their curriculum and practice. Specifically, teachers found themselves teaching in a transmission, didactic modality as a result of assessment content that needed review or that they failed to cover during the previous nine weeks. Teachers took time away from their instruction to review for this test. And, in some cases teachers admitted to providing students with direct hints in regard with what would be on the test. Mrs. Oakley provided an account of reading the test to students just prior to its administration, "Now what did I not cover that is on the test" (Interview Transcript, Paragraph 71)?

The Reo School District was not void of promising leads that guided instructional improvement. The 4th and 5th grade Gates-MaGinitie and the K-3 DIBELS assessments affected teaching strategies in a positive manner. Elementary teachers reported the utility

of these assessments. Three factors of the Gates- MacGinitie created meaning for teachers. First, because the tests were administered within the classroom, teachers received specific and timely feedback in regard to students reading skills. Second, the teachers worked with the building reading teacher to devise specific pedagogical strategies to help students in response to the feedback. Finally, the design of the assessments could be defined as value-added. Each student was evaluated to establish a baseline and subsequent tests tracked growth in regard to specific skill development. Consequently, adjustments were made for individual students in the form of re-teaching that included being taught in a different fashion. Because the reading teachers at the elementary schools were able to access the classroom and establish pull out programs, the time component of re-teaching which proved problematic for most classroom teachers, was managed effectively. Upper elementary teachers were able to work closely with the reading teacher to share strategies.

In the lower elementary grades, the DIBELS assessment proved of great value for the thirteen Reading First elementary schools. Like the Gates, K through 3 teachers administered the assessment and was provided a support specialist to assist them. In addition, the assessment was tied into a specific, on-going professional development component. Having a support specialist who worked with students and teachers in and out of class proved useful to teachers. And, the additional teachers allowed for more differentiated instruction and the ability to re-teach. The ability to share, learn, and implement new strategies allowed teachers to find a relative advantage in utilizing student assessment data. At Wood Street Elementary School, the reading teacher and coach regularly meet with grade level teachers. These meetings were described as highly

effective. The remaining elementary schools were unable to unlock the potential of the DIBELS assessments because they lacked the resources of professional development and a support person that the Reading First schools were provided.

An understanding of this promising lead can be viewed from two vantage points. First, an environment formed that merged the district's plan for assessment use with the teachers' interests. This environment built and sustained capacity on the premise of many of the normative principles previously outlined. Another vantage point may be the traditional willingness of elementary teachers to adhere to policy signals. That is, previous research indicated that elementary settings are more rational, centralized bureaucratic systems than secondary settings (Firestone & Herriott, 1982). Consequently, elementary settings may be more amenable to the rationalization of assessment accountability. In either case, lessons can be constructed from understanding the ability of assessment information to foster knowledge. This case illustrated a willingness by teachers to invest in a vision of assessment accountability when the dynamics of resources, training, and support were present, and when teachers found value in the tests in relation to their practice. The Gates and DIBELS provide promising leads to reconceptualizing teaching and learning. Teachers were given enough autonomy and support to utilize the assessments in their practice.

The promising use of select assessments to inform and guide instructional improvement highlights a number of support mechanisms that must accompany assessments including: professional development opportunities (on-going and specific to curricular standards), assessment feedback (immediate and tied to curricular standards), and communicative features (on-site literacy coaches to discuss assessment and create

instructional interventions in and out of class). These elements provided teachers with a relative advantage in utilizing the assessments. In other words, these assessments paid attention to both the will (motivation or advantage) and capacity (knowledge) of the teachers. As a result, their practice was directed toward improving their instruction.

Impediments to Student Data Utilization

The anchors of the resistance to the accountability measures were rooted in three features of the organization and the individual: organizational coherence, the utility of information, and the power of professional autonomy. This section utilized these anchors to explain the obstacles to use of student learning data to guide decision-making and pedagogical improvement.

Organizational Coherence and Capacity

Enacting policy is difficult enough when set on a steady platform. However, when policy is layered upon an incoherent organization numerous obstacles emerge. The Reo School District has been in constant organizational flux. Attempts to bring organizational coherence to Reo have been set within a steady stream of curricular, professional development, and structural changes. While Reo could not be categorized as a "House Out of Order," the unique dynamics that underperforming urban school systems, like Reo, make it difficult to build organizational stability and coherence. Complicating the matter further is the continually changing assessment scope, sequence, and standards in the state of Michigan. The organizational features that have led to dysfunction in Reo include the lack of organizational accountability, the reliance on external funding, and the lack of instructional leadership.

Lack of Organizational Accountability

The pressure to utilize achievement data in the Reo School District was felt throughout the school system. However, while the accountability pressures were being felt by the district via the federal and state governing systems, the district has not formalized assessment accountability in either principal or teacher evaluations. That is, the formal evaluation tool or performance conferences did not pressure building educators to utilize achievement data. As a result, the assessment stress in the district monitored curricular practice, but not pedagogical practice.

At the outset of the new accountability plan, the district provided input and support mechanisms to ensure the effective, meaningful use of student learning data outlined in the assessment plan. However, because of organizational restructuring and budgetary implications, the writing teams that invited teacher input were disbanded and the pressures to utilize district artifacts, specifically the Quarterly Discussion Guide, subsided. Consequently, the early discussions that showed signs of changing practice and student outcomes quickly vanished.

Dr. Pinkerton hired a new assistant superintendent for curriculum after the development and introduction of the district's assessment plan. Upon her arrival, Dr. Whitehurst replaced the district writing teams with district steering committees. These committees were comprised of a building administrator and a building representative (elementary) or department chair (secondary) in each subject area. This change marked a major shift in the purpose and utility of the Pacing Guides and Quarterly Assessments. While the writing teams assumed responsibility for gathering input and adjusting the assessments, the steering committees became overarching assessment committees. That

is, the steering committee meetings often revolved around all of the districts assessments. In addition, the professional development activities of building principals in regard to student learning data dramatically decreased. As previously mentioned, the district data consultant articulated the pervasive fear and lack of understanding of assessment data as a result of high principal turnover and the need for on-going trainings.

With each organizational restructuring effort[37], a new set of complications emerged. This year a new reading assessment and development program for all elementary grades was implemented. The DRA (Developmental Reading Assessment) replaced the DIBELS and Gates-MacGinitie in the elementary levels. However, the Reading First schools had to continue with the DIBELS until the completion of the grant. Where K-3 elementary teachers in non-Reading First schools were put in a disadvantaged position of utilizing the DIBELS without support mechanisms, the Reading First schools had to continue with a reading program and assessment that would be replaced. Utilizing both reading programs and assessments was a challenge for the thirteen Reading First schools.

Even more crippling was the effect of the rift between offices within central office. Specifically, different student learning data was collected, disaggregated, and utilized by different departments. The Office of Research, Evaluation, and Program Assessment (REPA) did not have anything to do with the Quarterly Assessments. Mrs. Redding, Director of Assessments, did not feel the Quarterlies were a valid instrument. As a result, this created a number of issues that will be addressed in a later section (access to data and the inability to create a single site for data warehousing).

---

[37] There have been two major central office-restructuring efforts in the past four years at in the Reo district.

As the organizational changes continued to increase, the accountability of data had simultaneously decreased. While the Reo Teachers Union at first embraced the assessment plan, new attempts to remove the Quarterlies have mounted[38]. Clearly, central office believes the union is a major barrier in the utility of assessments in the district. Mrs. Whitehurst and Mr. Nile both attributed the lack of support from the teachers union as an inhibiting factor. Dr. Pinkerton also noted the problem of having a strong principals union when she instituted a weekend student showcase program: "We have to give administrators a day off because I have them come in on a Sunday. So they get off the day before Memorial Day. Who am I having showcase for? Selling their schools so they can get higher enrollment, and showing them off" (Interview Transcript, Paragraph 142).

As a result, the levels of the organization became de-coupled from one another in regard to accountability. The pressures to comply with assessment trends have taken a different toll at different levels of the organization. The lack of evaluation of the assessment accountability plan further de-coupled the sub-units of the organization (Meyer & Rowan, 1991). To further confuse the issue, the district's reliance on external funding sources has created additional dilemmas and diversions.

Reliance on External Funding Mechanisms

Because of the budgetary constraints[39] placed on the district, there was a constant search for new revenue streams. When new programming was sought, new revenue streams were sought after. The most common form of additional revenue for districts is

---

[38] In September 2004, the teachers union distributed a survey to teachers about the beliefs and use of the Quarterly Assessments in the district in an attempt to highlight the dissatisfaction with the assessments.

[39] Recall the impact of the district's fiscal problems on the increase of split grade classes at the elementary level.

the application for grant monies. Both federal and private grants usually stipulate that new programs have learning and evaluative components. In the Reo School District, there was a constant flow of grant implementation. These grants have provided schools the ability to construct new learning opportunities for teachers and students. However, the cyclical completion of grants and implementation of new grants has created a new set of complications.

Issues of sustainability and new learning have emerged as a result of the constant search for new funding. The professional development, special programs, textbook adoption, and assessments that frequently accompany grants proved difficult to sustain upon the completion of the grant. Mrs. Rose remarked that a number of textbooks in use throughout the district, often associated with grants, did not coincide with the Pacing Guides. Yet, when the grant is complete, teachers are left with the remnants. Mrs. Oakley provided an account of how grants have created late summer curricular and textbook changes. Oakley described the common practice of receiving curricular materials just weeks before classes began, "I feel nothing but anger when I walk in- in the fall, and I have something brand new like a math series, and I am supposed to start using it, teaching it, and be familiar with it all at the beginning of the school year" (Interview Transcript, Paragraph 117).

Special programs were constantly being introduced while new ones were being dismantled. In recent years, the district has seen high school block scheduling, remedial high school reading classes, and middle school teaming come and go. While not all the programs were fully funded by grants, monies were either shuffled to fund programs or the professional development efforts were funded through grants. This was the case

during the last two years at River High School where the CSR grant provided two years of professional development for teaching in the block schedule. In 2004, the block schedule was eliminated at River High School due to budget cuts. Dr. Whitehurst cited the need to better understand reliance on grants for the district:

> And one of the problems we've had is having grants that are requiring assessments so they can receive baseline information. And to me this decision becomes, 'What is most important?' Grants are great; they provide us with money and help us to do certain things in professional development. That's fine. But our first priority should be student achievement. And helping teachers having more of this as much you can. A teacher-friendly environment, so they don't feel so overwhelmed and intimidated and controlling those kinds of things. (Interview Transcript, Paragraph 59)

The funding mechanisms of schools, especially urban schools, place pressure to seek outside assistance; yet, assistance came with more demands for tests. This tension was having an impact on the organization's coherence.

Lack of Instructional Leadership

Organizational coherence and capacity also entailed the need for instructional leadership. The building principals have been cited by the district leadership as the key component to teacher leaning and accountability in regard to effective use of assessments. Elmore (2000) stated, "Instructional leadership is the equivalent of the Holy Grail in educational administration" (p. 7). However, in many cases the building principals lacked the capacity and authority to promote the use of student learning data. As previously mentioned, the amount of training and press for principals to utilize data had subsided. There was an assumption by district leaders that principals had all of the necessary tools to lead with assessments. However, principals and the district data consultant frequently mentioned the need for additional professional development by building principals.

Instructional leadership was the rhetoric in the district, yet managerial leadership was the norm. Building leaders lacked both the use of accountability pressures as well as the understanding of how to utilize assessments in meaningful ways. Principals indicated that the managerial aspects of their work defined the "real" accountability in the district. Mrs. Platte, Assistant Principal at River High School, previously stated that the "real" accountability for building administrators involved appeasing parents and taking care of managerial issues. In addition, principals' local constituents, teachers, pressed for managerial leadership. As Mr. Seine indicated, "The school operates on a simple principle, 'is there a kid in the hall'" (Interview Transcript, Paragraph 211). The confluence of temporal pressure to be managers and the lack of skill to be instructional leaders stymied the principals' ability to utilize student achievement data.

The district's pressures to improve tests scores were squarely laid on the principals' shoulders. Mrs. Rose stated:

> In one district meeting principals were shown a chart of school averages on the MEAP. Dr. Pinkerton told them, 'if you're below the line, you must be above it next year.' However, the line is an average that means for every school that rises above it, another will fall. Principals know this. (Interview Transcript, Paragraph 189)

Not surprisingly, principals reported angst over the inability to meet Adequate Yearly Progress goals that could ultimately lead to their removal. This fear had led to leadership that was outcome oriented. As previously mentioned principals sought test specific interventions to mediate test scores. Rather than work with teachers on assessment data throughout the year, the results drive mandates drove principals to be less inclusive and creative and more behaviorally oriented (Rowan, 1990; Spillane, 2000).

Mrs. Lysander indicated the expectation that principals should know how to help teachers; however, this was not evident from the principals in the study. Lysander's assertion that principals were under-skilled and terrified of data became evident in this work. As a result, principals were unable to help teachers interpret tests scores let alone help with pedagogical interventions. Only in a few instances did principals work directly with teachers to understand how assessment data could change student outcomes through improved teaching practices. At Fairview Elementary, Mrs. Moriarty devoted time at staff meetings to analysis of assessment results and creating action plans of improvement. However, even Moriarty, like the other building administrators, had difficulty finding time to both meet with individual teachers and to incorporate assessment conversations in the evaluation of teachers.

An interesting consequence of the focus on managerial leadership was the common delegation of student assessment responsibilities by building principals. In the elementary schools, principals remained informed and involved with MEAP data, but the teacher building representatives were responsible for the Quarterly Assessment data and reading specialists were relied on for the elementary reading assessments. At the high school level the practice was even more problematic. Each secondary building had an assistant principal that was responsible for assessment data. Mrs. Rose, Assistant Director of Curriculum, stated that secondary building principals were supplied with the building comparison charts for the MEAP, but not the more detailed disaggregated data that is needed if they were to be informed about student learning data. In addition, Quarterly Assessment data were disseminated directly to department chairs. Mrs. Platte

who was the assistant in charge of data at River High School described her role with assessment as a coordinator. Platte stated:

> I understand why they want to publicize that we have a lot of data streams. Because it makes political sense. Yeah, we are a struggling district. We're really looking at student learning. Why haven't they taken that second step? Sitting down and saying, 'okay, we have a subgroup problem here between black students and white students in math in the 10th grade and especially in these three classrooms. What are we going to do about it?' (Interview Transcript, Paragraph 56)

Principals reported a lack of effort on their part to confront issues of assessments and teaching. Rather, principals focused on the managerial aspects of their work. They had to rely on encouragement and simple platitudes or simply disseminating the data, "Basically... I give it [the data] back to the teachers and ask them to tell me what they think it means" (Mr. Caulfield, Interview Transcript, Paragraph 35). Their lack of action may have been rooted in their lack of capacity (knowledge) and will (motivation and sentiment). However, the principals may have also become ineffective by the fears associated with the sanctions of failure. Many principals, including Mr. Caulfield who described a fear of being removed from his building, understood the reality of continually failing to meet the Adequate Yearly Progress standards and the interventions that would ensue. In the end, the confluence of poor training, lack of authority to evaluate teachers, lack of supervision from the district office, the delegation of assessment use, and the fear of failing caused administrative practice to become de-couple from the espoused utilization of assessment data.

******

The intent of Reo's assessment system was to tightly couple data and practice. However, the findings indicated that de-coupling occurred, both vertically and

horizontally (Meyer & Rowan, 1991). Vertically, the district began to cluster their understanding of data at three levels: central office, building principals, and the teaching corps. Data-driven professional development activities were exclusively targeted at each level of the organization. Perceptions of accountability were different at the various levels of the organization, and each level of the organization felt different forms of pressure to utilize achievement data. Brown and Duguid (2000) recognized that what looks simple and clear at the top is usually "opaque and confusing on the ground" (p. 100). This lack of vertical clarity led to finger pointing between the vertical levels of the district.

The ambiguity also pierced the organization horizontally. The professional development activities, district writing teams, venues of communication, and instructional pressures from building leaders were not sustained over time. Teachers were encouraged to re-teach content that the assessments identified as problematic. However, teachers reported pressures to move onto the next elements of the Pacing Guides so they would not fall behind. Additionally, there were no mechanisms in place to help teachers build pedagogical skills to re-teach even if time was not an impediment. There were no strategies to help teachers understand how to engage in re-teaching which entails differentiating instruction and hopefully teaching in a different manner. In the end, adding assessment policy onto an incoherent organization did not modify the system of the actors as intended (Elmore, 2003). The problems associated with the information itself added further confusion.

The Utility of Student Learning Data

The term "Assessment" became a vulgar term to building educators in Reo. While the assessment plan was designed to add knowledge, it was the assessment plan itself that was halting reform efforts. A number of issues about the student learning information became evident in this work. This section explores how the assessment plan became problematic to effective, meaningful use.

Too Much Information

The district's testing timeline was viewed as a complimentary set of assessments according to district officials: There was an assessments that was norm-referenced against national results (ITBS); there was a state level assessment to compare students across the state to the criteria of the state benchmarks (MEAP); there were a set of elementary reading assessments uncover specific skill development (DIBELS and Gates MacGinitie); and a specific district assessment was in place to hold teachers accountable to the district's Pacing Guides (Quarterly Assessments). In addition, the district had recently added the PROM/SE assessment and continued to utilize the Houghton-Mifflin theme tests in grades K-3. According to Mrs. Rose, Assistant Director of Curriculum, "All of these tests are needed because they provide different student information" (Interview Transcript, Paragraph 123).

For building educators, the set of assessments was viewed as too much, too varied, and too intrusive. In fact, the building educators saw the cross purposes of the assessments as competing with one another. Because of the impact of the MEAP on students and educators, many believed the district was exasperating the problem by adding layers of assessments. This was especially true for teachers who administered the

MEAP. One teacher stated, "If the MEAP assesses the state benchmarks and the Quarterlies assess the Pacing Guides, which are based from the benchmarks, why do we have to give both tests" (Mrs. Templeton, Interview Transcript, Paragraph 23). Central office administrators are quick to point out that Quarterlies were suspended in MEAP tested grades during the MEAP testing. Nonetheless, teachers think the number of tests had overlapping purposes and the assessment scope and sequence should be reconsidered.

The intent of the multiple streams of data was not only to inform teachers of the achievement level of students, but also to cause modifications in their practice. Specifically, teachers were encouraged to re-test subject matter. Teachers found this to be impossible in their practice. Teachers indicated that the performance feedback, while relatively timely for the Quarterlies—within two weeks, were delivered into the curricular coverage of the following nine week period. Because teachers were required to cover the content for the following nine weeks, they felt pressure to move on, in spite of the assessment results. Recall Mrs. Tiber's comment from the previous chapter, "There is no time to re-teach it. We are lucky to get the teach part in" (Interview Transcript, Paragraph 106). Mr. San Juan, high school teacher, went on to offer an interesting tension. That is, if re-teaching were to take place, wouldn't it be appropriate to re-assess the students?

The cases of Mrs. Templeton and Mrs. Pilgrim highlighted the real time dilemmas placed upon teachers in the district. Mrs. Pilgrim reported that twenty-seven hours of instruction were devoted to district mandated assessment preparation and administration. This accounts for 14.9% of time she has with students. And, when this was combined with three hours of classroom assessments per semester, the percentage of time spent on

assessment jumps to 17.6%. Mrs. Templeton reported that 15 days or 100 hours of her instructional time was devoted to district mandated assessment administration and preparation. When these numbers are subtracted from the total days of instruction and then compared to the respective Pacing Guides, there is little question that time is a valuable commodity.

Data Accessibility and Metrics

Two additional issues emerged in regard to student-learning data: the accessibility to data and the metrics of data. Access to user-friendly data was problematic for building administrators. Mr. Trout, Principal at Kilgore High School, articulated the multiple offices he had to call to generate just the type of student report he needed. Additionally, Mrs. Moriarty at Paradise Elementary indicated that the reports provided to her were difficult to understand. Moriarty stated, "We get caught in the minutia of the data itself rather than what the data is trying to tell us" (Interview Transcript, Paragraph 198). As a result, building educators had to interpret the data they were provided.

A number of district educators indicated their concern that the multitude of assessments only provided a single type of data, student learning. Proponents of using data in schools have called for an eclectic array of student data information including demographics, perceptional, student interest and motivational and school process data (Bernhardt, 1998; Popham, 2001). Mrs. Lysander, district data consultant, was working with the district to create a single record data warehouse. Lysander indicated that without one place where data is collected, it becomes difficult to analyze different types of data (interoperability). Because Lysander believed that the district should be looking at more than student achievement data, the creation of an accessible warehouse of data is critical.

Currently, the district disaggregated MEAP data by sub-groups of the population as per state law. However, Mrs. Rose indicated that there was not the ability to track Quarterly Assessments by sub-group or even longitudinally to student chart growth over the years. Mr. Trout stated that data has to be reported differently for it to be useful, "These teachers don't give a rip what a deviation from the norm is... Most teachers can't pick up a report and read the stupid thing" (Interview Transcript, Paragraph 258, 316). As a result, the savvy educator who wanted to look at data in a number of dimensions must be able to collect and analyze data themselves.

Finally, a number of concerns over the validity of the assessment have been catalogued. Specifically, teachers and central office administrators alike questioned the merits of the Quarterly Assessments. Whether this was a tactic to discredit the tests or a legitimate problem will have to be left to psychometrians.

******

Building educators viewed the collection of student learning data as an interruption to their practice. The multiple streams of data inundated and intruded on their work. In addition, concerns about the interpretation of data posed problems for the utility of student-learning data. And, when these concerns were compounded with the lack of capacity (knowledge) and the absence of conversations among teachers and between teachers and instructional leaders, it should be no surprise that potentially useful data was relegated to the outer core of teaching and learning.

The Privacy and Power of Teachers

Cohen and Spillane (1993) purported that the "appealing vision of new order... contains a devastating critique of existing realities" (p. 35). Traditional reforms that tried

to wedge into the historically personal profession of teaching have been rebuked. Similarly, teachers in Reo questioned the philosophy of a policy that had repercussions on their pedagogy. For teachers, adherence to the assessment plan created curricular, pedagogical, and ideological tensions. The lack of involvement in what constituted the district's curriculum disenfranchised teachers. Even more potent was the perceived intrusion on the privacy of their pedagogy. Teachers reported a change of pedagogical practice that they believed to be antithetical to their current perception of effective, normative practice and formal learning (understanding of the research literature and pre-service teaching education). The inability of the assessment policy to strike curricular and instructional coherence distanced their practice from the signaled goals of the organization. As a result, teacher ideology impacted (often implicitly) the district's assessment plan (Rowan & Miskel, 1999). Weiss (1995) stated, "However weakly integrated they [ideological beliefs] may be, they provide an emotionally charged orientation that provides a basis for taking a position" (p. 575).

The deeply personal aspects of one's interests and ideologies have clashed with the Reo assessment accountability demands. Teachers had deeply held beliefs about the autonomy of their craft. The assessments were incompatible with their commitment to their craft (Elmore, 1983). While teachers were involved in the assessment plan early on, the disbanding of the writing teams and subsequent inability to adjust the content and sequence of the Quarterly Assessments and the scope and content of the Pacing Guides disenfranchised teachers. Recall Mr. Seine who was involvement in the implementation of the assessments and now refuses to even administer the Quarterlies because of the absence of the writing teams. This may have led to the unwillingness of many teachers to

give up classroom level assessments in lieu of the Quarterly Assessments. Teachers in this study indicated that information gathered at the classroom level by teachers (anecdotal and classroom assessments) most informed their practice,

For many teachers, entrance into the teaching profession allowed for a certain amount of autonomy when it came to what is to be taught and how it would be taught. That is, while standards and curricular frameworks have always framed the teachers' subject matter, the profession has allowed individual teachers to interpret *what* was important to teach. For teachers at the operational level, there were strong indications that the relative advantage of curricular "fit" and pedagogical "fit" with assessments were major factors. That is, personal curricular interest (*what* should be taught) and their belief about pedagogy (*how* content should be taught) drove the teachers' understanding of and the ability to utilize student-learning data. Teachers' personal interests, experiences, and formal training mitigated the curricular decisions from classroom to classroom. When the Reo District enacted the multi-phased assessment accountability plan, teachers reacted through these filters of interpretation.

Mrs. Rhine, high school teacher, revealed an interesting tension when she indicated that research and pre-service teaching hinged on pedagogical practices that were hands-on or constructivist in nature while teaching practices were moving the content coverage by means of didactic teaching modalities. Every teacher in this study indicated that their pedagogy changed to more of a didactic model in lessons that preceded the Quarterly Assessments and to some extent the MEAP. As a result, many lessons which teachers believed to be "tried and true" or of personal interest had to be abandoned to cover the required material. Mr. Burch at Wood Creek elementary stated

that he felt de-valued as a professional because of the new pressures to teach in this manner, "Nowadays I feel that I am an instructor not a teacher. I just do what they tell me to do" (Mr. Burch, Interview Transcript, Paragraph 25).

The perceived infringement on teachers' curricular and pedagogical beliefs also led to increased levels of anxiety and stress. Teachers perceived that assessment results would be used to evaluate their instructional competence. Consequently, teachers had to choose between their strong beliefs in pedagogy and the fear of failing to adhere to the district's assessment plan. For administrators, anxiety stemmed from the fear of failing to meet Adequate Yearly Progress.

Teachers filtered the new policy through a deep-seated organizational memory (where student assessments were not utilized) that had developed over time. Teachers developed an expectation that new reform efforts in the district would be a common occurrence. Mrs. Pinecrest stated, "You know how they [new programs] come and go... [So you] stand back and wait" (Interview Transcript, Paragraph 69). The steady diet of grants and new district initiatives created disharmony at central office and at the classroom level. Teachers have often rejected reform, not on its merit, but on the memory of the constancy of reform efforts. Shulman (1983) stated "We act not to achieve our ideals but to fend off our nightmares" (p. 484). Teachers did not recognize a relative advantage for assessments in their practice. When policy took dead aim at changing practice, without sufficient involvement or resources, teachers often rejected the reform. Peter Marris (1974) stated, "If an innovation is imposed upon them, without the chance to assimilate it to their experience, to argue it out, to adapt it to their own interpretation... they will do best to fend it off" (p. 157).

The Reo assessment plan, based on competition and the mandates, combined with the looming accountability sanctions of NCLB to displace staff was creating a sense of stress and anxiety for the educators in the Reo School District. In isolation from each other, the concepts of assessment data and accountability had been non-threatening in the Reo School District. The forces of accountability that were exerting anxiety and stress on teachers and students promulgated teachers to question their purpose of teaching (e.g. teaching to the test) that had moral implications (e.g. cheating on the tests). And, the perceived belief that assessments would be used to evaluate their effectiveness was a dangerous mix for teachers. Dr. Pinkerton stipulated that together, assessment and accountability could indeed stir different emotions when she stated, "In defense of the teachers, the way that assessment has been tied into accountability at the national level, it is scary. Because they are trying to tie assessment into accountability as opposed to assessment as a learning tool" (Interview Transcript, Paragraph 139). Pinkerton believed that the district's assessment plan was less about measuring performance and more about helping kids learn. As the school district tried to facilitate the use of student assessment data as a learning tool, the fears associated with accountability continued to be a distraction. This distraction may have lead to the paradox of accountability pressures becoming a factor that inhibits use rather than one that motivates use. O'Day (2002) warned that such a focus on high-stakes pressures that rely on negative incentives "distracts from risk-taking and innovation and toward organizational survival" (p. 315).

The moral imperative of teachers (Sergiovanni, 2000), rooted in their strong interests and ideology, usurped the systemic pressures to utilize assessments. Not surprisingly, attempts to alter *what* and *how* teachers taught were met with resistance. As

the saliency of the signals grew, so did the anxiety, denial, rationalization, and the issue of professional identity. In the end, teachers began to de-couple the assessments from their practice as a result of the lack of knowledge needed to utilize the data, the lack of time permitted to collegially discuss the results, the timeliness of the data results, the inundation of multiple data streams, and the perceived incompatibility of data with their instruction. As a result, the normative effect of data use was not realized; rather, educators developed a strong logic of confidence that halted not only the efforts of the theory of action (Meyer & Rowan, 1991) but also the amount of their risk taken in pedagogy (O'Day, 2002). Habermas (1984) stated, "Norms do not claim validity because they are connected with sanctions; if they did, they could not *obligate* the actors to obey only force them into submission (p. 45, italics in original). In the midst of the constancy of reform, educators had nostalgia for the solace and efficacy in the privacy of teaching (Hargreaves et al., 2004; Lortie, 1975; Rosenholtz, 1989).

The Predictive Power of the Alternative Framework

Up to now, the district's actions and impediments to action have been summarized. However, to fully understand the dynamics of the findings the influences of the theoretical assumptions must be investigated. What was the impact of the alternative framework on the actions of the Reo School District? Specifically, what was the predictive nature of the institutional, organizational, and informational lenses outlined in the conceptual framework? Each aspect is reviewed based on the district actions.

Reo's Institutional Isomorphism

The Reo School District implemented an internal accountability system that not only complied with the state's Education Yes!, but also surpassed its requirements.

While most districts weigh only the state assessment system (MEAP), Reo had developed additional layers of assessments. These layers of assessment were not all intentional. That is, by creating new educational programs in the district based on external funding (e.g. grants) a number of assessments were added to the assessment accountability plan.

The Reo School District faced a number of pressures to modify its current practices. Like all public schools, Reo was mandated to fulfill the sanction-ladened elements of the NCLB legislation. Explicit in the federal mandate was the utilization of assessment data as the most important indicator of student success. Moreover, Reo, as a struggling urban district was challenged with becoming a legitimate institution. As a result, a quest for legitimization was anchored in current political pressures. However, while the political moment created coercive pressures (DiMaggio & Powell, 1991b) on the district, there was a clear indication that Dr. Pinkerton anchored reforms in a mimetic fashion. Because Pinkerton was dismayed with the uncertainty she was exposed to in the district, she sought harmony by retrofitting a perceived successful system of assessment accountability from her previous district onto the Reo School District. Consequently, the policy created to utilize student achievement data was interpreted by central office through their strong beliefs that data could be utilized to motivate decision-making (budget and professional development), monitor the curriculum (what is being taught), track student knowledge, identify re-teaching opportunities, inform parents, and establish new communication patterns about pedagogical practices (how the curriculum is taught).

The policy design began with clarity, authority and consistency; however, this type of policy was difficult to implement because the mechanisms to support instructional policy were embedded in people's will and capacity. Consequently, Meyer and Rowan's

(1991) predictions of institutional isomorphism held true. The phenomenon of the de-coupling of the subunits from each other (different beliefs at central office and the building) and from the activity (teachers and building administrators became de-coupled from the use of assessment data) took place. In addition, the rituals of past practice (strong belief in pedagogical practice) and the logic of faith that this reform would also pass (organizational memory) rang true. Finally, the avoidance of inspection and evaluation permeated all levels of the organization. Without the evaluative function the use of student learning data became a ritual, not a tool for real change.

As a result, the institution's goal or orientated focus became a prohibiting factor in the learning process for actors in the organization. The more "highly institutionalized the environment, the more time and energy managerial elites devote to managing their organization's public image and status and the less they devote to coordination and to managing particular boundary-spanning relationships" (Meyer & Rowan, 1991, p. 61). In the case of Reo, once policy was implemented, the strong institutional environment combined with the weak technical environment, distanced the theory of action from information for decision-making knowledge. Adding even more complexity was the instability of the institutional environment itself. Specifically, the state's constant changes in the MEAP scope and sequence placed a burden on the local school districts. While the purpose and goals to utilize student achievement data were clear to the administrative elites in the district, ambiguity among the street-level actors emerged, making the implementation of the theory of action ambiguous and, thus, problematic. In the midst of the complexity, the Reo central office staff could not understand why

teachers "love the Pacing Guides, [but] hate the assessments" (Interview Transcripts, Dr. Pinkerton, Paragraph 124).

The theory of action became one of legitimization, not of real change (Meyer & Rowan, 1991). The assessment plan became a public display of improving student achievement, rather than a process oriented plan of action. The Reo theory of action co-opted a novel ideal of education- to use evidence of student learning to guide instructional practice- into another command driven output oriented reform (Rowan, 1990). These institutional effects crippled the use of assessments by de-coupling administration from policy, teachers from administration, and teachers from pedagogy. This is what DiMaggio and Powell (1991) called the "Iron Cage of Institutional Isomorphism."

Reo's Coherence and Capacity

The institutional setting that evoked information as the catalyst for improvement in the Reo School District became increasingly inconsistent with organizational learning. That is, this district piloted an exploitive search process for solutions (March, 1999a). Specifically, actors in the district were pressed to utilize student assessment data without clear pedagogical goals and on-going, specific support mechanisms. Instead, educators had to interpret and create their own artifacts of learning. For the Reo principals this meant creating public learning artifacts (e.g. MEAP Camps, Science Clubs, etc.). However, the artifacts did not constitute new learning for students or teachers. Consequently, leadership became command and behaviorally-oriented (Rowan, 1990; Spillane, 2002). The classical, highly bureaucratic model of district leadership remained

rooted in place. As the reform stakes increased, so did the command level of the district administrators.

The clear focus on goal achievement confounded adaptation of the theory of action as a result of the need for an immediate solution (March, 1997; Rowan & Miskel, 1999). This search process led to the creation of new solutions in the neighborhood of old ones (March, 1997). That is, the search process to seek utility of student learning information to inform practice was marginalized as a result of teachers enacting data into their current practice. The press to find solutions led to cases of cheating and teaching to the test. This type of response was the most frequent response to the assessment streams by district teachers. Without taking into account the sensemaking process of individuals through their prior knowledge and social context (Spillane et al., 2002; Weick, 1995) the organization itself blocked new learning opportunities. The singular focus on the goals of student achievement and the simultaneous looseness found in the process handcuffed Reo's theory of action. As a result, the focus on compliance blocked the organization's capacity to learn. Feldman and March (1988) quipped that their ability to determine how organizations will utilize data leads to the conclusion that "organizations are simply stupid" (p. 414).

Individuals had strong belief systems (Weiss, 1995) that led to conservatism (Buchmann, 1996) of practice and the resistance to the bureaucratic controls (Goldring, 1995). While the policy signals stipulated a tightly coupled compliance at central office, the capacity at other levels of the organization remained uncertain. Consequently, compliance did not stipulate capacity (DeBray et al., 2001). Goertz (2001) posits, "Sufficient capacity is critical to making the 'horse trade' work. Without it, performance

based accountability is reduced to a 'horse whip'" (p. 58). As a result, the stakes of assessment accountability did not modify the organization (Elmore, 2003). Rather, the actors' interpretation "shape[d] the environment more than the environment shape[d] the interpretation" (Daft & Weick, 1984, p. 287). Without coherence and capable organization features, the stakes of reform mattered little. In the end, Reo was data rich, but organizational impoverished.

Reo's Informational Isomorphism

Information in Reo became a tool of surveillance to monitor teaching, but not to alter pedagogy. The technology of data itself did not have the ability to modify practice without its information being turned into knowledge for practice. The knowledge management framework, rooted in a process oriented and knowledge rich environment (Choo, 2001; Petrides & Guiney, 2002; Petrides & Nodine, 2003; Thorn, 2001), was not practice in Reo. Because the student assessment data was collected, analyzed, and disseminated at the bureaucratic level of the organization, information was not flowing in the social context of the organization (Brown & Duguid, 2000; March & Sevon, 1988; Weiss, 1995). Scott (1992) stated, "Among the various flow connecting elements, the flow of information is the most critical" (p. 88). Once the communication about student assessment data stopped so did the chances of real pedagogical change. This led to the assessments being utilized in a less than meaningful and effective symbolic fashion where data stood at the ready, but was not utilized for pedagogical improvement (Feldman & March, 1988).

As a result, educators marginalized student assessment data. Decision-making and information became de-coupled from one another. The glut of information and the

accessibility issues (types of information and timeliness of feedback) created a scenario where decisions were made ex post facto of information. For example, achievement charts were created and distributed, but not analyzed in regard to decision-making. Actors also discredited the relevance of information (questioning the metrics of the assessments) and were constantly asking for more information (Feldman & March, 1988). The decision-making trends were determining the information rather than information and decision-making working in concert. When data was disseminated, actors engaged in a search mode to seek solutions. This has been characterized by a number of learning artifacts building principals (e.g. MEAP Camp, Science Club, etc.) and teachers (e.g. teach to the test mode) created when they received data results. Decisions made at a district wide school improvement meeting provided a useful example of the district's behavior in regard to information. Mrs. Redding announced new target goals for the 2004-05 MEAP before the 2003-04 MEAP was even administered. Teachers and administrators were asked to develop plans aimed at new targets when results from recent assessments were not yet analyzed:

> How can new targets be made without a thoughtful critique of the previous year's scores? There is an inordinate amount of time talking about next year's goals and less time taking apart the current year's goals and activities. As a result, the activities and goals of the district are disengaged. Are the goals merely for Central Office, merely a bureaucratic process? (Field Notes, March 17, 2004)

The student achievement data was transformed through the existing structures of the district and the personal ideologies of actors throughout the organization. March and Levitt (1999) stated that, "Rules, procedures, technologies, beliefs, and cultures are conserved through systems of socialization and control" (p. 83). Not surprisingly, the organization and its actors had a profound isomorphic effect on student achievement data.

The influence of information that was not contextualized through the teachers' practice, rather information was modified to fit their current practice. Brown and Duguid (2000) posited that "knowledge lies less in it database than in its people" (p. 121). Because data was not embedded in practice in Reo, information itself further fragmented an unstable environment. What began as an attempt for student leaning data to shape the educational environment turned into a case of the environment shaping the information (Daft & Weick, 1984). Only when data was collected locally, supported with professional development materials, aligned with the curriculum, and collectively analyzed at the school did certain data streams (e.g. Gates and DIBELS) have a change to guide instructional improvement. Otherwise, information morphed into the similar patterns of use as predicted by the organizational information theories.

But Did the Theory of Action Work?

While the discussion thus far has evolved around the predictive powers of the alternative aspects of the conceptual framework, the theory of action must be revisited in regard to understanding if the reforms had an impact on student learning outcomes. The theory of assessment accountability in the Reo district was straightforward: practice can be mediated through the three prongs of setting standards, teaching to the standards, and monitoring the teaching via assessment mechanisms. The theory of action assumed that teachers' beliefs would follow the mandate to take action.

While the evidence that student-learning data guided instructional improvement remains inconclusive, there are indications that the theory of action did in fact achieve its goals. To begin, the theory of action augmented the effort of teachers to monitor the curriculum. The district Pacing Guides and Quarterly Assessments monitored the

adherence to the curriculum every nine weeks even without a formal evaluation tool. Even those who did not utilize the Quarterly data monitored the Pacing Guides. The impact may not have been uniform across the district, and many unintended consequences may have arisen (e.g. teaching to the test and didactic teaching modalities), nonetheless a primary goal of the theory was achieved.

While the actions of the district actors surfaced crucial findings, the student learning outputs were equally as important. Consequently, district demographic and student achievement data provided a set of outcome data in an attempt to track changes in each since the arrival of Dr. Pinkerton in 2000, the subsequent implementation of Pacing Guides in 2000, the implementation of the Quarterly Assessments in 2001, and the enactment of NCLB in 2001. Table #10 indicates that since the inception of Reo's Assessment plan and the arrival of Dr. Pinkerton there have been significant improvements in the district drop out rates and graduation rates (state averages in parentheses). In addition, the high School MEAP (HST) (see Table #11) shows marked improvements in reading, science, writing, and social studies (state averages in parentheses). The increases in scores have outpaced the increase in the state averages in these subjects. Finally, table #12 indicates gains in the elementary MEAP scores with the exception of 5th grade science and 7th grade writing (state averages in parentheses). The most significant gains, when compared to the state average gains, were made in 4th grade reading and writing as well as 7th grade reading. Other scores reflect the flux in the state averages.

Table 10: Reo and State Student Drop Out, Graduation, and Enrollment, 2000-04

	2000-01	2001-02	2002-03	2003-04
Drop Out	12% (5.7%)	14% (5.5%)	8% (3.7%)	9% (4.1%)
Graduation	61% (79.6%)	57% (80.9%)	71% (86.3%)	71% (84.8%)
Enrollment (Thousands)	18.1	17.6	17.5	17.2

- Enrollment in charter schools decreased by over 150 students from 1994-95 to 2003-04

Table 11: Reo and State HST Scores, 2000-04

	2000	2001	2002	2003	2004
Math	49% (65%)	52% (68%)	45% (67%)	47% (60%)	48% (59%)
Reading	59% (69%)	60% (74%)	52% (71%)	57% (67%)	74% (76%)
Science	39% (56%)	33% (60%)	39% (59%)	37% (61%)	55% (64%)
Writing	38% (58%)	33% (69%)	41% (68%)	43% (61%)	59% (58%)
Social Studies	17% (24%)	12% (27%)	18% (24%)	13% (26%)	30% (35%)

Table 12: Reo and State Elementary MEAP Scores, 2001-04

	2001-02	2002-03	2003-04
Read/ELA (4)		50% (75%)	75% (79%)
Math (4)	51% (65%)	54% (65%)	59% (73%)
Science (5)	63% (73%)	69% (77%)	65% (78%)
Writing (4)		41% (47%)	45% (48%)
Soc. Stds. (5)		16% (28%)	20% (31%)
Read/ELA (7)		36% (61%)	45% (61%)
Math (8)	30% (53%)	24% (52%)	42% (63%)
Science (8)	48% (67%)	44% (65%)	51% (66%)
Writing (7)		42% (56%)	26% (47%)
Soc. Stds. (8)		16% (33%)	16% (29%)

* Blanks indicate when a new test was administered.

It is problematic to correlate the improved scores above to the district's assessment plan, NCLB, or the arrival of Dr. Pinkerton. Looking at MEAP scores in this manner is suspect when considering the test changes from year to year and the lack of more than three years of achievement data. Without test equalizing procedures, comparisons can be utilized for simple interpretations only. Nonetheless, there is a clear

indication that the district is making marked improvements in a number of important measures.

Conclusion

Previously, achievement results were kept in-house and used externally by or local realtors who sold real estate based on the location associated with "good" schools. Today, student assessment data are reported in public venues (e.g. the local newspapers) and have critical monetary and personnel implications. As a result, the stakes of assessment accountability have risen. This chapter summarized the findings of how the Reo School District reacted to the pressure to institute new accountability measures. The Reo School District responded with a comprehensive assessment plan of action. As a result, the important first step of moving raw data to usable, disaggregated information had been reached. The district made gains in the development of much needed formative assessment tools in the district, but failed to root the assessments in student and teacher learning (Black & Wiliam, 1998). The collision of the theory of action and the human, institutional, and organizational forces in the district halted the efforts to utilize evidence of student learning to guide instructional improvement.

This is a case of central office policy creation based on institutional pressures. That is, district elites initiated policy mechanisms as a response to state and federal pressures. Reo *made* policy as a reaction *to* policy (Spillane, 2004). The policy that was implemented consisted of the creation of state benchmarked curricular guides and accompanying assessments to monitor the use of the guides. The district elites' interest to retain students, mandate to improve achievement scores, and belief in assessment use as a valuable pedagogical tool led to the creation of this assessment plan.

The institutional, organizational, and informational tensions created a case of maladaptive implementation. The competing theories of institutional press (mimetic isomorphism) and institutional effect (de-coupled units and actions) indicated that decision-making with information in Reo was fragmented and weak. The information introduced in the district was also filtered through the interests and ideologies of educators in the district. People made sense of data through their own knowledge, interests, ideology, and the social context of the organization (Spillane et al., 2002; Weiss, 1995). Spillane (2004) purported that change in practice is proportional to the amount of instructional and epistemological change that is demanded. However, the theory of action in the Reo School district was based on a different factor, motivation. The theory of action was not accompanied by technical mechanisms of sustained support and accountability to target instructional and epistemological changes. As a result, the pressure to utilize achievement data had an effect on *what* teachers taught (curriculum), but not on *how* they taught (pedagogy). For central office, the quest for institutional and constituent legitimacy propelled a quick bureaucratic response and may have simultaneously subverted implementation.

As pressure mounted to utilize assessment data to improve student achievement increased, so did the response by the Reo central office staff. The district's theory of action was an attempt to tightly couple information and pedagogical practice (activities). However, this rational view created strife between sub-units in the organization (central office, building principals, and teachers) that ultimately lead to de-coupling the activity (utilization of evidence of student learning to guide instructional improvement) and the organizational goal (improved student achievement). Each level of the organization had

its own understanding about why student assessment data were not impacting classroom practice. The possible effectiveness of evidence of student learning information to impact teacher pedagogy and student learning had not been fully realized in the Reo District.

The *active* use of achievement data expressed by the district's theory of action was countered by the finding that data are utilized in a *passive*, sporadic, and episodic manner. The pressures to utilize student achievement data initially shaped the district 's response; however in the end the potency of the environment shaped the data in a number of ways. The assessment plan of action created structures that "help[ed] define human activity, providing the rules and resources on which it is based; but structure is also created, reproduced, and potentially transformed by the same activity" (Spillane, 2004, p. 177). The ability of the activities created an informational isomorphism that shaped student achievement data in its current form. Specifically, the current achievement data were utilized to monitor curriculum, not student achievement; to seek new data, not utilize existing sources; to place students, not to teach them differently; to improve test scores, not student learning; to advocate for conversations, not to create meaningful dialogue; and to create additive measures, not embed new practices into the current structure. While there were some indications of data utilization informing decision-making and pedagogical practices, most practitioners deemed assessment data as a leadership fad or a distraction to their current professional work. That is, teachers marginalized or enacted (through their own interpretation) the reform efforts as a result of the policy ambiguity and the continuous bombardment of new grants, programs, and reorganizations of the district leadership. If student-learning data were to strike at the

core technology of schooling, teaching and learning, these inhibiting factors would have to be addressed. Otherwise, data, like many reforms that came before will be marginalized to the periphery of schools (D. K. Cohen, 1988).

The most consequential findings were the dramatic improvement in key district data (MEAP, HST, drop out, graduation), the inordinate amount of time the assessment administration and preparation takes away from the instructional day, and the impediments of the alternative frameworks of the institution, organization, and of information itself. While these findings shed light on the impact of the assessment accountability plan in the district, they also surface a number of new tensions, questions, and implications.

By virtue of the student outcome gains, an interesting paradox emerged: How can the outcome results be explained based on the findings of the institutional, organizational, and informational impediments previously reported? And, did the outcomes indicate the relevance and effectiveness of the rational theory of action? Finally, an obvious normative question emerged: Has the assessment plan propelled a sound educational practice? These questions along with the implications will be addressed in the final chapter.

CHAPTER SEVEN: CONCLUSION AND IMPLICATIONS

If schools are to cultivate sophisticated and independent instruction, they must be sophisticated and independent. ~ Ted Sizer

Man is born free and everywhere he is in chains. ~ Jean Jacques Rousseau

Introduction

Data generated from student assessments have become the catalyst and means for student improvement in education today. However, while information from data is essential to school improvement efforts, it has also become problematic (O'Day, 2002). This study addressed key questions about the viability and utility of assessment data to guide instructional improvement. It was inspired by the recent mandate-based, sanction-ladened, and hortatory assessment policy of school reform that aims to reconceptualize important aspects of decision-making and pedagogy in schools. The finding revealed not only the organizational and professional difficulties of implementing such reforms in schools, but also the consequences of a top-down, untested policy itself. More specifically, the work unpacked a district's response to federal and state policies as well as the effects of implementation by the actors within the organization. In the end, this study underscores another story of policy reform that is unfamiliar, but not unknown. That is to say, while assessment accountability is beginning to frame curricular changes, the meaningful and effective use of achievement data in pedagogical practice remains absent (D. K. Cohen & Hill, 2001; Sipple et al., 2004).

The purpose of this chapter is twofold. To begin, three tensions are contemplated: the rationale of the district's assessment plan, interpreting the contentious results of the plan, and the apparent gains and losses of the plan. Based on the knowledge acquired in

this work, the researcher offers informed speculation for each. Second, implications for policy, practice, and research are offered.

The Assessment Accountability Plan Redux

School districts, and struggling urban schools in particular, are constantly seeking legitimacy. Organizations gain legitimacy through the eyes of community constituents (e.g. state policy makers and parents). In turn, organizations risk losing legitimacy when constituents loose faith in them (March & Olsen, 1976; Meyer & Rowan, 1991). Consequently, when external pressure increases, organizations create new artifacts that signal compliance with the constituents' concerns (Cuban, 1990; Meyer & Rowan, 1991). Often, the quest for legitimacy entails a search process that culminates in artifacts that are fashioned after similar legitimized organizations[40] (DiMaggio & Powell, 1991b; Meyer & Rowan, 1991).

This was certainly the case of the Reo School District. When pressures increased from a number of angles (meeting the demands of NCLB, lackluster achievement performance of state measures, and poor community perceptions) the district engaged in a process to win back legitimacy. The mandates to make improvements in the district were clear; however, how it was to be achieved remained ambiguous. The Reo plan for improvement was left in the hands of the new superintendent. Dr. Pinkerton ultimately made policy based on her own interpretation of how to regain legitimacy for the school district (Baier, March, & Saetren, 1988).

[40] This is especially true in times of high saliency (Popkewitz, 1991) where the public perceives changes must take place in schools for social, economic, or credentializing purposes (Labaree, 1997).

Dr. Pinkerton began by introducing a plan of action that replicated an assessment plan from her previous district. Pinkerton was able to gain the support of her central office staff and the district's teacher union. For central office staff, the new wave of accountability, and their perception of new authority was an easy sell. Additionally, the district's teacher union felt pressure to support a new superintendent based on the their knowledge of the public perception of the district as well as the promise by Pinkerton not to utilize the assessments as an evaluation tool.

What resulted was a comprehensive assessment scope and sequence with mechanisms of support (e.g. professional development, discussion guides). While the creation of a rational plan of action took root with the district elites, the projected participation of teachers (collegial conversations, communication with parents, re-teaching, pedagogical changes) and building principals (instructional leadership) went unrealized. The complexity of organizational life and the nuances of individuals that comprise it did not mesh well with the plan. As March and Simon (1958) purport, that "[A] rationale theory of action calls for simplified models that capture the main features of a problem without capturing all its complexities" (p. 169).

The quest to install a plan of action to make improvements in the district was not accompanied by many of the features of assessment utilization offered in the literature. Specifically, a number of models that have emerged call for the development of a culture of data use, uses of multiple data typologies, detailed disaggregated achievement data that highlights patterns, and professional conversations and training centered on explaining and strategizing about the data (Choo, 2001; Keeney, 1998; NCREL, 2004; O'Day, 2002; Petrides & Guiney, 2002; Petrides & Nodine, 2003; Popham, 2001; Thorn, 2001).

Instead, the district plan relied on compliance through mechanisms of competition and the expectation that data reporting would generate discourse at the building level. In addition, rather than creating new learning opportunities and time for teachers and principals to work with the data, the district created artifacts of learning to drive improvement in test scores. In turn, building leaders created artifacts that reflected adherence to the utilization of data and improving student outcomes (e.g. posting competitive assessment result charts, creating after school and pull out programs for underperforming students). Finally, teachers signaled compliance through elements of pedagogical change that were unintended by the district's plan (e.g. teaching to the test and didactic modes of pedagogy).

The assessment accountability plan was a top-down, command oriented reform that failed to gain commitment from internal constituents. Conversely, the plan generated legitimacy from the external community that was solidified as reports of student achievement, drop out, and graduation rates improved. Elucidating the improvements in the midst of disillusioned and controversial actions of the district's actors is taken up next.

Elucidating the Improvements

The effects of the assessment plan on the outcomes of schooling and on the teachers in the district were not neutral or benign. As the plan unfolded there were marked improvements in test scores (MEAP/HST) and drop out and graduation rates. Surprisingly, teachers were synchronizing these publicly celebrated improvements with intensifying their efforts to modify the assessment plan[41]. A major question in this work

[41] Recall the teachers' union efforts to discredit the Quarterly Assessments.

involved the confluence of intended student outcome results with the unintended results of modifying practice. If impediments of the institution (e.g. de-coupling effects in the schools), organization (e.g. incoherent, command style reform), and information (e.g. metrics and access) became crippling to implementation, what accounted for the significant improvements in the state tests, drop out rates, and graduation rates after the implementation of the assessment plan[42]?

Answers to student outcome questions are most often found in the characteristics of the teacher and the learner. Because this work did not address the latter, looking into the black box of teaching provided an opportunity to answer the question. The rational explanation is rooted in the functionality of the assessment plan itself. The plan created a simplistic model: create standards, tell teachers to teach to the standards, and assess the students against the standard to make sure teachers are complying and students are growing. Teachers, while reluctant, use the Pacing Guides to plan their instruction. As a result, one account of the improvements in state testing can be attributed to the increased adherence to the curriculum.

Other scenarios can be utilized to explain the results as well. To begin, the reported unintended consequences of pedagogical changes may have impacted student learning. Specifically, teaching to the test, cheating, and teaching students how to take tests. The numerous accounts of reviewing material just prior to assessments (especially the Quarterly Assessments and the MEAP) may have had an impact on assessment results. Other factors that may have mitigated the results include the number of new programs instituted in the district in the past five years (e.g. magnet schools, Reading

[42] As indicated in the previous chapter, this claim of MEAP improvement in the district is a simple interpretation of data reported from 2000-04. Without additional longitudinal data and test equalizing factors, this data should be judged accordingly.

First Grant, Comprehensive School Reform Grant). New artifacts of learning that were specific to the tests were also instituted including: remedial math and reading programs, special MEAP programs outside of class time, and the integration of test preparation material and time into the teachers instruction. As reported, these programs taught directly to the state assessments with a specific focus on how to take the MEAP test.

The review of extant research also provides possible explanations. Studies indicate practice has engaged in teaching to the test, cheating, and a "skill and drill" modality (Klein et al., 2000; L. McNeil & Valenzuela, 2001). Additionally, Haney (2000) purports many schools are engaging in gaming techniques where groups of students are excluded for taking tests. While the changes in teacher practices were identified in the Reo's school, there was no evidence of gaming in the district. Less factual claims can also be made. The increase in scores may be a result of the initial improvement associated with many new reform efforts (Ogawa & Collom, 2000; Tyack & Cuban, 1995). Or, the fact that there are new governmental and public pressures to improve our urban schools may have created a rupture in time (e.g. Industrial Revolution, Sputnik launch of 1957, the Nation at Risk Report, 1983) that is often associated with achievement gains (Popkewitz, 1991).

Finally, because teachers reported changing their pedagogy and curricular attention and practice as a result of the assessment accountability plan, it raises the question of whether their previous forms of instruction were ineffective? Mrs. Rhine, high school teacher, provided an example of how the practice of teaching one novel with different mediums over a long period of time was restricted when Pacing Guides and Quarterly Assessments were instituted. Perhaps, teachers like Rhine, previous to the

plan, were spending an inordinate amount of time on topics based on their own interests, not curricular benchmarks. Buchmann (1986) purported that the informal, eclectic models of pedagogy learned (via pre-service teaching) and practiced (actions based on interests alone) inhibit movement toward best practices. Historically, schools aimed at the production of students, and were not held accountable[43] to learning targets (Cusick, 1983; Meyer & Rowan, 1991). As a result, the new accountability placed upon the teachers may have proven beneficial for student learning, curricular framework development, and the development of assessments to test student knowledge (Darling-Hammond, 2003). Not surprisingly, "mean" policy, such as NCLB, has been deemed worthy in an age of "nice" educational policies that have proven ineffective to change practice and outcomes (Hess, 2003).

Because of the myriad of possible effects on the outcomes, it would be premature to claim "Mission Accomplished" based on the district's assessment plan. The current press may have "nudged" test scores higher, but pedagogical changes did not occur as intended (Ogawa & Collom, 2000). Unless schools are willing and able to reconceptualize their organizational culture, teaching and learning practices will remain static (Fullan, 2001; Weiss, 1995). In light of this, the next section addresses the issue of what constitutes good teaching and learning.

Learning Gains and Losses

One final normative, question emerged from this work: Did the improvements previously outlined constitute good teaching and learning? The overall impact on

[43] Education has often been compared to the health care industry where outcomes (assessment scores and cure rates respectively) are not publicly reported.

teachers teaching and learners learning are of critical importance for reflection. The noted changes in pedagogy for teachers constituted teaching that was an affront to many educators. Specifically, a claim can be made that the movement toward didactic means of teaching to the tests, cheating, test preparation (how to take the test) and the lack of inquiry and experientially based instruction will have a negative impact on student learning, regardless of assessment results. The press to utilize assessments changes pedagogy constricts instruction (e.g. constructivist, experiential-based learning) and derails the use of authentic assessments. Multiple choice and limited constructed response questions did not persuade teachers to teach in a constructivist manner. Moreover, the ever-increasing breadth of curriculum will continue to promote a didactic type of pedagogy that most students despise, teachers are uncomfortable with, and educational researchers decry. Sykes (1997) states, "Such tasks press toward a narrow, impoverished conception of learning that comes to influence students and teachers alike" (p. 112). This narrow, impoverished learning is exacerbated by an assessment focus that fails to take students' social and cultural capital into account (Popham, 2001).

The contentious ability for assessments to represent student knowledge and the impact assessments were having on the curriculum, pedagogy, and the consumption of school resources raises an important question: Who is paying attention to the assessment results? And, is anyone concerned about the inordinate amount of time assessment administration and preparation is taking away from classroom instructional? If teachers, students, and parent constituents are not concerned with the formal assessment results, do the assessments and the results lose local credibility? And, as a result, should the force of external constituents weigh so heavily on the legitimacy efforts of local schools? What

constitutes good curriculum, good pedagogy, and good learning should be the focal point of educational debate; however, the question remains, can assessment accountability open this topic of discourse?

Undoubtedly, the institutional pressures have advanced the "technical capacity [of schools] to inspect educational outcome" (Rowan & Miskel, 1999, p. 365). However, in the current format and under the current conditions there arises the possibility of data used as a subversive tool to eliminate groups of students who are likely to perform poorly or to pervert the educational process by altering instruction (Ogawa & Collom, 2000). Ultimately, such a single-minded assessment policy may create a troubling paradox: the focus on tests can lead to better test scores while simultaneously lowering educational quality (Duke, Grogan, & Tucker, 2003). Under this guise, many would not want their student enrolled in such a highly assessment-heavy school.

Implications

Understanding the implications of this study for policymakers, practitioners, and researchers is important. Certainly the new educational federalism that has been mounted has no signs of relenting. Consequently, there is a need to explore, explain, and analyze assessment accountability in its current state. The purpose here is not to determine what kinds of policy, research, and practice are needed, but to indicate what insight this study provides for each.

Implications for Policy

Each accountability system has expectations, mechanisms, and incentives. NCLB marks a new era of school reform, an era characterized by the amalgamation of both market and state oriented approaches to accountability (Heinecke, Curry-Corcoran, &

Moon, 2003; Ogawa & Collom, 2000). Assessment has moved from the accounting of student learning to the accountability of schools writ large (Firestone et al., 1998). The focus has clearly changed from input driven policies to an output focus (Rowan & Miskel, 1999). This result driven focus that is attached to clear, even if untested, sanctions marks a new era of educational control and compliance (Elmore, Abelmann, & Fuhrman, 1996; Fuhrman, 1999).

Shulman (1983) states, "Teaching is the very prototype of the idiographic, individual, clinical enterprise. Policy connotes the remote, monolithic, and unresponsive" (p. 488). As a result, external demands on schools, and consequently goal-setting, are not axiomatic to action. Simply building a theory of action does not determine action. While pressures have mandated the development of professional educational practice that is void of interest-based habits (Buchmann, 1986), it has become obvious that an educator's actions and sentiments are strongly correlated (Homans, 1950). As a result, it is difficult to mandate what matters (McLaughlin, 1990). The consequences of such external pressures may include coercing action, forcing uniformity, and creating adversarial relationships (McDonnell & Elmore, 1987).

While the policy landscape continues to change, policy enactment itself continues to traverse a non-linear course from the statehouse to the schoolhouse (Spillane, 2004). The district plays an important role in the enactment and learning of policy as well as in the creation of new policy as was the case in Reo. As a result, policy should not treat organizations as uniform entities, but as the complex organizations that they are. Currently, the focus of assessment policy is on the organization of schools and its effect on student outcome measures, not the actions of teaching and learning for educators and

students. Since actors determine action in organizations, schools would be well served by policy architects who understand the policy ramifications (perceived or real) on teaching practices (Black & Wiliam, 1998).

Policy often stipulates an over-simplified theory of action, and adding sanctions and exhortations to the policy has shown little indication of promise. Top-heavy policies have led to street-level ambiguity. Ambiguous policy has been associated with inhibiting meaning, interpretation, and learning (Baier et al., 1988). As a result, autonomy is also inhibited, and a conservative impulse is enacted by molding new reforms into old structures (Malen, 2003; Marris, 1974). Of greater concern is the ability for policy to create damaging effects on leaders, teachers, and students[44]. If policy alone cannot mandate action, how does policy strike at the core technology of schooling, teaching and learning? And, if ambiguity is destructive to accountability efforts, how can accountability efforts seek clarity? Two possible policy archetypes may provide leads in the quest for policy to impact curriculum, pedagogy, and student learning: backward mapping (Elmore, 1983) and learning policy (Ball & Cohen, 1999; D. K. Cohen & Barnes, 1993; D. K. Cohen & Hill, 2001; D. K. Cohen & Spillane, 1993).

Elmore's (1983) backward mapping design begins with the end in mind. That is, if policy is to reconceptualize teaching and learning, then teaching and learning are where policy architects should begin. Elmore (1983) states, "Define the delivery- level unit with the greatest effect on the problem, describe what needs to happen in that unit to solve the problem, and then describe for each successive level above that unit what needs

[44] Spillane (2000) provided a description of district leaders' react to policy in cognitively behavioralistic way. Cohen's (1990) case study of Mrs. O, provided an account of a teacher who misinterpreted reform efforts. Finally, Valencia, et al. (2004) supplied a study of how assessment-driven reforms in Texas negatively affected students.

to be done to support activity at the delivery level" (p. 367). Backward mapping provides input from those whom policy affects the most, teachers. As a result, the policy should complement, not compete, with teacher practice. This type of policy design can be useful in the quest to utilize student achievement data (via assessments) in guiding instructional improvement. Backward mapping can be metaphorically represented by the construction of a puzzle. While most puzzle aficionados start by finding corner pieces and building-out, starting a puzzle with the middle piece is most assuredly more difficult and time consuming; however, this process frees one from the binds of being constrained by the constricting boarders.

Similarly, learning policy evokes empowerment, not obligation (Shulman, 1983). A learning policy typology aims to merge individual performance and institutional goals. Such policy allows organizations to create and "actively support teachers' learning of matters closely related to instruction" (D. K. Cohen & Hill, 2001, p. 121). Learning policy is rooted in the inquiry of pedagogical practices and student understanding rather than skill and technique development for the implementation. This inquiry-based development is an emotional and social process that constructs learning via engaged and open debate (D. K. Cohen & Hill, 2001). In addition, the inquiry process incorporates idiosyncratic practices of teachers, rather than projecting individual teachers practice as a "contagion" that prohibits systematic implementation. Brown & Duguid (2000) state that learners need "access to communities of learning, interpretation, exploration, and knowledge" (p. 232). If change in instruction is the goal, then "professional learning matters to instruction" (Ball & Cohen, 1999, p. 4).

Currently policy is manifested in short-term, highly visible effects. Implementation failures are countered with additional policy layers or exhortations for the need of a new professionalism in schools. Backward mapping and learning policy slow the goal oriented thrust and incorporate the wholly personal process orientation into the policy-implementation mix. In both cases, policy is accomplished *with* actors rather than *to* actors[45]. In other words, policy can challenge beliefs and actions without dictating them. The need for specific, site-based, and on-going resources, not just technical accommodations, may prove beneficial for policymakers. Adults in schools should be treated as learners, and, because learning is a social and cognitive process, the procedural *how* of policy is more important than the *what* (Elmore & McLaughlin, 1988). The alternative is a further balkanized organization. Complex organizations need such a nuanced approach if policy is to have a chance to strike at the core of education. The hope of assessment accountability to improve teaching and learning lies in hearts and minds of practitioners, not the pen of policymakers.

Policymakers must consider other possible inhibiting factors to assessment accountability. Without support and resources in compliance with policy demands, policy enactment is difficult at best (Barber & Phillips, 2000; O'Day & Smith, 1993), and professional and organizational learning is confounded by anxiety and stress (Rait, 1995). Orfield (2004a) states that "the transformation sought by the law [NCLB] was supposed to be lubricated by a huge infusion of new federal funds that would add resources to the

[45] Two highly celebrated cases of urban school improvement, New York Community District #2 (Elmore & Burney, 1999) and San Diego Schools (Hightower, 2002), utilized elements of policy for process learning which included: a focused vision, the development of professional cultures and learning networks, long-term professional development, targeted deployment of resources, and a constant collaboration of improvement efforts in instruction. In both cases, coercive administrative forces were also exerted on the district actors.

schools required to produce large improvements" (p. 4). The unrealized federal support mechanisms are having a profound effect on implementation. Policies like the assessment accountability provisions in NCLB embody a structure where "the intent is laudable, the prescriptions inadequate, the execution poor, and ultimately the results insufficient to narrow the achievement gap and truly leave no child behind" (Fusarelli, 2004, p. 89).

Finally, curricular and assessment demands, set at the policymaking level, must be understood as to its effect on the utility of assessments. The reports that over 100 hours of elementary instructional time and over 15% of high school instructional time is devoted to assessment administration and preparation is alarming. Additionally, the issue of curricular breadth and depth must be considered. Currently, the schools, like Reo, are constrained by the amount of content that has been mandated. Studies have indicated the effectiveness of reducing the breadth of curriculum in order to promote the depth of content (Ma, 1999). Policymakers must consider the possibility that the problem of assessment accountability may not lie in teachers' interest or practice, but in the curricular demands imposed upon teachers and students. That is to say, if the number of curricular demands decreased, would evidence of student learning guide decision-making and instructional improvement?

After 150 years of fumbling through reforms of accountability and assessments, a new wave of reform has forged assessment and accountability together. Whether the euphoria will soon wear off, only to be replaced with another wave of reform, is unknown. What is known is the fact that in order for policy to take root in teaching and learning, policy must be meaningful to practice with support mechanisms for its

implementers. And, if the current stakes (mandates, public exhortations, monetary sanctions and rewards, and competition) are not having the intended result, we must consider alternatives to the ineffectiveness of modifying behavior with rewards and punishment (Kohn, 1993).

Implications for Practice

Policy continues to be a chronic tonic for the ailments of schooling (Cuban, 1990); however, the implementation of policy in practice is episodic (Elmore & McLaughlin, 1988). This study reveals fundamental questions about the role practitioners play in reform efforts. Issues of will, capacity, ideology and interest, professional identity, and a relative advantage in teaching practice have emerged as key aspects of practice in this era of high-stakes accountability. Feldman and March (1988) state that only when "tactical uses of information are transformed into belief" will there be a useful function for information (p. 421). Consequently, understanding and mediating teachers' beliefs are part and parcel of pedagogical modifications. Rather than try to change the established behaviors of teachers, Elmore (2002) stipulates, "Only a change in practice produces a genuine change in norms and values. Or, to put it more crudely, grab people by their practice and their hearts and minds will follow" (p. 3). Because "our cognitive model is not linear... understanding can follow action" (Spillane et al., 2002, p. 421). Simply put, belief can follow action (McLaughlin, 1990; Spillane et al., 2002). While assessment accountability also stipulates that belief can follow action, the institutional effects have disengaged policy from practice. Learning and activity stipulate changes in practice, not policy alone. As a result, a framework for educators utilizing assessment data is offered.

Creating a culture that enables the use of data must mark the beginning of any attempt to effectively utilize evidence of student learning to guide decision-making and instructional improvement in schools. Developing such a culture stipulates a process orientation rather than a single focus on outcomes (Choo, 2001; O'Day, 2002). March (1999c) states that "process gives meaning to life, and meaning is the core of life" (p. 28). Consequently, the focus must be inverted "from a focus on given standards, embodied structures, and their influence on use- to a focus on human agency and the enactment of emergent structures" (Orlikowski, 2000, p. 421). Nardi and O'Day (1999) purported the use of information as an ecology "allows us to find individual points of leverage, ways into the system, and avenues of intervention" (p. 50). Additionally, the knowledge management frameworks describe a data-rich organization focus on people and process. If information can become embedded in a process of pedagogical learning and improvement real changes have a chance[46].

Fullan (1993) states, "Changing formal structures is not the same as changing norms, habits, skills, and beliefs" (p. 49). The outcomes focus of the institution placed the burden of data on changing behaviors, not the system. This process resulted in a learning myopia of short-term assimilation as a symbolic gesture of compliance (March & Levinthal, 1999) rather than leveraging teachers' "hearts, minds, and practices" (Elmore, 2003, p. 204). Consequently, at the heart of organizational learning is the ability to enter into a professional community (Lave & Wenger, 1991), to develop modes of inquiry (Rait, 1995), and to take risks (O'Day, 2002).

---

[46] See Choo (2001) in Chapter 2 of this study. Choo illustrated how information utilized as a process led to the eradication of smallpox.

Like previous policy failures which are based on organizational incapacity and incoherence (Chubb & Moe, 1990; Elmore, 2003) the assessment accountability movement has been stifled at the district and school level. As a result, building internal capacity is seen by many as the key to coherence, clarity, and enactment of any external demand (Elmore, 2003; O'Day et al., 1995). The failure for practitioners to address the needs of the organization will cause policy, regardless of the stakes, to be ineffective (Elmore, 2003). Consequently, the organization becomes the backbone to develop the will and capacity to utilize evidence of student learning to guide decision-making and instructional improvement.

Because information flows through institutional rites, rituals and structure (Bolman & Deal, 1997), individuals experience most of their salient learning from organizations (Brown, Collins, & Duguid, 1989; Weiss, 1995). Without learning, data may be transformed into information, but not knowledge that will guide instructional improvement. O'Day (2002) states, "At the heart of the learning process in any complex system is the role of information and the movement of information among agents and subunits through patterns of interaction" (p. 298). The more an organization can couple communication and coordination of information, the more egalitarian the organization will become (March & Sevon, 1988). Creating professional development opportunities that are content specific and on-going must be established. In addition, professional learning communities and opportunities must be transparent in organizations. Rait (1995) submits, "Organizations encounter situations that demand learning, which may be an arduous and disturbing process, when they are already weakened by stress and tension" (Rait, 1995, p. 73). Nonetheless, the organization must provide opportunities for

discourse among educators. For raw student achievement data to become tacit pedagogical knowledge, a process of professionalism fraught with conversation and learning must become a cultural norm. Reforms that shake the foundations of schools create dialectical discourse that has been largely absent in schools (Majone, 1989).

Petrides and Nodine (2003) purport, "Knowledge management is more likely to take root in communities that need to share knowledge to realize their goals, that have some information sharing norms in place, and that have a leader who is willing to sponsor the effort" (p. 18). As a result, a data-rich environment must include strong instructional leadership. While policy often initiates a command-style behavioral model (Spillane, 2000), this type of leadership is not compatible with change. Earl and Katz (2002) posit that "data provide a ready-made vehicle for engaging staff in planning. Leaders can broaden the base of inquiry by distributing leadership and developing a cadre of people who are competent and confident using data" (p. 1020). Spillane and Seashore Louis (2002) made a case for stretching explorative practice, "School principals who cannot engage others in leading will be unable to spread and mobilize the experts necessary for school improvement in their schools; they are thus unlikely to be very effective" (p. 98). School leaders, as boundary expanders, may begin to mitigate the partisanship of administration and pedagogy in schools (Goldring, 1995). Spillane (2004) purports that when district policy makers initiated policies that nurtured an understanding of the standards at deeper levels, "They created opportunities for teachers to make sense of the mathematics and science standards on a more substantive level" (p. 138). Consequently, district leaders must develop a vision for professional development where assessments and standards are understood and hence utilized (D. K. Cohen & Spillane, 1993).

District leaders and teachers must create and care for a professional accountability model where data are collected, disseminated and utilized in appropriate ways (O'Day, 2002). The use of value-added data is crucial for educators to find a relative advantage for its use. The school organization must establish a data warehouse that is inclusive of the multiple data streams (student learning, perceptional, process, and demographic data) that provide a means to create specific, user-friendly reporting mechanisms. Thorn (2002) states, "A robust knowledge management system must reflect the information and knowledge management needs of all levels. In particular, data must be gathered at a level of aggregation appropriate to the user with the most fine-grained analytical needs" (p. 7). As a result, the types of data, the accessibility of the data, and the presentation of the data become important organizational mechanisms.

A key to creating a professional model in schools is the balance between professional commitment and bureaucratic controls (O'Day, 2002; Rowan, 1990). This can be achieved through the process of building the will and capacity of educators, stretching school leadership practices, and promoting communication using specific, value-added achievement data. Finally, educators must be provided opportunities to enact and learn pedagogical strategies into their practice based on the learning that is achieved from student learning data. This process may necessitate the development of a new curricular frameworks that seek to highlight the major concepts and themes from the given state standards and benchmarks. Narrowing the breadth of curriculum may allow practitioners to better utilize student-learning data in ways previously restricted (e.g. opportunities to re-teaching).

In the end, the way forward may be "paradoxically to look not ahead, but to look around" (Brown & Duguid, 2000, p. 8). Equilibrium must be found between the formal goal oriented process and informal practice (Brown & Duguid, 1991); between organizational search and organizational slack (March, 1999a); between organizational efficiency and individual adaptation; between change and stability; between organizational and professional accountability (O'Day, 2002); between rational concepts and clear-headed logic and powerful intuition (Mintzberg, 1973); and between diversity and unity (March, 1999a). These are contradictory terms that are often poorly integrated. As March (1999b) made clear, "Balance is a nice word, but a cruel concept" (p. 5). Balancing "requires developing coupling loose enough to allow groups to develop their own knowledge, but tight enough to be able to push the knowledge along the lines of process" (Brown & Duguid, 2000, p. 115). Such balance is crucial to making organizations tight enough to promote clear goals that are accountable to outcomes with the loose features of individual practice (Bossert et al., 1982). When a balance is struck, a window of opportunity or local zone of enactment may be opened (Spillane, 1999). Creating a culture of assessment data and individual knowledge generation hinges on such a balance.

Implications for Research

There is a clear and present need to study the viability and impact of the current high-stakes assessment policies. How policies are mediated and implemented is more complex than a simple theory of action. Research must not only look at the summative outcome of assessment policy, but also investigate the utility of formative assessments and the practices of teachers (Black & Wiliam, 1998). Daft and Weick (1984) posit that

interpretation "shapes the environment more than the environment shapes the interpretation" (p. 287). Street-level actors provide a grain-sized view of implementation. In addition, understanding the district is also crucial because of its role in mediating and generating policy. When the units are combined (students, teachers, parents, district, etc.), a holistic framework will provide insights into the artifacts district leaders create, the impact on teachers, and the fallibilities of the policy itself.

School implementation failures indict the organization and its actors, not the institution itself (Clune, 1987). However, failures can point to systemic dysfunction just as well as the organizational issues. In fact, "institutions often get in the way of the extraordinary work of education" (Sykes & Elmore, 1988, p. 77). Consequently, the design and expectations of policy themselves must be considered.

The rational perspective focuses on the bureaucracy and its ability to effect student achievement outcomes by modifying behaviors. This perspective also assumes local context is static. Alternative analysis in this work uncovered the importance of the individual's belief and action to morph information and the effect of assessment policy. As a result, teaching, learning, and the educational environment must, together, be the focus of investigation rather than the individual units of analysis. In depth studies like this begin to tease out both forces and dynamics. Understanding both the institutional accountability and professional responsibilities are important (Adams & Kirst, 1999; Buchmann, 1986). Professional autonomy is undeniable in schools (Hoy & Miskel, 1991); however modifying the conservatism of local actors (Buchmann, 1996) cannot be the sole function of the bureaucracy.

The multi-focal framework utilized in this research shifts the research lens away from the question: What are the effects of assessment policies on student achievement? Rather, the study should inspire a new set of questions such as:

- Can mandate-based, sanction-ladened, and hortatory policy modify behavior?
- What is the impact of high-stakes policy on professional autonomy (curricular and pedagogical changes)?
- How is organizational capacity and coherence changing, based on high-stakes policies?
- What is the impact of assessments on student efficacy and achievement (especially between population sub-groups)?
- How have the metrics of the assessments themselves impacted achievement outcomes and teacher practice?
- What has been the utility of learning and of instructional artifacts created as a response to assessment policies?
- What is the impact of curricular breadth on the use of assessment data?
- Are the models of assessment use becoming prevalent and working in schools?

Conclusion

The original intent of the district's assessment plan was to track student achievement, change instructional practice, create re-teaching opportunities, identify trends in student achievement, inform parents, and monitor adherence to the curriculum. This study began with two hypotheses to analyze the district assessment plan. The normative hypothesis situated the utility of assessment data in the construction and maintenance of an inclusive, learning-based knowledge environment. This hypothesis held true in regard to the promising leads of assessment uncovered in Reo. Specifically, the Gates and DIBELS assessments were grounded in a collegial and supportive knowledge environment including: temporal proximity to the administration and results of the assessments, curricular and assessment correlation, on-going, content specific professional development, on-site professional assistance, and continuous professional

conversations about the student assessments. When such a knowledge environment was embedded in the classroom, the curriculum, and in the practice of teachers, assessment data guided decision-making and practice.

While promising leads supported the normative hypothesis, the alternative, behavioral hypothesis proved highly predictive. The institutional pressures, organizational incoherence and incapacity, and the strength of professional beliefs had profound effects on student assessment information. The ambiguity of the meaning of accountability and evaluation; the lack of building level commitment and command oriented reform; the lack of structured conversations; and the inability of building leaders to become instructional leaders all led to the predictive power of how information is utilized in organizations. Mr. Burch's analogy of student learning also held true for teachers and principals: "[You can't teach] how to drive a car... [and then] put [them] in a big truck" (Interview Transcript, Paragraph 73). If educators are to utilize a new base of knowledge to guide decision-making and pedagogical practice, new learning must shape a new understanding.

In the end, federal, state, and local search for the optimization of student achievement data was confounded by local sub-optimal technologies (March, 1997). And, when the unintended effects of assessment accountability were uncovered—inordinate amounts of time taken from instruction, changes in pedagogy that were incongruent with beliefs, and perceived effects on students—the assessment reform was rejected. Historically it can be said that schools have generally gotten in the way of reform, at best, and overpowered reform, at worst (Weiss, 1995). While the pressures are changing in schools, the technologies of teaching and learning, pre-service instruction,

and current practices have remained the same. The continual streams of reform push street-level actors to reminisce "a romanticized past of the 'ideal' home, family, and school" (Apple, 2004, p. 15). As a result, accountability pressures are at odds with deep seated, personal practices. Wave after wave of school reforms (Cuban, 1988) have not been able to penetrate the technical core of schools, student learning, and teacher pedagogy (D. K. Cohen, 1988; Elmore, 2000; Sykes, 1999). Instead, reform efforts have been marginalized (D. K. Cohen, 1988; Plank, 1987) as a strategy to relieve external pressures of policy (Meyer & Rowan, 1991). The constancy of reform is perpetuated in "shifting watchwords in education: excellence to equity, efficiency to empathy, unity to pluralism and then back again" (Tyack & Cuban, 1995, p. 44). Consequently, educators have become skeptics of reform efforts, both acute and chronic (Cuban, 1986).

Currently, the ubiquitous assessment data are over-hyped and under-utilized; yet, schools continue to be fed a steady diet of tests. The real power of assessments lies in the transformation of raw data and disseminated information into explicit knowledge to guide instructional improvement. Gathering information is distinct from judging information. Processing and thinking are anchors in judging information: "Knowledge is something we digest rather than merely hold. It entails the knower's understanding and some degree of commitment" (Brown & Duguid, 2000, p. 120). However, the rational theory of action strikes as a non sequitur of knowledge generation. There exists a false sense of rationality that clear goals and a simple plan, even when accompanied with sanctions and incentives, will result in supportive actions and positive outcomes. In the current state, explicit knowledge is being reinforced, not reconceptualized. Consequently, assessment accountability will merely "perpetuate a status quo in need of transformation" (Sykes,

1997, p. 118). Until policy architects are willing to invite teachers to the table, and until teachers are willing to open the door to their own practice, such policy will continue to fall short. It should be no surprise that the efforts to collect, analyze, and utilize data are social, cognitive, and political processes in schools (Weiss, 1995). As a result, information must have both a pragmatic and ideological utility: "information that can mitigate choice has staying power" (Feldman & March, 1988, p. 411).

The road that connects assessments and student learning most assuredly runs through the core technology of schooling, teaching practices and student learning. Assessment accountability must be anchored in meaningful, coherent, and clear demands. Tight mandates must supply ample room for capacity building, which, in most cases, is a slow, arduous, expensive, and ambiguous proposition. This path to effective, meaningful use of student assessment may lead to design characteristics congruent with organizational and information theories and may have less to do with the omnipresent theories of rational accountability. Consequently, the aspects of policy and practice must be simultaneously ripened (D. K. Cohen & Barnes, 1993) if evidence of student learning is to guide decision-making and pedagogical practice.

Schools continue to have difficulty combining information with users (educators) and uses (pedagogy and curriculum). The use of data has been relegated to aggregate data utilized for public consumption and as fodder to debate. Using data for pedagogical decision-making is under utilized. Instead, data are used for student placement, teacher curricular accounting, and school evaluation. Determining how Jane is learning is being subsumed by how Jane can do better on the test. Until new learning in an environment of assessment data takes place, educators will be unable to generate assessment knowledge

from data use as a subversive and symbolic activity to data use that affects teaching and learning. The real triggers of school improvement continue to be good teaching and student motivation, not punitive accountability measures (Darling-Hammond, 2003). As educators once again stand at the cliff's edge, they must not retreat-- as expected-- or jump-- as hoped. Rather, if assessments are to strike at the heart of teacher pedagogy and student learning, where they must be anchored, we must reconceptualize assessments as a bridge over the uncharted chasm to guide decision-making and instructional improvement.

APPENDICES

APPENDIX A

METHODOLOGY

The objects beauty is based on what you see and what you don't see makes it useful.
~ Zen saying

Introduction

Understanding the complexity of policy design and implementation is problematic in a nested organizational setting such as a school district. This work offers an inside-out view of how school districts think about and utilize student achievement data to guide instructional improvements. This study utilized an in-depth approach from multiple perspectives of the school environment to best capture the "story" of schools, assessment accountability, and decision-making. Additionally, the methodological approach allowed the "story" to be analyzed through a multi-focal conceptual framework of institutional, organizational, and professional lenses. Schoenfeld (1999) stated, "Sometimes the only way you can understand complexity is to study complex things" (p. 12). In this vein, this methodological approach supported the intention of understanding and analyzing the implementation of assessment accountability in a complex school environment.

This research uncovered one district's collection and analysis of student assessment data used to guide instructional improvement. The research questions previously posed were purposely open-ended in order to evoke natural responses. The questions sought to inductively and analytically unpack the demands of student achievement data through the eyes of multiple actors in a school district. Consequently, a

research design was selected to answer the research questions in this emergent field of inquiry.

Research Design

Qualitative research is good for in-depth analysis of complex processes (Miles & Huberman, 1994; Stake, 1994). Creswell (1998) stated that qualitative methodology follows in the "traditions of inquiry that explore a social or human problem. The researcher builds a complex, holistic picture, analyzes works, reports a detailed view of informants, and conducts the study in a natural setting" (p. 15). As a result, the research design employed was a qualitative case study (Creswell, 2002; Thomas, 2003; Yin, 1994). This study utilized an embedded single site case study design. Such an approach was needed to develop an overview view of the district while simultaneously understanding the multiple layers of the organization and an understanding of the actions of involved individuals. This analytical approach allowed for the discovery of meaning and theory building through the in-depth investigation, the diverse collection of data, and the continuous analysis of the data.

Single-Site Embedded Case Study

This research employed an embedded single-case design that allowed the researcher to simultaneously understand the features of the organization writ large and the behaviors, attitudes, and perceptions of the individuals in the organization (Yin, 1994). Specifically, the design allowed for exploration at three levels of the organization (central office, building principals, and teachers) and two levels of schooling (elementary and secondary). Additionally, the inductive nature of this work called for a methodology that was designed as a process of discovery, not an outcome. That is to say, "The interest

was in process rather than outcomes, in context rather than a specific variable, in discovery rather than confirmation" (Merriam, 1988, p. 19). Case study research is useful when "how" or "why" questions are asked, when the investigator has little control over events, and when the focus is a contemporary phenomenon within some real-life context (Yin, 1994). Yin (1994) stated that this is especially relevant "when the boundaries between phenomenon and context are not clearly evident" (p. 13). In the end, the context of the study can make meaning of the phenomenon as much as the phenomenon itself (Lincoln & Guba, 1985). As a result, understanding the culture, behavior, and actions of the actors in schools required such a deep methodological approach (Geertz, 1973; Wolcott, 1994).

The case study design has proven important to understand systemic patterns in schools. The unit of analysis, the school district, was treated as the "lived" organization that it is (Cobb, McClain, de Silva Lamberg, & Dean, 2003). In the end, the robustness of the work was in the researcher's ability to develop a full, rich understanding of actors in the organization at multiple levels.

This study sought to understand the phenomenon or activity of data (e.g. utilization in administrative decision-making and pedagogy, resource allocation, professional development, etc.) in one school district bounded by time (Creswell, 1998; Merriam, 1988; Yin, 1994). Yin (1994) cited two elements that characterize the need for a single site case study design. First, there must exist critical elements to the case. This study was designed to challenge a theory of action in regard to assessment accountability. Additionally, the analytical framework constructed from this study uncovered critical elements and dynamics untold and unforeseen in the theory of action. Second, Yin

(1994) posited that the case must be unique and relevant. The No Child Left Behind legislation of 2001 created unprecedented demands, incentives, and sanctions to utilize student achievement data in schools. This type of work is especially relevant in an undefined territory such as assessments in the new era of high-stakes assessment accountability.

Because the study was centered on the activity of the use of assessment data, understanding the relationship between the organization and the data was crucial. Additionally, how the individuals made meaning out of their lived experiences was important to understand (Moustakas, 1994). Such inquiry required the researcher to bracket or suspend the phenomenon to understand it through the voices of the informants who are living the phenomenon (Moustakas, 1994). The researcher's immersion into fieldwork helped uncover gatekeepers and locate key informants (Creswell, 1998; Wolcott, 1994). As a result, the researcher's ability to be embedded in the context of the school district allowed for a myriad of data sources.

Issues of Trustworthiness and Soundness

Issues of credibility (internal validity), transferability (external validity), dependability (reliability), confirmability (objectivity), and construct validity were addressed in this study (Lincoln & Guba, 1985; Yin, 1994). The study's construct validity and credibility were addressed by the multiple sources of evidence collected, by a prolonged engagement in the field, through the analysis of patterns in the data, and through the member checks of the findings. Such a design assured that the inquiries reflected the respondents' views and disengaged the research from any bias (Lincoln & Guba, 1985). The transferability of this case was only to the site studied. Case studies

are "generalizable to theoretical propositions and not to populations or universes... and the investigator's goal is to expand and generalize theories" (Yin, 1994, p. 10). The research design attempted to analyze a single site in order to understand one critical case of assessment accountability in schools. In order to generalize the finding of this work, additional studies, including replication studies, would have to be employed. Because this was a single site case study, the attention to dependability and confirmability was more vital to this work. The dependability of the work was rooted in the case study protocol. The triangulation of multiple sources of data that was collected and placed in a database was integral to this work (Merriam, 1988; Miles & Huberman, 1994; Patton, 1990; Yin, 1994). Additionally, the reflexive design of coding (open) and recoding (axial) of data, based on the emergent framework, provided dependability. Finally, the confirmability of the study was also anchored in the triangulation of data. Confirming and comparing what people said and did at different times was vital to this work (Patton, 1990). In the end, the design of multiple sources of data and an in-depth approach preserved the issues of trustworthiness.

Site Selection and Population

The researcher selected an urban school district as the unit of study. The district as the unit of analysis provided a dynamic view of multiple organizational levels. Individual schools provide only an atomistic view while the district must mediate state and federal pressures between the statehouse and schoolhouse (O'Day & Smith, 1993; Spillane, 2004; Sykes et al., under review). Elmore (2004) stated that the district could act as a reinforcement of "patterns of volunteerism, idiosyncrasy, and instability of goals in the way they deliver assistance to teachers and schools" (p. 288). However, the district

can also amplify reforms. As a result, understanding dimensions and actions of educators required a systemic district view, not just a school view (McLaughlin & Talbert, 2002; Spillane & Thompson, 1997).

An urban school district was selected for a number of reasons. Urban districts continue to be publicly scrutinized on the basis of poor performance. As a result, urban schools have been testing grounds for school reform (Resnick & Glennan, 2002). Burch (2002) supported this assertion, "Rather than static backdrops for reform, urban school districts are dynamic settings that create both challenges and strategic opportunities for schools and the intermediary organizations that work for them" (p. 125).

Because all districts across the United States must attend to the NCLB legislation, a random site selection was possible for this study. However, other considerations including size, location, and accessibility to a site had to be considered. The investigator selected a district that was large enough that variations of use would occur (Miles & Huberman, 1994). Also, the proximity of the site had to be considered as a result of the study's design, specifically the researcher's immersion at the site. Finally, access into a school district to conduct a research project required an opportunistic strategy (Marshall & Rossman, 1999). Marshall & Rossman (1999) describe a realistic researchable site to be where: Entry is possible; There is a high possibility that a rich mix of the processes, people, programs, interaction, and structures of interest are present; The researcher is likely to be able to build trusting relations with the participants in the study; and Data quality and credibility of the study is reasonably assured (p. 69). As a result, a purposeful sampling technique was utilized to select an urban school district that provides opportunities to learn a great amount of information about assessments and schools

(Merriam, 1988). Lincoln and Guba (1985) support the notion of purposeful or theoretical sampling because of the range of data collected as well as the "full array of multiple realities [that] will be uncovered" (p. 40). Consequently, an urban district in proximity to the researcher was selected for this study. It was the intention of the researcher to develop a quid pro quo relationship with the selected district. That is, the district was privy to the emerging findings as a result of this work, and, more importantly, district personnel were afforded the opportunity to engage in conversations of pedagogy and policy.

The study required data to be collected from three levels of the organization (central office, building principals, and teachers). In addition, the researchers aim to become immersed in the functions of elementary and secondary schools within the district led to the selection of two focus schools. Two focus schools in the district were selected (one elementary and one high school) for in-depth analysis.

Once the sites for this work were procured, the researcher created population clusters (central office, principals, teachers, and a consultant). From each cluster of participants the researcher utilized a rolling sample technique (Miles & Huberman, 1994). Data were collected from each cluster until a saturation of data were established (Glaser & Strauss, 1967; Miles & Huberman, 1994). A number of factors were instrumental in the selection of participants in each cluster: identification of key actors via immersion in various district settings, consent of actors to participate, and a diverse vector of demographic characteristics of actors. Specifically, the researcher paid close attention to the demographic vectors of race, age, gender, subject area taught, and years of teaching to ensure a representative sample of the population within each of the

clusters. Additional building administrators (n=4) and teachers (n=2) from non-focus schools were selected to participate in order to add depth to the work and to provide a member check of the early analysis of data. Finally, one external consultant was added to the study as a result of the relevance of her work to the study's research questions.

Data Collection

This qualitative design allowed for a myriad of data collection strategies. The collection strategy could be best categorized as emerging because of the on-going analysis element of this design (Creswell, 1998). The goal of the collection of data was to generate an on-going conversation and collection of artifacts to uncover facts, opinions, and insights (Yin, 1994). The primary methods of data collection included in-depth, semi-structured interviews (Merriam, 1988; Patton, 1990), observations, and strategic collection of artifacts. This array of data supported Yin's (1994) recommendation of the six types of information that can be collected: documentation, archival records, interviews, direct observation, participant observation, and physical artifacts.

Interviews

A multi-layered interview strategy was employed in this study. To better understand the organizational strategy, a number of central office personnel were interviewed. Subsequently, the tactical views of managing the district strategy led to interviews of school principals. Finally, the operational perspectives of the implementers of assessment accountability led to interviews with a number of teachers.

The utilization of in-depth interviews provided the researcher with data based on contexts, human activity, and intuitive knowledge (Marshall & Rossman, 1999). As

Patton (1990) stated, "The purpose of interviewing is to find out what is in and on someone else's mind. Qualitative interviewing begins with the assumption that the perspective of others is meaningful, knowable, and able to be made explicit" (p. 278). The interview protocols focused on both the organizational features and individual actions of how data were collected, analyzed, and utilized in the school district. The main themes of the interview protocol (see Appendix C) included: (a) contextual and demographic characteristics about interviewees and their settings; (b) their individual perceptions of student assessment data; and (c) their own experience and behaviors with data collection, analysis, and use. A well-structured protocol ensured consistency of information obtained (Miles & Huberman, 1994; Patton, 1990). The semi-structured design allowed for additional responses and follow-up questions (Merriam, 1988; Patton, 1990). The interviews were audio-taped and transcribed (verbatim). Each interview lasted between forty-five and ninety minutes. The transcribed interviews were then housed in an electronic database, *NVivo*, for subsequent analysis. Table 13 provides a catalogue of interview participants.

Table 13: Study Interview Participants

Pseudonym	***Title***	***Interview Date***
Reo School District (n=6)	School District	
Dr. Pinkerton	Superintendent	December 10, 2003; June 10, 2004
Mrs. Greenly	Chief Academic Officer	April 5, 2004
Dr. Whitehurst	Assistant Superintendent for Curriculum	May 13, 2004
Mrs. Redding	Director of Student Assessments	March 18, 2004
Mrs. Rose	Assistant Director of Curriculum	March 22, 2004; November 8, 2004
Mrs. Grey	Area Director	June 18, 2004
River High School (n=9)	High School	
Mr. Nile	Principal	March 29, 2004; April 26, 2004; June 10, 2004
Mrs. Platte	Assistant Principal	June 9, 2004
Mrs. Charles	Counselor	May 10, 2004

Pseudonym	*Title*	*Interview Date*
Mrs. Hudson	Teacher	May 20, 2004
Mr. Thames	Teacher	May 5, 2004
Mr. Seine	Teacher	May 5, 2004
Mrs. Rhine	Teacher	May 27, 2004
Mr. San Juan	Teacher	May 4, 2004
Mrs. Tiber	Teacher	May 10, 2004
Wood Street Elementary (n=6)	Elementary School	
Mrs. Willow	Principal	March 1, 2004; March 8, 2004
Mrs. Oakley	Teacher	May 21, 2004
Mrs. Pinecrest	Literacy Coach	May 21, 2004
Mrs. Balsa	Literacy Teacher	May 21, 2004
Mrs. Rosewood	Teacher	May 13, 2004
Mr. Burch	Teacher	May 28, 2004
External Consultant (n=1)		
Mrs. Lysander	Professional Development Consultant, Athens University	May 4, 2004
Additional Building Administrators (n=4)		
Mr. Trout	Principal, Kilgore High School	May 24, 2004
Mr. Caulfield	Principal, Salinger Middle School	May 26, 2004
Mrs. Moriarty	Principal, Paradise Elementary School	May 26, 2004
Mr. Lombardi	Principal, Packer Elementary School	May 21, 2004
Additional Teachers (n=2)		
Mrs. Templeton	Teacher, Zuckerman Farm Elementary	November 8, 2004
Mrs. Pilgrim	Ilium High School	November 8, 2004

Observations

Fieldwork was an important aspect of data collection for this work. A field journal was utilized to document formal and informal meetings observed and held with school personal. The observational fieldwork included attending a number of site-specific and district wide meetings and professional trainings (see Table 14). Additionally, a number of hours were spent at the two schools that were the focus of this work (Wood Street Elementary and River High School).

Table 14: Fieldwork Timeline

Observation	*Location*	*Date*
Meeting Attendance		
Wood Street School Improvement Meeting	Wood Street Elementary	March 10, 2004
District-wide School Improvement Meeting	Local Hotel Conference Center	March 17, 2004
River High School Improvement Meeting	River High School	April 26, 2004
River High Staff Meeting	River High School	February 28, 2004
Wood Street School Improvement Meeting	Wood Street Elementary	March 8, 2004; May 3, 2004
Wood Street Professional Training (LTRS)	Wood Street Elementary	March 5, 2004
District-wide Professional Development Workshop	City Conference Center	August 26, 2004
School Observations		
Wood Street Elementary	Wood Street Elementary	Multiple Visits: March-June 2004
River High School	River High School	Multiple Visits: March-June 2004

Artifacts

While case study methodology relies on direct observation and systemic interviewing for data collection, additional data collection methods were important to this work. Because the research context consisted of both individual educators as well as resources artifacts created in schools (D. K. Cohen, Raudenbush, & Loewenberg Ball, 2003), a number of artifacts needed be collected from the school district (see Table 15). The collection of documents and artifacts provided a "ready-made source of data easily accessible to the imaginative and resourceful investigator" (Merriam, 1988, p. 104).

Table 15: Artifacts Collected

Artifacts
Minutes form School Improvement Meetings (Wood Street and River High)
K-12 District Pacing guidelines
District Quarterly Assessments (Grades 3-12)
Board of Education Policies on Assessment Data (Policy #5121)
District Evaluation Instrument
District Testing Calendar
District School Improvement Checklist and Guidelines
District Quarterly Assessment Rationale

Artifacts
Assessment Item Analysis Worksheet (Quarterlies)
Quarterly Assessment Analysis Reports (Item Frequency, Student Scores and Responses, Building Exam Report, Building Test Summaries, Comparison Charts)
Handouts from Training Sessions (LETRS and Data Workshop)
LETRS Training Packet
DIBELS Assessment Reports (Class Report, School Report)
District Annual Reports (2000-2004)
School Annual Reports (Wood Street Elementary and River High)
School Newsletters
Standards and Poors Documents (2003)
Disaggregated MEAP Student Assessment Reports (2000-04)
Individual School Report Cards Provided by the District
Released state MEAP and HST assessments

Data Analysis

Data analysis was conducted concurrently with data collection. The data collected informed the analysis and vice versa. As data were collected and analyzed, the collection protocol was modified to collect missing data or to attain additional data to address issues of validity. This technique ultimately lead to a point of saturation where the on-going analysis was repeating and confirming prior data collected (Creswell, 1998; Glaser & Strauss, 1967; Marshall & Rossman, 1999; Miles & Huberman, 1994). Finally, the member check strategy of sharing the initial findings with the host school district proved helpful, and conducting additional building administration interviews determined both the saturation and accuracy of the data collected (Rudestam & Newton, 2001).

The analysis techniques employed in this study were a product of the data collected. Specifically, the interview database, field notes, memos, and artifacts collected were coded by themes and analyzed with a general content analysis (Lincoln & Guba, 1985) and specific pattern matching (Creswell, 1998; Stake, 1994). Themes emerged from the constant comparative analysis of data (Creswell, 1998; Glaser & Strauss, 1967; Lincoln & Guba, 1985) and patterns from the findings (Creswell, 1998; Miles &

Huberman, 1994; Stake, 1994). Pulling together a number of small chunks of data allowed the researcher to find "Gestalts which pull many pieces of data together" (Lofland, 1971, p. 246).

Coding the Qualitative Data

The analysis of the data collected can be characterized by three distinct stages. The first stage of analysis involved memo creation from field notes and artifact review. Artifact documents and field notes were scanned against the research questions as well as participant responses and then formulated into memos (Merriam, 1988). Miles and Huberman (1994) stated, "Memos are primarily conceptual in intent. They don't just report data; they tie different pieces of data together in a cluster, often to show that those data are instances of a general concept" (p. 72). The creation and coding of memos was another method to expose the patterns in the data collected (Bogdan & Biklen, 1992). The memos and transcribed interviews were then exported into a computer database. The qualitative software program *NVivo* was utilized for this work.

Stage two of the analysis involved an initial coding of the memos and transcripts into the *NVivo* database. The identification of codes in the qualitative data was vital to the analytical work (Bogdan & Biklen, 1992). In order for pattern analysis to take place, themes first had to be uncovered and categorized. To begin, an open coding technique was employed utilizing "Free Nodes" in the *NVivo* software (Bogdan & Biklen, 1992; Gibbs, 2002; Miles & Huberman, 1994). This technique allowed the researcher to extract relevant information from the large amount of data collected. The open coding developed nodes or themes that were shaped by the interview protocol (see Table 16).

Table 16: Open Coding- Free Nodes

Current Job and Responsibilities	Policy Mediate Instruction
Effective Instruction	Effective Leadership
Lesson Plan Artifacts	Types of Assessments
Belief of Data Improving Achievement	District Expectations Purpose
Resources and Support	Assessment Likes
Assessment Dislikes	Barriers to Use (Will or Capacity)
Effects of DM and Pedagogy	Use of Sub Group Data
HOTs and Teach to Test	Teacher Evaluation
Student Evaluation	Communication
Rate Your Ability	Pressures from Assessments
Compliance with Assessments	Needs to be Done Different
Ideal Use of Assessments	Other

The third stage of the analysis utilized an axial coding where the researcher identified phenomenon, conditions, contexts, consequences, and intervening conditions from the data (Creswell, 1998; Strauss & Corbin, 1990). Utilizing the Free Nodes, the data were inductively analyzed (Yin, 1994). The progression of data from a managed, warehouse format to an analytical procedure began to reveal features and relationships (Wolcott, 1994). As a result, "Tree Nodes" were created in *NVivo* to illustrate "families" or themes and relationship characteristics (Gibbs, 2002). Table 17 depicts the Tree Nodes that were created. The *NVivo* software allowed the researcher to run node reports that could be viewed in both a narrative and graphical context (Gibbs, 2002).

Table 17: Axial Coding- Tree Nodes

District Plan of Action	Data for Student Placement
Data for Pedagogical Decision-Making	Impediments to Use
Teachers Beliefs	CO Beliefs
Instructional Leadership	Data Metrics
Curricular Breadth and Pacing	Organizational Coherence

Confidentiality and Ethical Considerations

The security of the data collected and the confidentiality of the participants were of the utmost importance in this study. Pseudonyms were utilized for the district, the

schools, and all of the participants in the study. In addition, the transcription of interviews and tapes, field notes, and documents collected was kept in a secure, locked location. Finally, none of the material collected from district personnel was replicated or disseminated in any way.

As previously mentioned, the district was privy to an executive summary report as well as member checking of the initial findings. When any findings or reports were shared with the district every attempt to mask the identity of subjects was made. The superintendent of the district made it clear that the executive summary that would be submitted would be used for future decision making in regards to district assessments and not for the purpose of seeking out school personnel who did not support the assessment accountability efforts of the district.

In order for the researcher to conduct the study, a formal application was made and approved by the district's Office of Research and Evaluation as well as the University Committee on Research Involving Human Subjects (UCRIHS). Finally, a UCRIHS consent letter of participation was signed and on file for each participant of the study (see Appendix D for the study's consent forms).

Limitations of the Study

Leonardo de Vinci warned that, "all our knowledge has origins in our perceptions." Similarly, Max Weber cautions that people bring with them a number of theoretical and experiential presuppositions. Because this researcher came into the study with eleven years of public school experience, attempts to bracket personal understandings and experiences from those being studied were essential (Creswell, 1998). This researcher made every attempt to suspend personal beliefs and prior knowledge of

schooling in order to allow the data to emerge and to construct a theory of understanding. In other words, the researcher entered the district as a learner, not an expert (Creswell, 1998). Miles and Huberman (1994) stipulated that an espoused conceptual framework is the "current version of the researcher's map of the territory being investigated" (p. 20). As a result, while the conceptual framework guided the analysis of this work, the researcher also looked at other theoretical explanations during the analysis. Finally, the study was limited by the sample size and the purposeful sampling methodology. Consequently, the conclusions for this case are intended for this case alone.

Conclusion

This design was holistic enough to tell a district wide story and robust enough to understand individuals actions. This methodological design also fulfilled Yin's (1994) three essential elements of case study methodology by establishing: a case study database, a collection of multiple sources of evidence, and a chain of evidence. Moreover, the design of this work allowed for the data and analysis to work in tandem in order for a framework to develop. Any one theoretical perspective did not hamper the study. That is, this was a cyclical process of the collection of data informing analysis and analysis informing data collection, or as Wolcott (1990) eloquently stated, "The nexus between description and analysis in the written account is also dialectical- each process informing the other" (p. 50).

APPENDIX B

STUDY RESPONDENTS

Dr. Pinkerton, Superintendent

Dr. Pinkerton (African American) was named superintendent of the Reo School District in 2000. She earned her doctorate in 1986 in educational administration and supervision. Prior to her current position, Pinkerton had worked in an urban school system in a neighboring state. She spent ten years in the classroom as a teacher and afterward took other positions in the district including: counselor, parent and Title I coordinator, high school assistant principal, and dean of schools. Dr. Pinkerton left the district to take the position of chief of staff for the city's mayor for three years. Upon her return to the district, Pinkerton became an area director in charge of nineteen buildings.

Pinkerton's tenure at Reo has been marked by a number of successes. District dropout rates have decreased and graduation rates are up nearly 10% since her arrival. In addition, grant money supporting district programs increased dramatically, including awards totaling $40 million in 2002-03. Under Pinkerton's leadership, Reo successfully passed a $67.5 million building improvement bond in 2003. Dr. Pinkerton was named the state superintendent of the year in 2001. She continues to meet Board expectations and earns many of her salary performance bonuses.

Mrs. Greenly, Chief Academic Officer

Mrs. Greenly (African American) was previously an elementary school principal in Reo. Her new position was created when Dr. Pinkerton came to the district. The position involves working with schools to increase student achievement. Specifically,

Greenly is charged with helping schools to increase scores on the State Educational Assessment Program (MEAP).

Dr. Whitehurst, Assistant Superintendent for Curriculum

Dr. Whitehurst (African American) began her career in education in 1970 as an elementary teacher. Whitehurst was also a Title I resource teacher in grades K-8 for fourteen years and an elementary principal for ten years. This is her third year in the position of assistant superintendent. Her responsibilities include the direction of the curriculum and professional development. In addition, her office has oversight of all state and federal grants. Dr. Whitehurst also supervises the district's assessment writing groups. Finally, Whitehurst leads the district school improvement team.

Mrs. Redding, Director of Student Assessments

Mrs. Redding (African American) has worked in the district for the past thirty-three years. All of her positions in the district have involved research and evaluation. She has a masters degree in measurement and program evaluation. As the director of student assessment, Redding is responsible for overseeing student assessments administered in the district save the district's Quarterly Assessments. Additionally, her office evaluates state and federal special funded programs and board initiatives. Redding's office employs four secretaries, five project development specialists, and five evaluation assistants.

Mrs. Rose, Assistant Director of Curriculum

Mrs. Rose (Caucasian) coordinates the district's Quarterly Assessments. Her work includes making copies of the assessments and following up with principals in terms of teachers who have not completed assessments. Rose has worked in the district

since 1997 when she was a curriculum facilitator. While Rose has earned a teaching certificate in secondary science, her previous career focused on quality control issues in area hospitals. Mrs. Rose became involved in the school district as a parent of an elementary child. She was an active member of a district committee on grade level expectations. In 1996, she was asked to join the district and begin work on helping the schools align curriculum to the state's curricular framework.

Mrs. Grey, Area Director

Mrs. Grey (African American) has been an educator in the district for the past thirty-two years. Grey was an elementary principal for twenty years and then the director of elementary education in the school district. Grey supervises principals in one of the district's high schools and all of its feeder schools. Grey meets formally with all the principals once a month. They discuss management and student achievement issues.

Mrs. Willow, Principal, Wood Street Elementary School

Mrs. Willow (Caucasian) just completed her first year as principal of Wood Street Elementary. Willow was previously a middle school assistant principal in the district. Willow took over an underachieving school (as per the MEAP) that has a high mobility rate. The school that has a great deal of pride as the district's oldest elementary school continues to report poor assessment scores.

Mrs. Oakley, Teacher, Wood Street Elementary

Mrs. Oakley (Caucasian) teaches 3rd grade. She is also a graduate of the Reo District. She was hired in the district in 1990 and began teaching 1st grade. Since 1994 she has taught 3rd grade. While she has had many opportunities to transfer to other

schools in the district, Oakley feels deeply committed to the students in the neighborhood.

Mrs. Pinecrest, Literacy Coach, Wood Street Elementary

Mrs. Pinecrest (Caucasian) has taught early elementary for thirty-four years in the Reo District. This is her first year as a literacy coach, a position that is funded by the Reading First Grant. Mrs. Pinecrest has received training and continued support through Dr. Whitehurst's office. She also attends special training sessions provided by the textbook publisher Houghton-Mifflin. Most of her time is spent with students and teachers in grade K-3. She works with teachers during their planning time to talk about implementation of the LETERS training and the analysis of DIBELS reports. In addition, Pinecrest works with students who are not reading at grade level. Finally, Pinecrest has established a literacy library at Wood Street as a support for teachers.

Mrs. Balsa, Literacy Teacher, Wood Street Elementary

Mrs. Balsa (Caucasian) is in her fourth year of teaching. Prior to becoming the schools literacy teacher, she taught three years at Wood Street Elementary in the 5th grade. Each school in the district has a literacy teacher position funded with Title I monies. Balsa is required to work with students who are in the 29th percentile in reading comprehension as determined by the Iowa Basic Skills Test. She also utilizes achievement data from the DIBELS and Gates-MacGinitie, as well as teacher recommendations. Balsa currently works with sixty-three students, grouped by grade level, on guided reading strategies. She also meets with teachers before or after school.

Mrs. Rosewood, Teacher, Wood Street Elementary

Mrs, Rosewood (African American) teaches 2nd grade. She is a member of the Wood Street School Improvement Team. Rosewood grew up in the area and graduated from the Reo schools. She has worked in Wood Street Elementary for the past nineteen years, all in the lower elementary grades. Rosewood believes her upbringing in a single parent home has helped her connect to the needy children of the neighborhood.

Mr. Burch, Teacher, Wood Street Elementary

Mr. Burch (Caucasian) teaches a 3rd/4th multi-age class. He is currently a member of the school improvement team. Burch has been a teacher in the district for his entire twelve-year career. During his first year in the district, he taught at Salinger Middle School. Since that time he has been a member of the Wood Street faculty, teaching various early elementary grades.

Mr. Nile, Principal, River High School

Mr. Nile (African American) is in his 2nd year as principal of River High School. He previously spent eight years as a high school assistant principal in an urban school on the east side of the state. Nile was also a director of bands in an urban school on the west side of the state. Nile began his career in education after a stint as a professional jazz musician.

Mrs. Platte, Assistant Principal, River High School

Mrs. Platte (Caucasian) is in her 30th year in the district. She began her career as a special education teacher. She has been at River High since 1987 as an assistant principal. She is currently the assistant principal for curriculum. Her role entails

conducting a majority of the teacher evaluations (thirty this year). She is also the designated test coordinator for the school.

Mrs. Charles, Counselor, River High School

Mrs. Charles (Caucasian) is one of three River H.S. counselors. Prior to becoming a counselor, Charles taught English and psychology for five years at River H.S. Her responsibilities include career exploration and counseling, class scheduling, and teaching an advanced placement psychology course. Charles stated that the counselors do very little personal counseling.

Mrs. Hudson, Teacher, River High School

Mrs. Hudson (Caucasian) worked in the federal prison system in Texas prior to a career in education. She became a teacher in Texas under an emergency certification while completing her degree program that included a reading specialist endorsement. Hudson returned to the state in 1998 and found work at a local charter school. In December of 1999 she was hired at River H.S. as an English teacher and literacy coach with monies from the Comprehensive School Reform Grant. As a result, Hudson worked with other English teachers to design and teach a 9th grade reading class. Hudson currently teaches two 9th grade English courses, two 10th grade American Literature courses, and is a member of the school improvement team. She is provided one planning period to continue work as the school's literacy coach.

Mr. Thames, Teacher, River High School

Mr. Thames (Caucasian) teaches English and journalism and serves as the co-department chair. He is currently a member of the school improvement team. Thames

worked for ten years in the newspaper business and then five for an automotive company on publications. He has been a teacher at River H.S. for the past nine years.

Mr. Seine, Teacher, River High School

Mr. Seine (Caucasian) has taught science (physics, chemistry, and forensic science) at River High since 1998. Previously, Seine was a restaurant manager for eight years. Upon moving to the area, Seine entered a post B.A. teacher certification program at Athens University. He is very involved in the school improvement team as the assessment chair (mandated by the Comprehensive School Reform Grant) as well as acting as the advisor for the school's Science Club, Quiz Bowl, and Academic Decathlon. Seine is an original member of the science writing team for the district's Pacing Guides and Quarterly Assessments.

Mrs. Rhine, Teacher, River High School

Mrs. Rhine (Caucasian) began her teaching career in Texas in 1980. Upon returning to the state in 1987, she was hired as an English teacher at River High. Rhine currently teaches a 9th grade reading class, America Literature, and 9th and 11th grade English. She also serves as co-chair of the school's English Department.

Mr. San Juan, Teacher, River High School

Mr. San Juan (Hispanic) is a social studies teacher at River High. San Juan began his career in 1975 teaching English as a second language and elementary math. He became a vocational education and career coordinator at one of the district's three high schools prior to taking an administrative position for two years in a local district. San Juan returned to Reo in his current position ten years ago.

Mrs. Tiber, Teacher, River High School

Mrs. Tiber (Caucasian) is a social studies teacher and department chair at River H.S. The district hired her in 1995 as a middle school teacher. Tiber came to River High in 1998 and currently teaches U.S. History, Honors World History (10th grade), and AP History (11th and 12th grade). As a department chair, Tiber conducts monthly meetings and distributes and collects the district's Quarterly Assessments. Tiber acknowledged that she is currently the department head even though she has the next to last amount of seniority in her department.

Mr. Caulfield, Principal, Salinger Middle School

Mr. Caulfield (Caucasian) has been the principal of Salinger since 2002. He began his career in education in the district in 1978. Prior to his current position, he had been an elementary teacher and administrator. Salinger M.S. has a student enrollment of 866 students of which 64.8% are economically disadvantaged and 69% receive free or reduced lunch. The school has a 39.1% MEAP pass rate and a high mobility rate.

Mrs. Moriarty, Principal, Paradise Elementary School

Mrs. Moriarty (Caucasian) has been the principal of Paradise Elementary School since 2001. She began her career as a bilingual teacher in the district and then became a principal of an elementary school that eventually was closed as a result of the declining district enrollment. Paradise Elementary School has a student population of 231. 50% of the student population is considered economically disadvantaged. However, the school has a 59.2% MEAP pass rate, highest elementary rate in the district. As a result, Paradise is considered the district's showcase elementary school. The school has a low mobility rate, and 48.6% of the students receive free or reduced lunch. Mrs. Moriarty described

the school as a district showcase for three reasons: high student test scores, supportive parents, and teachers who are deeply committed.

Mr. Lombardi, Principal, Packer Elementary School

Mr. Lombardi (Asian) has been the principal of Packer Elementary School for the past two years. Mr. Lombardi has a mixed cultural background. He grew up in the Bronx of New York with a Japanese mother and a West Indian father. Prior to his current position he was a teacher and assistant middle school principal in a neighboring district for twenty-one years. Mr. Lombardi describes his school as a community. He has a strong sense of social justice and advocates for his students. He works at involving the parents and the neighboring business community in the school. As a result, he has been able to provide the students of Packer many cultural opportunities and experiences. Packer Elementary enrolls 276 students. 83.5% of the students are considered economically disadvantaged, and 80.7% of the students receive free or reduced lunch. The school has a 40.6% MEAP pass rate and a high mobility rate (approximately 30% per year).

Mr. Trout, Principal, Kilgore High School

Mr. Trout (Caucasian) has been employed by the Reo School District since 1967. He has held a number of teaching and administrative positions in the district. He claims he has been moved to different administrative positions to solve various problems. He believes he has survived in the district for so long because he is politically astute. Kilgore H.S. enrolls 1697 students of which 37.7% are labeled as economically disadvantaged, and 43.3% of the students receive free or reduced lunch. The school has a 38.7% MEAP pass rate. Mr. Trout is very proud of the school's programs and believes

that the high student mobility rate (200-300 students per year) impairs the schools assessment scores.

Mrs. Templeton, Teacher, Zuckerman Farm Elementary School

Mrs. Templeton (Caucasian) has been a teacher in the Reo district since 1992. Upon completion of her teaching certification in 1992, Templeton began her second career in education. She has taught 4th and 5th grades during her employment in Reo. Currently, Templeton teaches a 4th/5th grade split class at Zuckerman Farm Elementary School. Zuckerman Farm consistently performs in the middle to upper range of assessment scores and state's report card for the district. The school has a low mobility rate as compared to other elementary schools in the district.

Mrs. Pilgrim, Teacher, Ilium High School

Mrs. Pilgrim (Caucasian) began her second year of teaching in the 2004-05 school year at Ilium High School. A recent graduate of Athens University, Pilgrim teaches two different algebra classes: algebra and a special algebra class which covers the regular algebra curriculum over a year and a half. Ilium High School is one of the three high schools in the district. Ilium has similar demographic characteristics and performanced indicators as the other two high schools in Reo. Ilium is currently in the midst of implementing an International Baccalaureate program in the high school.

Mrs. Lysander, Data Consultant, Athens University

Mrs. Lysander (Caucasian) is an employee of Athens University. She was a social studies teacher prior to working for the university. She was hired at Athens in 1981 to create an on-line database that combined multiple types of information into a single accessible system. Since that initial contract, Mrs. Lysander has been employed by

the university and has worked in the area of student data. She describes herself as an "applied demographer" rather than a statistician. Mrs. Lysander conducts seminars with local school districts on utilizing student achievement data. Recently, Mrs. Lysander has been contracted to work with the Reo School District. She has conducted a series of workshops for district administrators over the past three years. The workshops introduce administrators to the various forms of data and its utility to drive educative decisions.

APPENDIX C

INTERVIEW PROTOCOL

Part A- Contextual/Demographic General Information

1. Tell me about yourself.
    a. Educational background
    b. Tenure in your position
    c. Past experiences
    d. Current job and responsibilities
    e. Structure of your day (including planning time)
    f. Any special programs in your district (especially those that may create additional assessments)

2. Please tell me what effective instruction means to you.

3. Please tell me what effective leadership means to you.

4. What materials do you have (should teachers have) in front of you (them) for lesson planning?

5. Can policy and procedures mediate good instruction? If so, how?

Part B- Perceptions of Student Achievement Data & Assessments

1. Please rate from 1 to 4 the extent to which using data can increase student achievement, where "1" means "hardly any" and "4" means "substantially"? Please explain.

2. What specific assessments involve you and your students?
    a. What are the district's expectations for principals and teachers?
    b. Are the schools complying with the district expectations? If so, in what ways?
    c. Have teachers' changed their practice to meet the demands of assessments in your district? [For administrators]

3. What kinds/forms of raw and disaggregated data do you receive from the district/school?
    a. What are the district's expectations for the use of the data?
    b. What resources and support do you receive from the district/school (SIT/Dept.) in regard to utilizing student achievement data?

c. Is the data you receive timely?
d. Do you conduct any data analysis yourself? If so, explain.

4. Tell me at least two things you like about the district's focus on student assessments and what assessment data you receive is most useful to your practice.

5. Tell me at least two things you dislike about the district's focus on student assessment.

6. Is there too much testing in the Schools?

Part C- Student Achievement Data and Your Practice

1. [TEACHERS] Please describe a couple of examples, if any, of ways that you have used student achievement data to change your instruction?
    a. Have you created special lessons, rubrics, tests, etc. to meet the demands of assessments in your district?
    b. Do you use the district's Quarterly Assessments as part of the classroom level assessment of student achievement?
    c. Are there any aspects of your previous teaching that have been forsaken as a result of the assessment focus (HOTS, teaching to the test, etc.)?
    d. Have you made a concerted effort to look at the student sub-group data? If so, what have you done with the data?

2. [FOR CENTRAL OFFICE] Please describe a couple of examples, if any, of ways that you have used student achievement data to change your practice?
    a. Please list up to three decisions you make that are affected by assessment data.
    b. Please list up to three decisions that principals make that should be affected by assessment data.
    c. Please list up to three decisions that teachers make that should be affected by assessment data.
    d. What resources and support do you provide your principals at the building level? [e.g. professional development activities, conferences]
    e. What resources and support do you provide your teachers at the building level?
        i. Is there a problem with the varied planning times at the HS and elementary schools?
    f. What resources and support do you provide your students at the building level? [e.g.- remedial classes, test prep, grants, etc]
    g. Have you made a concerted effort to look at the student sub-group data? If so, what have you done with the data?

3. [FOR PRINCIPALS] Please describe a couple of examples, if any, of ways that you have used student achievement data to change your practice?

a. Please list up to three decisions you make that are affected by assessment data.
b. Please list up to three decisions that teachers make that should be affected by assessment data.
c. What resources and support do you provide your teachers at the building level? [e.g. professional development activities, conferences, SIT]
d. What resources and support do you provide your students at the building level? [e.g.- remedial classes, test prep, grants, etc]
e. How do you use the district's Quarterly Assessments to inform your practice?
f. Have you made a concerted effort to look at the student sub-group data? If so, what have you done with the data?

4. Are you/do you evaluate(d), formally and informally, in regard to student achievement data (report cards, district quarterly assessments, and MEAP)? If so, how and by whom?

5. Do you communicate with colleagues, administrators, parents, and/or students about student achievement data? If so, how?

    a. If Department Chair- How has your leadership position affected by student assessment data? [Decision making, leadership, support, artifacts]

6. How would you rate the your use of data (level 1= none (emergent), level 2= low (island), level 3= moderate (integrated), level 4= expert (exemplary)? Explain.

7. How do you feel when you receive student achievement data from the district that indicates poor performance?
    a. Are the demands to use data in your instruction creating stress in your professional life? Pressures for students and administration?
    b. Has the building and district leadership changed since the focus on student achievement data? If so, in what ways?
    c. Do you feel more pressure from the accountability associated with NCLB or the districts system? Why?

8. What do you think the district needs to do differently in order for you to better utilize student achievement data?
    a. Is your biggest obstacle your will/motivation or your capacity/knowledge?
    b. What is the principals' biggest obstacle will/motivation or capacity/knowledge?
    c. What is the teachers' biggest obstacle will/motivation or capacity/knowledge?

APPENDIX D

CONSENT FORMS: DISTRICT, UCRIHS, PARTICIPANT

February 3, 2004

Matthew Militello

Dear Mr. Militello:

Your request to conduct the study "An Exploratory Case Study of One District's Collection, Analysis, and Use of Student Achievement Data Required by No Child Left Behind," in the [redacted] School District has been approved. The identification of the [redacted] School District shall not appear in any publication without the written permission of the school district. It is understood with this approval, that the researcher will submit to the district a written report ninety days (90) after the conclusion of the study.

The following comments apply to the study:

> Participation in the study is strictly voluntary. Please contact me for technical assistance in carrying out the research activities.

If you have any questions or need additional information, please contact me

Thank you,

[redacted]

Director of Research, Evaluation and Pupil Accounting

lel

c: Research Review Committee

Research, Evaluation & Pupil Accounting

MICHIGAN STATE
UNIVERSITY

Revision Application Approval

November 12, 2004

To: Gary Sykes
410a Erickson Hall
Msu

Re: **IRB # 04-266** Category: EXPEDITED 2-6, 2-7
Revision Approval Date: November 5, 2004
Project Expiration Date: March 22, 2005

Title: AT THE CLIFF'S EDGE: UTILIZING EVIDENCE OF STUDENT ACHIEVEMENT FOR INSTRUCTIONAL IMPROVEMENT IN A SCHOOL DISTRICT

The University Committee on Research Involving Human Subjects (UCRIHS) has completed their review of your project. I am pleased to advise you that **the revision has been approved.**

This letter notes approval for the change made in the project title.

The review by the committee has found that your revision is consistent with the continued protection of the rights and welfare of human subjects, and meets the requirements of MSU's Federal Wide Assurance and the Federal Guidelines (45 CFR 46 and 21 CFR Part 50). The protection of human subjects in research is a partnership between the IRB and the investigators. We look forward to working with you as we both fulfill our responsibilities.

Renewals: UCRIHS approval is valid until the expiration date listed above. If you are continuing your project, you must submit an ***Application for Renewal*** application at least one month before expiration. If the project is completed, please submit an ***Application for Permanent Closure***.

Revisions: UCRIHS must review any changes in the project, prior to initiation of the change. Please submit an ***Application for Revision*** to have your changes reviewed. If changes are made at the time of renewal, please include an ***Application for Revision*** with the renewal application.

Problems: If issues should arise during the conduct of the research, such as unanticipated problems, adverse events, or any problem that may increase the risk to the human subjects, notify UCRIHS promptly. Forms are available to report these issues.

Please use the IRB number listed above on any forms submitted which relate to this project, or on any correspondence with UCRIHS.

Good luck in your research. If we can be of further assistance, please contact us at 517-355-2180 or via email at UCRIHS@msu.edu. Thank you for your cooperation.

Sincerely,

Peter Vasilenko, Ph.D.
UCRIHS Chair

c: Matthew Militello
403 Erickson Hall

OFFICE OF
RESEARCH ETHICS AND STANDARDS
University Committee on Research Involving Human Subjects
Michigan State University
202 Olds Hall
East Lansing, MI
48824
517/355-2180
FAX: 517/432-4503
Web: www.msu.edu/user/ucrihs
E-Mail: ucrihs@msu.edu

Interview Consent Form

The educational player piano: An exploratory case study of the collection, analysis, and use of student achievement data in a school district

Thank you for agreeing to participate in this study. By participating in this study, you will help researchers at Michigan State University document the collection, analysis, and use of English and social studies student achievement data in your school district. Interviews with district personnel will help the researchers better understand the role data plays in a school district. The interview questions will ask you about your practices under the district mandates and support to utilize student achievement data. As a result, you may be placed in a precarious position of reporting your actions that do not meet district mandates.

Your participation in this study is voluntary and confidential to the maximum extent allowable under federal, state and local law. Feel free to skip any question that you do not want to answer in the interview. The interview will take about 50 minutes. The conversation will be taped and transcribed by the doctoral candidate. Please know that at any time you can request the have the tape recorder turned off.

Your name will never be associated with the information you provide in this study, and no identifying information will ever be released to anyone. All the information gathered in this study will be kept confidential in a locked file cabinet. Any information that could identify you will be destroyed at the conclusion of the study. Finally, please be aware that you can withdraw from this study at any time with no repercussions.

If you have any *questions regarding this study*, please contact Matt Militello the doctoral student or his advisor Dr. Gary Sykes. If you should have any *questions or concerns regarding your rights as subjects and the duties of the investigators, or are dissatisfied at any time with any aspect of the study*, you may contact- anonymously, if you wish- Dr. Peter Vasilenko, University Committee on Research Involving Human Subjects at (517) 355-2180, fax (517) 432-4503, email: ucrihs@msu.edu or regular mail: 202 Olds Hall, East Lansing, MI 48824.

Sincerely,

Matt Militello
Doctoral Candidate, Michigan State University
403 Erickson Hall
East Lansing, MI 48824
(517) 353-5461
militel1@msu.edu

Gary Sykes, Ph.D.
Advisor, Michigan State University
410A Erickson Hall
East Lansing, MI 48824
(517) 353-9337
garys@msu.edu

Your signature below indicates your voluntary agreement to participant in the interview portion of the study.

Print Name:_____ Date:_____ Phone No._____ Email:_____

Signature:_____

Your signature below indicates your voluntary agreement to allow the researcher to audio tape record the interview.

Signature:_____

UCRIHS APPROVAL FOR THIS project EXPIRES:

MAR 2 2 2005

SUBMIT RENEWAL APPLICATION ONE MONTH PRIOR TO ABOVE DATE TO CONTINUE

BIBLIOGRAPHY

BIBLIOGRAPHY

AASA. (2002). *Using Data to Improve Schools: What's Working.* Arlington, Va.: American Association of School Administrators.

Adams, J. E., & Kirst, M. W. (1999). New demands for educational accountability: Striving for results in an era of excellence. In J. Murphy & K. S. Louis (Eds.), *Handbook of research in educational administration* (Second ed., pp. 463-489). San Francisco, CA: Jossey-Bass.

Allison, G. T. (1971). *Essence of decision: Explaining the Cuban missile crisis.* Boston: Little, Brown.

Amrein, A., L., & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Educational Policy Analysis Archives, 10*(18), Retrieved from http://epaa.asu.edu/epaa/v10n18/.

Anderson, G. (2004). Promoting educational equity in a period of growing social inequity: The silent contradictions of the Texas reform discourse. In L. Skrla & J. J. Scheurich (Eds.), *Educational equity and accountability: Paradigms, policies, and politics* (pp. 189-200). New York, NY: RoutledgeFalmer.

Apple, M. (2004). Creating differences: Neo-liberalism, neo-conservatism, and the politics of educational reform. *Educational Policy, 18*(1), 12-44.

Asp, E. (2000). Assessment in education: Where have we been? Where are we headed? In R. S. Brant (Ed.), *Education in a new era* (pp. 123-158). Alexandria, VA: Association of Supervision and Curriculum Development (ASCD).

Baier, V. E., March, J. G., & Saetren, H. (1988). Implementation and ambiguity. In J. G. March (Ed.), *Decisions and organizations* (pp. 150-164). New York: Basil Blackwell.

Ball, D., & Cohen, D. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3-32). San Francisco: Jossey-Bass.

Barber, M., & Phillips, V. (2000). Should large-scale assessment be used for accountability? The fusion of pressure and support. *Journal of Educational Change, 1*, 277-281.

Bernhardt, V. L. (1998). *Data Analysis for Comprehensive School Reform.* Larchmont, NY: Eye on Education.

Berry, B., Turchi, L., Johnson, D., Hare, D., Owens, D., & Clements, S. (2003). *The impact of high stakes accountability on teachers' professional development: Evidence from the south.* Chapel Hill, NC: Southeast Center for Teaching Quality, Inc.

Bidwell, C. (1965). The school as a formal organization. In J. G. March (Ed.), *Handbook of organizations* (pp. 972-1032). Chicago: Rand McNally.

Bidwell, C. (2001). Analyzing schools as organizations: Long-term performance and short-term change. *Sociology of Education, Extra Issue*, 100-114.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Bogdan, R. C., & Biklen, S. K. (1992). *Qualitative research for education: An introduction to theory and methods.* Boston: Allyn & Bacon.

Borman, G., Hewes, G., Overman, L., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*(2), 125-230.

Bossert, S. T., Dwyer, D. C., Rowan, B., & Lee, G. V. (1982). The instructional management role of the principal. *Educational Administration Quarterly, XVIII*(3), 34-64.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*, 32-42.

Brown, J. S., & Duguid, P. (1991). Organizational learning and communities of practice: Toward a unified view of working, learning, and innovation. *Organizational Science, 2*(1), 40-57.

Brown, J. S., & Duguid, P. (2000). *The social life of information.* Boston, MA: Harvard Business School Press.

Buchmann, M. (1986). Role over person: Morality and authenticity in teaching. *Teachers College Record, 87*(4), 529-543.

Burch, P. E. (2002). Constraints and opportunities in changing policy environments: Intermediary organizations' response to complex district contexts. In A.

Hightower, M. S. Knapp, J. A. Marsh & M. W. McLaughlin (Eds.), *School district and instructional renewal* (pp. 111-126). New York, NY: Teachers College Press.

Burkhardt, H., & Schoenfeld, A. H. (2004). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher, 32*(9), 3-14.

Carnoy, M., & Loeb, S. (2004). Does external accountability affect student outcomes? A cross-state analysis. In S. H. Fuhrman & R. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 189-219). New York, NY: Teachers College Press.

Carnoy, M., Loeb, S., & Smith, T. L. (2003). The impact of accountability policies in Texas high schools. In M. Carnoy, R. Elmore & L. S. Siskin (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 147-174). New York, NY: Routledge Falmer.

Chabran, M. (2003). Listening to talk from and about students on accountability. In M. Carnoy, R. Elmore & L. S. Siskin (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 129-146). New York: Routledge Falmer.

Chakwera, E., Khembo, D., & Sireci, S. (2004). High-stakes testing in the warm heart of Africa: The challenges and successes of the Malawi National Examination Board. *Education Policy Analysis Archives, 12*(29), Retrieved from http://epaa.asu.edu/epaa/v12n29/.

Choo, C. W. (2001). The knowing organization as learning organization. *Education & Training, 43*(4/5), 197-205.

Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, and America's schools.* Washington, D.C.: Brookings Institute.

Cizek, G. J. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practice, 19*(2), 16-23, 33.

Clune, W. H. (1987). Institutional choice as a theoretical framework for research on educational policy. *Educational evaluation and Policy Analysis, 9*(2), 117-132.

Cobb, P., McClain, K., de Silva Lamberg, T., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher, 32*(6), 13-24.

Cohen, D. K. (1988). Teaching practice: Plus que ca change... In P. Jackson (Ed.), *Contributing to educational change: Perspectives on research and practice* (pp. 27-84). Berkeley, CA: McCutchan Publishing Corporation.

Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis, 12*(5), 311-329.

Cohen, D. K., & Barnes, C. A. (1993). Conclusion: A new pedagogy for practice. In M. W. McLaughlin & J. E. Talbert (Eds.), *Teaching for understanding: Challenges for policy and practice* (pp. 240-275). San Francisco: Jossey-Bass.

Cohen, D. K., & Hill, H. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record, 102*(2), 294-343.

Cohen, D. K., & Hill, H. (2001). *Learning policy: When state education reform works.* New Haven: Yale University Press.

Cohen, D. K., Raudenbush, S. W., & Loewenberg Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis, 25*(2), 119-142.

Cohen, D. K., & Spillane, J. (1993). Policy and practice: The relationship between governance and instruction. In S. H. Fuhrman (Ed.), *Designing coherent educational policy: Improving the system* (pp. 35-88). San Francisco: Jossey Bass.

Cohen, M. D., & March, J. G. (1974). *Leadership and ambiguity: The American college president.* New York: McGraw-Hill.

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly, 17*(1), 1-25.

Creighton, T. B. (2001a, April). Data Analysis in Administrators' Hands: An Oxymoron? *The School Administrator, 58,* 6-11.

Creighton, T. B. (2001b). *Schools and data: The educator's guide for using data to improve decision making.* Thousand Oaks, California: Corwin Press, Inc.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage.

Creswell, J. W. (2002). *Research design: Qualitative, quantitative, and mixed method approaches.* Thousand Oaks, CA: Sage.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920.* New York: Teachers College Press.

Cuban, L. (1990). Reforming again, again, and again. *Educational researcher, 19*(1), 3-13.

Cusick, P. (1983). *The egalitarian ideal and the American high school: Studies of three schools.* New York: Longman.

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm.* New York: Blackwell.

Daft, R. L., & Weick, K. E. (1984). Toward a model of organizations as interpretive systems. *The Academy of Management Review, 9*(2), 284-295.

Darling-Hammond, L. (2003). Standards and assessments: Where we are and what we need. *Teachers College Record,* Retrieved from http://www.tcrecord.org/Content.asp?ContentID=11109.

Darling-Hammond, L., & Sykes, G. (2003). Wanted: A national teacher supply policy for education: The right way to meet the "highly qualified teacher" challenge. *Educational Policy Analysis Archives, 11*(33), Retrieved from http://epaa.asu.edu/epaa/v11n33/.

Davenport, T. H., DeLong, D. W., & Beers, M. C. (1998). Successful knowledge management projects. *Sloan Management Review, 39*(2), 43-57.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge managing what your organization knows.* Boston: Harvard Business School Press.

DeBray, E., Parson, G., & Avila, S. (2003). Internal alignment and external pressure. In M. Carnoy, R. Elmore & L. S. Siskin (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 55-86). New York, NY: RoutledgeFalmer.

DeBray, E., Parson, G., & Woodworth, K. (2001). Patterns of response in four high schools under state accountability policies in Vermont and New York. In S. H. Fuhrman (Ed.), *From the capital to the classroom: Standards-based reform in the states* (pp. 170-192). Chicago, IL: University of Chicago.

DiMaggio, P. J., & Powell, W. W. (1991a). Introduction. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 1-40). Chicago: University of Chicago Press.

DiMaggio, P. J., & Powell, W. W. (1991b). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 63-82). Chicago, IL: University of Chicago Press.

Dobbs, M. (2003). *Education "miracle" has a math problem.* Retrieved November 10, 2003, from www.washingtonpost.com/ac2/wp-dyn/admin/

Doherty, K. M. (2001). Poll: Teachers support standards- with hesitation. *Education Week, 29*(17), 20.

Duke, D. L., Butin, D. W., & Sofka, A. (2003). The impact of Virginia's accountability plan on high school English departments. In D. L. Duke, M. Grogan, P. D. Tucker & W. F. Heinecke (Eds.), *Educational leadership in an age of accountability: The Virginia experience* (pp. 135-154). Albany, NY: State University of New York Press.

Duke, D. L., Grogan, M., & Tucker, P. D. (2003). Educational leadership in an age of accountability. In D. L. Duke, M. Grogan, P. D. Tucker & W. F. Heinecke (Eds.), *Educational leadership in an age of accountability: The Virginia experience* (pp. 198-214). Albany, NY: State University of New York Press.

Earl, L., & Katz, S. (2002). Leading schools in a data-rich world. In K. Liethwood & P. Hallinger (Eds.), *Second international handbook of educational leadership and administration, Part Two.* Dordrecht: Kluwer Academic Publishers.

Earl, L., & Torrance, N. (2000). Embedding accountability and improvement into large-scale assessment: What difference does it make? *Peabody Journal of Education, 75*(4), 114-141.

Easton, J. Q., & Luppescu, S. (2004). *Teacher and principal responses to the Grow Network: A preliminary evaluation of use in Chicago Public Elementary Schools.* Chicago, IL: Consortium on Chicago School Research (CCSR).

ECS. (2004). *ECS report to the nation: State implementation of the No Child Left Behind Act.* Denver, CO: Education Commission of the States.

Elmore, R. (1983). Complexity and control: What legislators and administrators can do about implementing public policy. In L. S. Shulman & G. Sykes (Eds.), *Handbook of teaching and educational policy* (pp. 342-369). New York: Longman Inc.

Elmore, R. (2000). *Building a new structure for school leadership.* Washington, D.C.: Albert Shanker Institute.

Elmore, R. (2002). *The limits of "change".* Retrieved February 2002, from http://www.edletter.org/past/issues/2002-jf/limitsofchange.shtml

Elmore, R. (2003). Accountability and capacity. In M. Carnoy, R. Elmore & L. S. Siskin (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 195-209). New York, NY: RoutledgeFalmer.

Elmore, R. (2004). The problem of stakes in performance-based accountability systems. In S. H. Fuhrman & R. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 274-296). New York, NY: Teachers College Press.

Elmore, R., Abelmann, C., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65-98). Washington, DC: Brookings Institution.

Elmore, R., & Burney, D. (1999). Investing in teacher learning: Staff development and instructional improvement. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession* (pp. 263-291). San Francisco: Jossey-Bass.

Elmore, R., & McLaughlin, M. W. (1988). *Steady work: Policy, practice and the reform of American education.* Santa Monica, CA: RAND.

Elmore, R., Peterson, P. L., & McCarthy, S. (1996). *Restructuring in the classroom: Teaching, learning and school organization.* San Francisco: Jossey-Bass.

English, F. (2002). Cutting the Gordian Knot of educational administration: the theory-practice gap. *UCEA Review, XLIV*(1), 1-3.

eSchool News staff. (2003, October 2003). School officials punished for false reporting of data. *eSchool News, 6,* 10.

Farkas, S., Johnson, J., & Duffet, A. (2003). *Rolling up their sleeves: Superintendents and principals talk about what's needed to fix public schools.* New York, NY: Public Agenda.

Feldman, M. S., & March, J. G. (1988). Information in organizations as signal and symbol. In J. G. March (Ed.), *Decisions and organizations* (pp. 409-428). New York: Basil Blackwell.

Firestone, W. (1989). Using reform: Conceptualizing districts initiatives. *Educational Evaluation and Policy Analysis, 11*(2), 151-164.

Firestone, W., & Herriott, R. (1982). Two images of schools as organizations: An explication and illustrative empirical text. *Educational Administration Quarterly, XVIII*(2), 39-59.

Firestone, W., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis, 20*(2), 95-113.

Fisher, C., Dwyer, D. C., & Yocam, K. (Eds.). (1996). *Education and technology: Reflections on computing in classrooms.* San Francisco: Jossey Bass.

Fuhrman, S. H. (1999). *The new accountability* (No. Policy Brief 27). Philadelphia, PA: Consortium for Policy Research in Education (CPRE).

Fuhrman, S. H. (2004). Introduction. In S. H. Fuhrman & R. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 3-14). New York, NY: Teachers College Press.

Fullan, M. (2001). *Leading in a culture of change.* San Francisco: Jossey-Bass.

Fuller, B. (2004, February 1). Only the politicking gets an 'A'. *Washington Post,* p. 4.

Fusarelli, L. D. (2004). The potential impact of the No Child Left Behind Act on equity and diversity in American education. *Educational Policy, 18*(1), 71-94.

Gayler, K., Chudowsky, N., Kober, N., & Hamilton, M. (2003). *State high school exit exams put to the test.* Washington, DC: Center on Education Policy (CEP).

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In C. Geertz (Ed.), *The interpretation of cultures: Selected essays* (pp. 3-30). New York: Basic Books.

Gibbs, G. R. (2002). *Qualitative data analysis: Explorations with NVivo.* Philadelphia, PA: Open University Press.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

Gleick, J. (2000). *Faster.* New York: Vintage.

Goertz, M. E. (2001). Standards-based accountability: Horse trade or horse whip? In S. H. Fuhrman (Ed.), *One Hundredth Yearbook of the National Society for the Study of Education, Part II: From the capitol to the classroom: Standards-based reform in the states*. Chicago, IL: University of Chicago Press.

Goertz, M. E., Duffy, M. C., & with LeFloch, K. C. (2001). *Assessment and accountability systems in the 50 states: 1999-2000* (No. RR-046). Philadelphia: Consortium fo Policy Research in Education, University of Pennsylvania.

Goldring, E. B. (1995). Striking a balance: Boundary spanning and environmental management in schools. In S. B. Bacharach & B. Mundell (Eds.), *Images of schools: Structures and roles in organizational behavior*. Thousand Oaks, CA: Sage.

Gootman, E. (2004, June 4). 10,000 3rd graders to be held back in NYC. *New York Times*, p. 1.

Green, J. P., Winters, M. A., & Forster, G. (2003). *Testing high stakes tests: Can we believe the results of accountability tests*. New York, NY: Manhattan Institute for Policy Research.

Grogan, M., & Roland, P. (2003). A study of successful teachers preparing high school students for the standards of learning tests in Virginia. In D. L. Duke, M. Grogan, P. D. Tucker & W. F. Heinecke (Eds.), *Educational leadership in an age of accountability: The Virginia experience* (pp. 114-134). Albany, NY: State University of New York Press.

Habermas, J. (1984). *The theory of communicative action* (T. McCarthy, Trans. Vol. 2). Boston, MA: Beacon Press.

Hallinger, P. (1992). The evolving role of American principals: From managerial to instructional to transformational leaders. *Journal of Educational Administration, 30*(3), 35-48.

Haney, W. (2000). The myth of the Texas miracle in education. *Educational Policy Analysis Archives, 8*(41), Retrieved from http://epaa.asu.edu/epaa/v8n41/.

Haney, W. (2002). Lake Woebeguaranteed: Misuse of test scores in Massachusetts, Part I. *Education Policy Analysis Archives, 10*(24), Retrieved from http://epaa.asu.edu/epaa/v10n24/.

Hannaway, J., & Kimball, K. (2001). Big isn't always bad: School district size, poverty, and standards-based reform. In S. H. Fuhrman (Ed.), *From the capitol to the*

classroom: Standards-based reform in the states: One hundredth yearbook of the National Society for the Study of Education, Part II. Chicago, IL: University of Chicago Press.

Hargreaves, A., Giles, C., & Foote, M. (2004, April). *The sustainability of innovative schools as learning organizations and professional learning communities during standardized reform.* Paper presented at the American Educational Research Association, San Diego, CA.

Heinecke, W. F., Curry-Corcoran, D. E., & Moon, T. R. (2003). U.S. schools and the new standards and accountability initiative. In D. L. Duke, M. Grogan, P. D. Tucker & W. F. Heinecke (Eds.), *Educational leadership in an age of accountability: The Virginia experience* (pp. 7-35). Albany, NY: State University of New York Press.

Herman, J. L. (2004). The effects of testing on instruction. In S. H. Fuhrman & R. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 141-166). New York, NY: Teachers College Press.

Hess, F. (2003). The case for being mean. *Educational Leadership, 61*(3).

Heubert, J. P. (2004). High-stakes testing in a changing environment: Disparate impact, opportunity to learn, and current legal protections. In S. H. Fuhrman & R. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 220-244). New York, NY: Teachers College Press.

Hightower, A. (2002). San Diego's big boom: Systemic instructional change in the central office and schools. In A. Hightower, M. S. Knapp, J. A. Marsh & M. W. McLaughlin (Eds.), *School districts and instructional renewal* (pp. 76-93). New York, NY: Teachers College Press.

Holcomb, E. L. (2004). *Getting excited about data: Combining people, passion, and proof to maximize student achievement* (Second ed.). Thousand Oaks: Corwin.

Homans, G. (1950). *The human group.* New York: Harcourt, Brace and Co.

Hoy, W., & Miskel, C., G. (1991). *Educational administration: Theory, research, and practice* (4th ed.). New York, NY: McGraw-Hill.

Jones, B. D., & Egley, R. J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education Policy Analysis Archives, 12*(39), Retrieved from http://epaa.asu.edu/epaa/v12n39/.

Keeney, L. (1998). *Using data for school improvement.* Rovidence, RI: Annenberg Institute for School Reform.

Kim, J., & Sunderman, G. L. (2004). *Large mandates and limited resources: State response to the No Child Left Behind Act and implications for accountability.* Cambridge, MA: The Civil Rights Project, Harvard University.

Klein, S. P. (2004). Response to Skrla et al.: Is there a connection between educational equity and accountability? In L. Skrla & J. J. Scheurich (Eds.), *Educational equity and accountability: Paradigms, policies, and politics* (pp. 91-98). New York, NY: RoutledgeFlamer.

Klein, S. P., Hamilton, L. S., McGaffrey, D. F., & Stecher, B. (2000). *What do test scores in Texas tell us?* Washington, DC: RAND.

Kohn, A. (1993). *Punishment by rewards.*

Kornhaber, M. L. (2004). Appropriate and inappropriate forms of testing, assessment, and accountability. *Educational Policy, 18*(1), 45-70.

Labaree, D. (1997). *How to succeed in school without really learning.* New Haven, CT: Yale University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.*

Learning First Alliance. (2003). *Beyond islands of excellence: What districts can do to improve instruction and achievement in all schools- a leadership brief.* Washington, DC: Association for Supervision and Curriculum Development.

Leithwood, K., & Prestine, N. (2002). Unpacking challenges of leadership at the school and district level. In J. Murphy (Ed.), *The educational leadership challenge: Redefining leadership for the 21st century: One hundred-first yearbook of the National Society for the Study of Education* (pp. 42-64). Chicago, IL: University of Chicago Press.

Lemons, R., Luschei, T., & Siskin, L. S. (2003). Leadership and the demands of standards-based accountability. In M. Carnoy, R. Elmore & L. S. Siskin (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 99-128). New York, NY: RoutledgeFalmer.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage.

Linn, R. K. (2003). Accountability: Responsibility and reasonable expectations. *Educational Researcher, 32*(7), 3-13.

Linn, R. K., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher, 31*(6), 3-16.

Lofland, J. (1971). *Analyzing social settings: A guide to qualitative observation and analysis.* Belmont, CA: Wadsworth.

Lortie, D. (1975). *Schoolteacher.* Chicago: University of Chicago Press.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States.* Mahwah, NJ: Lawrence Erbaum Associates.

MacDonald, C., & Feighan, M. (2004, February 3). Schools give selves A's, overt failure. *The Detroit News.*

Madaus, G., & Clarke, M. (2001). The impact of high-stakes testing on minority students. In G. Orfield & M. L. Kornhaber (Eds.), *Raising standards or raising barriers: Inequality and high stakes testing in public education* (pp. 85-106). New York, NY: The Century Foundation.

Majone, G. (1989). *Evidence, argument, and persuasion in the policy process.* New Haven: Yale University Press.

Malen, B. (2003). Tightening the grip? The impact of state activism on local school systems. *Education Policy, 17*(2), 195-216.

March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics, 9*(2), 587-608.

March, J. G. (1988). Ambiguity and accounting: The elusive link between information and decision-making. In J. G. March (Ed.), *Decisions and organizations* (pp. 384-408). New York: Basil Blackwell.

March, J. G. (1997). Understanding how decisions happen in organizations. In Z. Shapira (Ed.), *Organizational decision making* (pp. 9-31). Cambridge, MA: Cambridge University Press.

March, J. G. (1999a). Exploration and exploitation in organizational learning. In J. G. March (Ed.), *The pursuit of organizational intelligence* (pp. 114-136). Malden, MA: Blackwell.

March, J. G. (1999b). Introduction. In J. G. March (Ed.), *The pursuit of organizational intelligence* (pp. 1-10). Malden, MA: Blackwell.

March, J. G. (1999c). Understanding how decisions happen in organizations. In J. G. March (Ed.), *The pursuit of organizational intelligence* (pp. 13-38). Malden, MA: Blackwell.

March, J. G., & Levinthal, D. (1999). The myopia of learning. In J. G. March (Ed.), *The pursuit of organizational learning* (pp. 193-222). Malden, MA: Blackwell.

March, J. G., & Olsen, J. P. (1976). *Ambiguity and choice in organizations*. Bergen, Norway: Universitetsforlaget.

March, J. G., & Olsen, J. P. (1988). The uncertainty of the past: Organizational learning under ambiguity. In J. G. March (Ed.), *Decisions and organizations* (pp. 335-358). New York, NY: Basil Blackwell.

March, J. G., & Sevon, G. (1988). Gossip, information and decision-making. In J. G. March (Ed.), *Decisions and organizations* (pp. 429-442). New York: Basil Blackwell.

March, J. G., & Simon, H. (1958). *Organizations*. New York: Wiley.

Marris, P. (1974). *Loss and change*. New York: Pantheon Books.

Marshall, C., & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

Mason, S. (2002). *Turning data into knowledge: Lessons from six Milwaukee Public Schools* (No. Working Paper 2002-3). Madison, WI: Center for Education Research.

Massell, D. (2001). The theory and practice of using data to build capacity: State and local strategies and their effects. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states: One hundredth yearbook of the National Society for the Study of Education, Part II* (pp. 148-169). Chicago, IL: The University of Chicago Press.

Massell, D., & Goertz, M. E. (2002). District strategies for building instructional capacity. In A. M. Hightower, M. S. Knapp, J. A. Marsh & M. W. McLaughlin (Eds.), *School districts and instructional renewal.* New York: Teachers College Press.

McDonnell, L. (2004). *Politics, persuasion, and educational testing.* Cambridge, MA: Harvard University Press.

McDonnell, L., & Elmore, R. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis, 9*(2), 133-152.

McLaughlin, M. W. (1990). The Rand change agent study: Ten years later. In A. Odden (Ed.), *Implementation* (pp. 143-155). New York: State University of New York Press.

McLaughlin, M. W., & Talbert, J. E. (2002). Reforming districts. In A. Hightower, M. S. Knapp, J. A. Marsh & M. W. McLaughlin (Eds.), *School districts and instructional renewal* (pp. 173-192). New York, NY: Teachers College Press.

McNeil, L., & Valenzuela, A. (2001). The harmful impact of TAAS system of testing in Texas. In G. Orfield & M. L. Kornhaber (Eds.), *Raising standards or raising barriers? Inequality and high stakes testing in public education.* New York, NY: Century.

McNeil, L. M. (2000). *Contradictions of reform: The educational costs of standardized testing.* New York, NY: Routledge.

Merchant, B. M. (2004, April). *Tick, TAKS, toe-- Fail and stay, pass and go: 3rd grade accountability, Texas style.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Merriam, S. B. (1988). *Case study research in education: A qualitative approach.* San Francisco: Jossey-Bass.

Meyer, J. W., & Rowan, B. (1991). Institutional organizations: Formal structure as myth and ceremony. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 41-82). Chicago: University of Chicago Press.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Beverly Hills, CA: Sage.

Mintzberg, H. (1973). *The nature of managerial work.* New York: Harper & Row.

Moustakas, C. (1994). *Phenomenological research methods.* Thousand Oaks, CA: Sage.

Nardi, B., & O'Day, V. (1999). *Information ecologies: Using technology with heart.* Cambridge, MA: MIT Press.

NCREL. (2004). Using data to guide school improvement. *Notes and Reflections*(7), 1-15.

Neill, M., Guisbond, L., & Schaeffer, B. (2004). *Failing our children: How No Child Left Behind undermines quality and equity in education.* Cambridge, MA: Fair Test: The National Center for Fair and Open Testing.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation.* New York: Oxford University Press.

O'Day, J. (1996). Incentives and school improvement. In S. H. Fuhrman & J. O'Day (Eds.), *Rewards and reform: Creating educational incentives that work.* San Francisco, CA: Jossey-Bass.

O'Day, J. (2002). Complexity, accountability and school improvement. *Harvard Educational Review, 72*(3), 293-329.

O'Day, J., Goertz, M. E., & Floden, R. E. (1995). *Building capacity for educational reform.* Philadelphia, PA: Consortium for Policy Research in Education.

O'Day, J., & Smith, M. A. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent educational policy: Improving the system.* San Francisco: Jossey-Bass.

Ogawa, R. T., & Collom, E. (2000). Using performance indicators to hold schools accountable: Implicit assumptions and inherent tensions. *Peabody Journal of Education, 75*(4), 200-215.

Ogawa, R. T., Sandholtz, J., Martinez-Flores, M., & Scribner, S. P. (2003). The substantive and symbolic consequences of a district's standards-based curriculum. *American Educational Research Journal, 40*(1), 147-176.

Olson, L. (2004, May 26). User-friendly reports on student test scores help guide instruction. *Education Week, 23,* 9.

Orfield, G. (2004a). Forward. In G. L. Sunderman & J. Kim (Eds.), *Inspiring vision, disappointing results: Four studies on implementing the No Child Left Behind Act.* Cambridge, MA: The Civil Rights Project, Harvard University.

Orfield, G. (2004b). Forward. In G. L. Sunderman & J. Kim (Eds.), *Expansion of federal power in American education.* Cambridge, MA: The Civil Rights Project, Harvard University.

Orlikowski, W. J. (2000). Using technology and constituting structures: A practice lens for studying technology in organizations. *Organizational Science, 11*(4), 404-428.

Paige, R. (2001). *Testimony on No Child Left Behind before the House Committee on Education and the Workforce.* Retrieved March 7, 2004, from http://www.ed.gov/speeches/03-2001/010307.html

Parker, L. (2004). Statewide assessment triggers urban school reform: But how high the stakes for urban minorities? In L. Skrla & J. J. Scheurich (Eds.), *Educational equity and accountability: Paradigms, policies, and politics* (pp. 183-188). New York, NY: RoutledgeFlamer.

Patton, M. Q. (1990). *Qualitative evaluation and research methods.* Newbury Park, CA: Sage.

Pedulla, J. (2003). State-mandated testing- What do teachers think? *Educational Leadership, 61*(3), 42-46.

Petrides, L. A., & Guiney, S. Z. (2002). Knowledge management for school leaders: An ecological framework for thinking schools. *Teachers College Record, 104*(8), 1702-1717.

Petrides, L. A., & Nodine, T. R. (2003). *Knowledge management in education: Defining the landscape.* Half Moon Bay, CA: The Institute for the Study of Knowledge Management in Education.

Piaget, J. (1975). *The equilibrium of cognitive structures: The central problem of intellectual development* (T. Brown & K. J. Thampy, Trans.). Chicago, IL: University of Chicago Press.

Plank, D. N. (1987). Why school reform doesn't change schools: political and organizational perspectives. In W. L. Boyd & C. T. Kerchner (Eds.), *The politics of excellence and choice in education: 1987 yearbook of the Politics of Education Association* (pp. 143-152). New York: The Falmer Press.

Plank, D. N., & Sykes, G. (Eds.). (2003). *Choosing choice: School choice in international perspective*. New York, NY: Teachers College Press.

Popham, W. J. (2000). Should large-scale assessments be used for accountability? Answer: Depends on the assessment silly! *Journal of Educational Change, 1*, 283-289.

Popham, W. J. (2001). *The truth about testing: An educator's call to action*. Alexandria, VA: Association for Supervision and Curriculum Development.

Popkewitz, T. S. (1991). *A political sociology of educational reform: Power/knowledge in teaching, teacher education and research*. New York: Teachers College Press.

Popkewitz, T. S. (2004). The alchemy of the mathematics curriculum: Inscriptions and the fabrication of the child. *American Educational Research Journal, 41*(1), 3-34.

Rait, E. (1995). Against the current: Organizational learning in schools. In S. B. Bacharach & B. Mundell (Eds.), *Images of schools: Structures and roles in organizational behavior* (pp. 71-107). Thousand Oaks, CA: Corwin Press.

Resnick, L. B., & Glennan, T. K. (2002). Leadership for learning: A theory of action for urban school districts. In A. Hightower, M. S. Knapp, J. A. Marsh & M. W. McLaughlin (Eds.), *Schools districts and instructional renewal* (pp. 160-172). New York, NY: Teachers College Press.

Roderick, M., & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low-achieving students to high stakes testing. *Educational Evaluation and Policy Analysis, 23*(3), 197-222.

Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.

Rothman, R. (2004). Benchmarking and alignment of state standards and assessments. In S. H. Fuhrman & R. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 96-114). New York, NY: Teachers College Press.

Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. In C. B. Cazden (Ed.), *Review of Research in Education* (Vol. 16, pp. 353-389). Washington, DC: American Educational Research Association.

Rowan, B. (1995). School organizations by design: The organizational design of schools. In B. Mundell (Ed.), *Images of schools: Structures and roles in organizational behavior* (pp. 11-42). Thousand Oaks, CA: Sage.

Rowan, B., & Miskel, C., G. (1999). Institutional theory and the study of educational organizations. In J. Murphy & K. S. Louis (Eds.), *Handbook of research on educational administration* (Second ed., pp. 359-384). San Francisco, CA: Jossey-Bass.

Rudestam, K. E., & Newton, R. R. (2001). *Surviving your dissertation: A comprehensive guide to content and process.* Thousand Oaks, CA: Sage.

Scheurich, J. J., Skrla, L., & Johnson, J. F. (2004). Thinking carefully about equity and accountability. In L. Skrla & J. J. Scheurich (Eds.), *Educational equity and accountability: Paradigms, policies, and politics* (pp. 13-28). New York, NY: RoutledgeFalmer.

Schoenfeld, A. H. (1999). Looking toward the 21st century: Challenges of educational theory and practice. *Educational Researcher, 28*(7), 4-14.

Scott, W. R. (1992). *Organizations: Rational, natural, and open systems* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Sergiovanni, T. (2000). *The lifeworld of leadership: Creating culture, community and personal meaning in our schools.* San Francisco: Jossey-Bass.

Shulman, L. S. (1983). Autonomy and obligation: The remote control of teaching. In L. S. Shulman & G. Sykes (Eds.), *Handbook of teaching and educational policy* (pp. 484-504). New York: Longman, Inc.

Simon, H. (1957). *Administrative behavior.* New York: Free Press.

Sipple, J., Killeen, K., & Monk, D. (2004). Adoption and adaptation: School district response to state imposed learning and graduation requirements. *Educational Evaluation and Policy Analysis, 26*(2), 143-168.

Siskin, L. S. (2003). Outside the core: Accountability in tested and untested subjects. In M. Carnoy, R. Elmore & L. S. Siskin (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 87-98). New York, NY: RoutledgeFalmer.

Siskin, L. S. (2004). The challenge of the high schools. In S. H. Fuhrman & R. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 167-188). New York, NY: Teachers College Press.

Skrla, L., & Scheurich, J. J. (2004). Displacing deficit thinking in school district leadership. In L. Skrla & J. J. Scheurich (Eds.), *Educational equity and*

accountability: Paradigms, policies, and politics (pp. 109-132). New York, NY: RoutledgeFalmer.

Skrla, L., Scheurich, J. J., Johnson, J. F., & Koschoreck, J. W. (2004). Accountability for equity: Can state policy leverage social justice? In L. Skrla & J. J. Scheurich (Eds.), *Educational equity and accountability: Paradigms, policies, and politics* (pp. 51-78). New York, NY: RoutledgeFalmer.

Smith, J. B., Lee, V. E., & Newmann, F. (2001). *Instruction and achievement in Chicago elementary schools.* Chicago, IL: Chicago Consortium on School Research (CCSR).

Spillane, J. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zone of enactment. *Journal of Curriculum Studies, 31*, 143-175.

Spillane, J. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition and Instruction, 18*(2), 141-179.

Spillane, J. (2002). Local theories of teacher change: The pedagogy of district policies and programs. *Teachers College Record, 104*(3), 377-420.

Spillane, J. (2004). *Standards deviation: How schools misunderstand educational policy.* Cambridge, MA: Harvard University Press.

Spillane, J., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher, 30*(3), 23-28.

Spillane, J., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research, 72*(3), 387-431.

Spillane, J., & Seashore Louis, K. (2002). School improvement processes and practices: Professional learning for building instructional capacity. In J. Murphy (Ed.), *The educational leadership challenge: Redefining leadership for the 21st century* (pp. 83-104). Chicago: University of Chicago Press.

Spillane, J., & Thompson, C. L. (1997). Restructuring conceptions of local capacity: The local education agency's capacity for ambitious instructional reform. *Education Evaluation and Policy Analysis, 19*(2), 185-203.

Sproull, L. S., & Zubrow, D. (1981). Performance information in school systems: Perspectives from organizational theory. *Educational Administration Quarterly, 17*(3), 61-79.

Stake, R. (1994). *The art of case study research.* Thousand Oaks, CA: Sage.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury Park, CA: Sage.

Sunderman, G. L., & Kim, J. (2004a). *Expansion of federal power in American education: Federal-state relationships under the No Child Left Behind Act, year one.* Cambridge, MA: The Civil Rights Project, Harvard University.

Sunderman, G. L., & Kim, J. (2004b). *Increasing bureaucracy or increasing opportunities? School district experience with supplemental educational services.* Cambridge, MA: The Civil Rights Project, Harvard University.

Supovitz, J. A., & Taylor, B. (2003). *The impact of standards-based reform in Duval County, Florida, 1999-2002.* Philadelphia, PA: Consortium for Policy Research in Education (CPRE).

Sykes, G. (1997). On trial: the Dallas value-based accountability system. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 110-119). Thousand Oaks, CA: Corwin Press.

Sykes, G. (1999). Teacher and student learning: Strengthening their connection. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 151-179). San Francisco: Jossey-Bass.

Sykes, G., & Elmore, R. (1988). Making schools manageable: Policy and administration for tomorrow's schools. In J. Hannaway & R. Crowson (Eds.), *The politics of reforming school administration: The 1988 yearbook of the Politics of Education Association* (pp. 77-94). New York: Falmer Press.

Sykes, G., Thompson, C., & Skrla, L. (under review). The district role in improving student learning: A research agenda. *Educational Administration Quarterly.*

The White House. (2001). *No Child Left Behind Act of 2001.* Washington, DC: United States Department of Education.

Thomas, R. M. (2003). *Blending qualitative and quantitative research methods in theses and dissertations*: Corwin.

Thompson, J. (1967). *Organizations in action.* New York: McGraw Hill.

Thorn, C. A. (2001). Knowledge for educational information systems: What is the state of the field. *Educational Policy Analysis Archives, 9*(47), Retrieved from http://epaa.asu.edu/epaa/v9n47/.

Tucker, P. D. (2003). The principalship: Renewed call for instructional leadership. In D. L. Duke, M. Grogan, P. D. Tucker & W. F. Heinecke (Eds.), *Educational leadership in an age of accountability: The Virginia experience* (pp. 97-113). Albany, NY: State University of New York Press.

U.S. Department of Education. (2002). *Strategic plan for 2002-2007.* Retrieved February 4, 2004, from http://www.ed.gov/pubs/stratplan2002-07/index.html

U.S. Department of Education. (2003). *The facts about No Child Left Behind: Getting results.* Retrieved February 4, 2004, from http://www.ed.gov/nclb/accountability/results/gettingresults.html

Valencia, R. R., Valencia, A., Sloan, K., & Foley, D. E. (2004). Let's treat the cause, not the symptoms: Equity and accountability in Texas revisited. In L. Skrla & J. J. Scheurich (Eds.), *Educational equity and accountability: Paradigms, policies, and politics* (pp. 29-38). New York, NY: RoutledgeFalmer.

Vaughan, D. (1996). *the challenger launch decision: Risky technology, culture and deviance at NASA.* Chicago: University of Chicago Press.

Voke, H. (2002). What do we know about sanctions and rewards? *ASCD Info Brief*(31).

Wayman, J., Stringfield, S., & Yakimowski, M. (2004). *Software enabling school improvement through analysis of student data* (No. 67). Baltimore, MD: Center for Research on the Education of Students Placed at Risk (CRESPAR).

Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Sciences Quarterly, 21*, 1-19.

Weick, K. (1979). *The social psychology of organizing.* New York, NY: Random House.

Weick, K. (1995). *Sensemaking in organizations.* Thousand Oaks, CA: Sage.

Weiss, C. (1995). The four I's of school reform: How interests, ideology, information and institution affect teachers and principals. *Harvard Educational Review, 65*(4), 571-592.

Western States Benchmarking Consortium. (1997). *Improvement stages.* Retrieved October 26, 2003, from www.wsbenchmark.org/westernstates/exemplars.htm

Wilson, S. M., & Floden, R. E. (2001). Hedging bets: Standards-based reform in classrooms. In S. H. Fuhrman (Ed.), *One-Hundredth Yearbook of the National Society for the Study of Education: Part II: From the capitol to the classroom: Standards-based reform in the states* (pp. 193-216). Chicago: University of Chicago Press.

Wolcott, H. F. (1990). *Writing up qualitative research.* Newbury Park, CA: Sage.

Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation.* Thousand Oaks, CA: Sage.

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage.

Zucker, L. G. (1991). The role of institutionalization in cultural persistence. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 83-107). Chicago: University of Chicago Press.