THE IMPACT OF ON-GOING ASSESSMENT DATA ON TEACHER PRACTICES

By

Stephanie McKinney Williams

A Thesis Presented to

The Faculty of Humboldt State University

In Partial Fulfillment of the Requirements for the Degree

Master of Arts in Education

Committee Membership

Ann Diver-Stamnes, Professor Emerita, Committee Chair

David A. Ellerd, PhD, Associate Professor, Committee Member

Eric Van Duzer, Associate Professor, Graduate Coordinator

December 2013
ABSTRACT

THE IMPACT OF ON-GOING ASSESSMENT DATA ON TEACHER PRACTICES

Stephanie McKinney Williams

With the rising demands for improvement and accountability required under the No Child Left Behind Act, schools are beginning to see the value of making data-based decisions to dramatically increase student achievement on high-stakes tests. Unfortunately, the information from these end-of-the-year tests often arrives too late to afford teachers the time to analyze the resultant data and make necessary curricular adjustments, both for specific students and groups as a whole. In an effort to make annual yearly progress, many educational agencies are starting to make data-driven decisions to help struggling students make the necessary academic gains to be successful on high stakes tests. In order to assess the Common Core State Standards and ensure that students are meeting these benchmarks, some schools utilize formative assessments called Measures of Academic Progress (MAP), designed by Northwest Evaluation Association. The disaggregated assessment data provided by Measurement of Academic Progress highlight the strengths and weaknesses of every child and provide teachers with the necessary information to target instruction. The aim of this study was to reveal if and how teachers at a highly successful school used assessment data to make instructional decisions, and what impact, if any, this had on the school’s on-going high performing status.
ACKNOWLEDGEMENTS

At different times in our lives, we all dream the unimaginable. Making our dreams a reality is sometimes the result of help from special people in our lives. I would like to thank the wonderful people in my life that helped me obtain my Master of Art in Education Degree.

Ann Diver-Stamnes, my committee chair, helped me accomplish this overwhelming endeavor. Not only did she provide valuable help in the writing process, she consistently inspired me through the rough times. Her expertise, passion, and dedication were truly motivating.

I am also thankful to Eric Van Duzer, my professor and graduate coordinator. Eric not only pushed me academically but helped me get through a tough semester while dealing with my daughter’s unfortunate accident. His compassion was very heart warming.

I would also like to thank David Ellerd for serving on my committee. David was my qualitative professor who provided assistance with the design of this thesis. His expertise was certainly appreciated.

I would also like to thank the wonderful teachers I work with. Not only did they take time from their busy lives to participate in my study, they consistently offered encouragement and support. I am honored to work with such dedicated professionals.
Doug White is also someone who helped me pursue this accomplishment. For years he has been my mentor and during this venture his love and support was readily available. Thank for being my number one cheerleader in good times and bad.

During this entire endeavor, my parents have been a huge source of inspiration. Thanks, Mom and Dad, for believing in me when I didn’t. Thanks for coming to Wyoming to help shuttle kids, cook dinners, and do whatever so I could focus all my time to write. I cannot express how much your help was appreciated. It was wonderful to have you stay for so long! I am blessed to have such wonderful parents.

Finally, I am extremely thankful to my family. The last year has been a rough one dealing with medical issues while attending graduate school. Despite these challenges, my husband and two children have worked together and supported me attain this master's. We have grown stronger since we are so far away from our California family. We will look back on this time and remember the sacrifices that helped solidify what really matters, and this is our healthy family.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CHAPTER 2: LITERATURE REVIEW</td>
<td>3</td>
</tr>
<tr>
<td>CHAPTER 3: METHODOLOGY</td>
<td>29</td>
</tr>
<tr>
<td>CHAPTER 4: RESULTS</td>
<td>40</td>
</tr>
<tr>
<td>CHAPTER 5: ANALYSIS</td>
<td>58</td>
</tr>
<tr>
<td>CHAPTER 6: CONCLUSIONS</td>
<td>74</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>78</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>80</td>
</tr>
</tbody>
</table>
CHAPTER ONE
INTRODUCTION

My involvement in education for the past twenty years has afforded me the opportunity to witness many changes. Currently, demands for improvement and accountability required under the No Child Left Behind Act are rising. Because of these demands, schools are beginning to see the value of making data-based decisions to dramatically increase student achievement on high-stakes tests. Unfortunately, the information from these end-of-the-year tests often arrives too late in the school year to impact learning.

The purpose of this observational case study was to answer the question: What is the impact of on-going assessment data on teacher practices and attitudes in a small rural elementary school? The aims of the research were to reveal if and how teachers use assessment data to make instructional decisions, and what impact, if any, this has on the school’s on-going high performing status. Information gleaned from this study could be helpful to other schools in an effort to achieve annual yearly progress as measured by the No Child Left Behind Act.

The literature review in Chapter Two provides an historical overview of educational assessment and accountability practices used in the United States. This background knowledge is useful in understanding the rationale for the No Child Left Behind Act that has had such a huge impact on states, schools, teachers, and students. The final section reviews the types of assessments currently utilized in schools for these
accountability purposes and highlights how successful schools have utilized data from
high stakes testing to impact student learning.

Chapter Three provides the methodology I used to gain insight from teachers in
grades three through five and the third through fifth grade Professional Learning
Community Coordinator from a top performing elementary school that has made annual
yearly progress (AYP) every year since No Child Left Behind Act was implemented in
2001. Using pre-determined questions and one-on-one interviews, I was able to identify
attitudes regarding testing and the ways in which these teachers use their test data to
increase student achievement.

Chapter Four provides an overview of the results of the research described in the
methodology chapter. The responses include demographic information as well as a
detailed summary of the open-ended interview questions describing how each teacher
utilizes results from Measures of Academic Progress (MAP) to increase academic
growth. Chapter Five offers an analysis of the data collected, broken down into the
following categories: setting, perspectives, and strategies.

In the Conclusion chapter, I offer my deductions to my central question,
limitations of the research, and implications for future research.
CHAPTER TWO
LITERATURE REVIEW

Introduction

The world is rapidly changing every day. With these different challenges come new demands on our education system to prepare students to function in the perplexing world in which they live. Education is the key to success in society and policy makers have turned to assessment to hold schools accountable for ensuring that all students are educated to high standards.

This literature review explores the ways in which data based decisions can help schools improve test scores. The first section provides a historical overview of educational assessment and accountability practices used in the United States. This background knowledge is useful when trying to understand the rationale for the No Child Left Behind Act since it has had such a huge impact on states, schools, teachers and students. No Child Left Behind has forced schools to initiate practices that will improve the overall student achievement levels and raise standardized test scores. The final section reviews the types of assessments currently utilized in schools for these accountability purposes.

Historical Overview of Assessment and Accountability Practices

Testing has always been a major part of public education, but how it has been used has changed with time (Ravitch, 2002). During the 19th century students who did not pass a particular grade remained in that grade for another year, and educators and others believed that the students themselves were at fault for being retained (Ravitch,
Few school children went onto high school, and even fewer onto college (Ravitch, 2002). Students who wanted to pursue a higher education were usually accepted unless they were applying to a prominent university (Ravitch, 2002). Students became frustrated because each institution had a different entrance exam (Ravitch, 2002). In an effort to standardize and help students prepare for these tests, educators created College Entrance Examinations (Ravitch, 2002).

The twentieth century brought a complete change as educators perceived subject matter, discipline and grades to be less important than children’s social adjustment (Ravitch, 2002). Educators tested children in schools, but they did not believe tests should be used for accountability (Ravitch, 2002). As a result, struggling students who were not meeting the required skills in a particular grade level were socially promoted to the next grade in order to keep them with their peers (Ravitch, 2002).

Accountability started to become a factor as policy makers began to notice school performance compared to the amount of resources the schools received (Coleman, 1966). In 1970, the National Assessment of Educational Progress (NAPE) was established in hopes of spotting trends and patterns in education for future study (Ravitch, 2002). By the 1980s, education was the biggest budget item for each state (Ravitch, 2002).

During the Reagan administration, United States students’ test results were far lower than other countries (Klein, Zevenbergen, & Brown, 2006). This prompted a perceived need for better instruction and assessment to hold schools accountable (Klein et al., 2006). In 2001, the No Child Left Behind Act (NCLB) was passed unanimously in the U.S. Congress (Caillier, 2007). This Act resulted in a huge shift in U.S. education
policy as test scores, rather than academic standards, became the norm for how schools increased proficiency as a means of accountability (Jennings, 2012).

*Rationale of No Child Left Behind.*

The No Child Left Behind Act of 2001 is the biggest reform to the educational system and is now referred to as standards-based accountability (National Conference of State Legislature, 2012). The major component of standards-based accountability requires states to set standards and then assess students to see if academic growth has been made (National Conference of State Legislature, 2012).

The primary goal of NCLB is for every child in America to be proficient in both math and reading by 2014 (Caillier, 2007). This Act of 2001 calls for more frequent testing by schools in an effort to reorganize their accountability system (Education Commission of the States, 2012). Each year, students starting in grade 3 must take a reading and math test in order to demonstrate proficiency (Caillier, 2007). The proficiency percentage of each school is used to determine if that local agency is making adequate yearly progress (AYP) so they will be on target for 100% proficiency in 2014 (Caillier, 2007). The basic idea of NCLB is to set small academic growth goals each year and then test to monitor progress (Wiener & Hall, 2004).

Education varies greatly throughout all 50 states and the quality of U. S. education has caused concern over the attainment of high academic achievement (National Conference of State Legislature, 2012). Factors such as race, ethnicity, economic status, geography, and parental education have proven to be areas of huge
disparities in terms of students’ test scores (Martin, 2011; National Conference of State Legislature, 2012).

*No Child Left Behind.*

Because of existing achievement gaps, NCLB requires that every school identify subgroups (Caillier, 2007). These further divide students into groups according to ethnicity, socio-economic status, English proficiency status, and inclusion in special education (Caillier, 2007). Under NCLB, no child is overlooked (Caillier, 2007). NCLB requires schools to show steady growth as a school each year and within each subgroup, as measured by adequate yearly progress, in order to close the achievement gap between advantaged and disadvantaged students (Education Commission of the States, 2012). By isolating the data into subgroups, schools are able to identify students who did not make gains and provide specific intervention as needed (Wiener & Hall, 2004).

Prior to NCLB, schools reported their test results as an average (Wiener & Hall, 2004). This type of accountability did nothing for the under-achievers and served only to hide them in the statistics, creating a bigger gap in achievement (Wiener & Hall, 2004). As part of NCLB, schools analyze the results in subgroups in an attempt to highlight achievement and to ensure high performance for all students, regardless of ethnicity, background, income, or other special circumstances (Jennings & Rentner, 2006).

Under NCLB, schools that fail to make adequate yearly progress for two consecutive years are labeled as failing and must allow students to transfer to a school of their choice (Caillier, 2007). If they do not make adequate yearly progress the third year, they must provide preapproved tutoring services free of charge (Caillier, 2007). On the
fifth year of not making adequate yearly progress, schools may be faced with a takeover and reconstruction of staff (Caillier, 2007).

Another key component of NCLB is the highly qualified teacher (Wiener & Hall, 2004). The impact of a good teacher or bad teacher can continue to impact children four years after they have left that classroom (Porter-Magee, 2004). Five consecutive years of good teachers can bring the lowest child’s test scores up significantly (Wiener & Hall, 2004). There is no substitute for a good teacher (Wiener & Hall, 2004). In an effort to improve test scores, NCLB requires lessons to be delivered by effective instructors within their discipline (Education Commission of the States, 2012).

To comply with NCLB, schools are held responsible to meeting standards and notifying parents the test results of their children, school, and district (National Conference of State Legislature, 2012). Many schools publish a report card in the paper highlighting proficiency percentages (Caillier, 2007).

Impact on states and schools.

NCLB has definitely had an impact on schools (Jennings & Rentner, 2006). As part of NCLB’s standard-based assessment, testing is more frequent and the format is different because it reports whether or not an individual correctly answered a question (Jennings & Rentner, 2006; Klein et al., 2006). Before NCLB, norm-referenced testing was administered which showed how the student compared to other students in the same age group (Jennings & Rentner, 2006; Klein et al., 2006). Norm-referenced tests were administered every three to four years in grades 4, 8, 11 and 12; currently testing is done annually (Klein et al., 2006).
Today teachers find that in order for their students to perform well on standardized tests, they must first teach them the test format and ways to take the test (Klein et al., 2006). Standard-based tests primarily evaluate math and reading so schools have begun the process of trying to align the curriculum in hopes of increasing scores (Jennings & Rentner, 2006). Unfortunately, in the process of preparing for standardized tests, some schools have started to reduce the amount of time spent teaching other subject areas (Jennings & Rentner, 2006). Teachers feel a great deal of anxiety as they explore ways to utilize test questions to promote higher order thinking in their classrooms (Klein et al., 2006). Many teachers express frustration because they spend most of their instruction time preparing their class for the end of the year test (Byrd-Blake et al., 2010; Klein et al., 2006).

Pulling subgroups out of the larger school population in order to isolate and identify strengths and weaknesses within each subgroup can either be viewed as a strength or weakness (Rose, 2004). To make adequate yearly progress, schools must show growth in every subgroup (Kim & Sunderman, 2005). The assumption of NCLB is that if adequate yearly progress increases each year, then the achievement gap is shrinking (Kim & Sunderman, 2005).

Since adequate yearly progress is determined by mean proficiency, it can be a disadvantage for schools and their subgroups (Kim & Sunderman, 2005). Most schools with high poverty have lower mean proficiency than schools with low poverty (Kim & Sunderman, 2005). A high correlation exists between race and poverty and low-test
scores so schools with high poverty and racially diverse populations have a difficult time making adequate yearly progress (Kim & Sunderman, 2005; Stiggins, 2005).

Following the NCLB model, students must demonstrate small academic gains each year to meet adequate yearly progress targets (Rose, 2004). Unfortunately, the proposed adequate yearly progress growth model is not always an uphill slope of improvement (Rose, 2004). When learning new concepts, students experience plateaus, setbacks, and sometimes slow growth periods that NCLB does not take into consideration if the end of the year target is not reached (Rose, 2004). NCLB does not take into account where schools start from, only the end result (Kim & Sunderman, 2005; Rose, 2004).

Under the NCLB model, the schools serving the highest need will fail first even if they do make gains each year because of the number of subgroups within the population (Rose, 2004). Students in subgroups usually have to work the hardest since they have such a long way to grow in order to be proficient (Kim & Sunderman, 2005; Rose, 2004). Following adequate yearly progress guidelines, the high poverty schools will not make the proficiency growth target even though they show consistent growth in both math and reading (Kim & Sunderman, 2005). However, many disadvantaged schools demonstrate growth by the National Assessment of Educational Progress (NAEP) standards (Rose, 2004). Further, making adequate yearly progress poses another challenge because for many high-poverty schools, these same low-income students are also in the minority subgroup (Kim & Sunderman, 2005). The more subgroups in a school, the greater is the chance of not making adequate yearly progress since it only takes one subgroup to fail for
the entire school to also fail (Kim & Sunderman, 2005; Rose, 2004). By pure probability smaller schools with fewer subgroups will have an easier time making adequate yearly progress each year (Rose, 2004). To complicate things even more, high poverty schools may meet all accreditation guidelines set forth by their state but fail to make adequate yearly progress (Kim & Sunderman, 2005). This double standard makes it confusing for teachers, students, and parents (Kim & Sunderman, 2005).

NCLB’s goal is to ensure that all subgroups are getting the same quality education and showing growth (Martin, 2011). Subgroups clearly decide if a school makes adequate yearly progress, and there are strong reasons why isolating these specific subgroups as part of adequate yearly progress and failing a school are unfair (Martin, 2011). Low socio-economic background and low-test scores are strongly correlated so showing growth every year is a difficult task for most schools that serve diverse populations (Kim & Sunderman, 2005; Martin, 2011). NCLB requires English Language Learners who have been enrolled for a year to take the achievement test (Martin, 2011). Standard-based assessments are geared for native speakers so schools with English Language Learners subgroups are being evaluated on English fluency and on average score lower (Martin, 2011). Students with a disability are defined as at least two years below grade level and are required to take the test as well (Martin, 2011). Accommodations to the test can be administered; however, due to the huge range and degree of disabilities, some would argue whether the accommodations really help (Martin, 2011). The reliability of the subgroups is questionable because test results vary so greatly from year to year (Martin, 2011).
Students who have the hardest time learning are pushed to make the fastest gains (Martin, 2011). Adequate yearly progress reports that are publically announced tell whether or not a school is proficient (Martin, 2011). These reports do not describe the growth made in each subgroup when the adequate yearly progress target was not reached (Martin, 2011). Some schools, because of their student population, are destined to fail if passing means reaching certain proficiency benchmarks in a specific time frame (Kim & Sunderman, 2005; Martin, 2011). Schools are either stamped as meeting adequate yearly progress or not, and in the view of the public this translates to passing or failing regardless of the degree of success attained by a school within its subgroups (Martin, 2011).

In an effort to bridge the gap between low-income students and other students, the U.S. Department of Education provides supplemental funding to local school districts to meet the needs of at-risk and low-income students (Porter-Magee, 2004). In order to receive federal Title I money under NCLB, schools have been asked to raise teacher standards (Porter-Magee, 2004). Schools must prove that the instructors within a classroom are highly qualified because good teachers have such an effect on their students (Porter-Magee, 2004). In order to be highly qualified, teachers must have a college degree and state certification, and they must demonstrate mastery of the subject they teach (Porter-Magee, 2004).

Determining whether or not a teacher is highly qualified is another area in which states have no clear guidelines under NCLB (Porter-Magee, 2004). To be highly qualified, teachers must have a bachelor’s degree (Porter-Magee, 2004). The second
component requires state certification (Porter-Magee, 2004). The process of how each state defines their certification requirement is different everywhere (Porter-Magee, 2004). The third requirement asks teachers to demonstrate mastery within a particular subject (Porter-Magee, 2004). This final requirement has been controversial (Porter-Magee, 2004). Authors of the NCLB feel content knowledge is far more beneficial than pedagogy (Porter-Magee, 2004).

The rationale for hiring highly qualified teachers is to provide excellent learning opportunities for students (Porter-Magee, 2004). Through a provision called High Objective Uniform State Standard of Evaluation (HOUSSE) states are able to decide if teachers demonstrate subject area mastery (Porter-Magee, 2004). Unfortunately, High Objective Uniform State Standard of Evaluation does not have set requirements defining those standards (Porter-Magee, 2004). Because of a lack of standards, many states have created their own definition of what it means for teachers to be considered highly qualified (Porter-Magee, 2004). This has resulted in a large numbers of highly qualified teachers who were not required to demonstrate subject area mastery (Porter-Magee, 2004). Allowing states to define their own rules for being highly qualified has had the unintended consequence of covering up struggling teachers (Porter-Magee, 2004). The intent of the highly qualified teacher was to establish teacher standards and to provide help to those teachers not meeting the standards with help, which would ultimately improve the educational outcomes for students. (Porter-Magee, 2004).

As part of NCLB, test scores must be made public (Klein et al., 2006). By reporting assessment results, schools within a district may become competitive as they try
to show the most gains (Klein et al., 2006). Teachers and school realize the importance of high stake testing and the impact the results have not only on their school but also on the housing market, school taxes, and an array of neighborhood issues (Klein et al., 2006; Triplett & Barksdale, 2005).

**Impact on students.**

NCLB’s goal is to improve education in order for all students to be academically successful (Wiener & Hall, 2004). To accomplish this task, NCLB requires setting academic goals and monitoring student achievement in the form of standardized testing (Wiener & Hall, 2004). Since the frequency of tests taken has increased, more importance is placed on test results (Sloane & Kelly, 2003). The term high-stakes has become synonymous for standardized tests since the scores often determine if a child is failing or passing (Sloane & Kelly, 2003).

In most classrooms across the nation, teachers begin preparing students for high-stakes testing when school begins and increase the amount of time as the test date approaches (Triplett & Barksdale, 2005). Typically students are given practice tests that mirror the format of what will be administered (Triplett & Barksdale, 2005). Students are instructed how to eliminate answers when given multiple choice questions and practice essays and short response items (Triplett & Barksdale, 2005). Some schools even hold assemblies to encourage effort and promise parties for the students if scores improve (Triplett & Barksdale, 2005).

Advocates of high-stake testing argue that the tests direct the instruction and curriculum in a positive way by providing clear learning outcomes (Madaus, 1991).
Testing can also motivate students to work hard and focus on what they need to study if they know they are going to be evaluated (Sloane & Kelly, 2003; Stiggins, 2005). Through testing, students soon see the connection between effort and high-test scores (Sloane & Kelly, 2003; Stiggins, 2005).

Opponents believe that high-stake tests narrow the curriculum (Madaus, 1991; Shepard, 1991). Many teachers express that they only have time to teach what is on the test because of the increased amount of testing (Madaus, 1991; Shepard, 1991). This approach to the curriculum limits students’ experience to hands-on learning (Madaus, 1991). Some students become overwhelmed from the competition as a result of all the tests administered (Sloane & Kelly, 2003; Stiggins, 2005). Others become frustrated and decide it is not worth the effort to try because they have failed so many times before (Sloane & Kelly, 2003; Stiggins, 2005).

Standardized tests are usually given at a predetermined time set by the state that does not generally coincide with closure of learning (Kleinsasser, 1995). Students express anxiety over the length of the test, the long testing periods without the possibility of talking with peers, and the chance of failing a grade (Triplett & Barksdale, 2005). With this fear of failure comes a lack of motivation (Madaus, 1991). Students who typically do not perform well on tests would rather drop out of school than fail another exam (Horn, 2003; Madaus, 1991). Adversaries of high-stakes testing believe that the tests themselves put more students at risk of failing without showing a significant increase in student learning (Horn, 2003).
A relationship exists between high-stakes testing and student outcomes from student subgroups (Horn, 2003; Triplett & Barksdale, 2005). Typically students in subgroups score below basic on standardized tests (Horn, 2003; Triplett & Barksdale, 2005). The intent of high stakes testing is to improve student learning; however, high-test scores do not always mean increased learning since students may only be taught that on which they will be tested (Horn, 2003).

*Current response to No Child Left Behind.*

The types of tests administered and the targets for meeting adequate yearly progress are determined by each state (Caillier, 2007; Usher, 2012). Therefore, comparisons from adequate yearly progress results cannot be made because of the lack of standardization among states (Caillier, 2007; Usher, 2012). In addition, schools not making adequate yearly progress should not be considered inferior educational facilities because the rigor of state tests varies so greatly (Usher, 2012).

Over the past six years, the nation went from 29% failing to make adequate yearly progress to 49% in 2011 (Usher, 2012). The trend in that time period was a slight increase each year; however, in 2011 there was a 10% increase in the number of schools not making adequate yearly progress when it went from 39% to 49% (Usher, 2012). In 2011, roughly 49% of the schools in the United States failed to make adequate yearly progress since implementation of NCLB (Usher, 2012).

There are different reasons to explain this trend (Riddle & Kober, 2011). The national percentage can change greatly when data are taken from some of the medium to larger states like California, Illinois, Missouri, New York, North Carolina, Oklahoma,
South Carolina and Florida (Riddle & Kober, 2011). When a majority of the schools in these states either make or fail to make adequate yearly progress, then the national average is greatly affected (Riddle & Kober, 2011).

Probably the biggest reason schools are not making adequate yearly progress is that schools do not make their annual measurable objective (Riddle & Kober, 2011). If schools cannot make small gains each year, then they will not make adequate yearly progress (Riddle & Kober, 2011). When a state the size of California does not show growth in student achievement, the national average will be affected proportionally (Riddle & Kober, 2011). In Oklahoma, the annual measurable objective only changes once every three years and enables more schools in that state to make adequate yearly progress (Riddle & Kober, 2011).

However, as 2014 approaches, more schools will possibly fail to make adequate yearly progress due to back loaded achievement trajectories (Usher, 2012). Many schools made small proficiency gains in the early years and now they must make larger gains in order to be 100% proficient in 2014 (Usher, 2012).

Other factors that determine adequate yearly progress results are states deciding to change the test or allow the proficient scores from retested students to be used as part of adequate yearly progress (Riddle & Kober, 2011). In 2008, North Carolina administered a new reading test for students in grades 3 through 8 and established a new base line to be proficient which resulted in more schools failing to make adequate yearly progress (Riddle & Kober, 2011). That same year, North Carolina also retested students who were not proficient (Riddle & Kober, 2011). Those who were retested and scored proficient
provided enough gain for some schools to make adequate yearly progress (Riddle & Kober, 2011). In addition, in South Carolina state officials lowered the cut scores in order to be proficient which allowed more schools to make adequate yearly progress between 2008 and 2010 (Riddle & Kober, 2011). However, New York in 2009-2010 changed their cut scores in order to align their proficiency standards so they experienced a significant drop in the number of school meeting adequate yearly progress (Riddle & Kober, 2011). At first glance one would think New York’s schools are underperforming; however, changing proficiency scores pushed New York students harder in an attempt to better prepare them for college (Riddle & Kober, 2011).

Some schools have shown improvement through the safe harbor provision (Riddle & Kober, 2011). Safe harbor allows schools to meet adequate yearly progress guidelines if the percentage of students scoring below proficient decreases by 10% or more (Riddle & Kober, 2011). Illinois went from 2% of its students making adequate yearly progress to 15% in 2010 (Riddle & Kober, 2011). This provision allows many schools to meet adequate yearly progress guidelines; without it the national percentage of schools failing to make adequate yearly progress would be greater (Riddle & Kober, 2011).

Utilizing a growth model enables some schools to decrease the number of schools failing to make adequate yearly progress (Riddle & Kober, 2011). In 2006, 71% of schools in Florida did not make adequate yearly progress (Riddle & Kober, 2011). In 2007, that percentage fell to 66% (Riddle & Kober, 2011). However, during the next four years the percentage continually increased (Riddle & Kober, 2011). These data
suggest that the initial implementation of a growth plan only serves to delay a school from making adequate yearly progress (Riddle & Kober, 2011).

NCLB was intended to align accountability among schools in the United States (Riddle & Kober, 2011). However, huge discrepancies exist among states and within them (Riddle & Kober, 2011). The adoption of common core standards will do nothing if states continue to vary so greatly in accountability (Riddle & Kober, 2011). Ten years after implementation, key stakeholders have noted problems with the policy (Usher, 2012). While states are still responsible to show growth on yearly assessments, waivers have been granted to schools not reaching their goals (Usher, 2012). The next section highlights the types of assessments used for accountability purposes and discusses some of the issues of utilizing the test data.

Assessment in Public Schools

Prior to No Child Left Behind states used overall averages for accountability purposes; however, this resulted in achievement gaps. NCLB developed the idea of setting adequate yearly progress to establish goals for student learning. The measure of student achievement thus began to be measured by national assessments and state adopted reading and math tests.

The next section reviews the types of examinations being utilized nationally and in my home state of Wyoming. The availability of test data to teachers and the issues with accessibility are then discussed. The final section highlights how teachers use data to impact instruction and ends with effective teacher and school practices that have utilized data to improve the academic performance of their students.
The National Assessment of Educational Progress (NAEP).

The National Assessment of Educational Progress is the only national standardized assessment currently being administered in the United States (Hombo, 2003). This assessment was designed to assess academic achievement in a variety of subject areas every two years (Hombo, 2003). When it was first initiated nationally in 1969, the goal was to measure academic trends from students in grades 4, 8, and 12, as well as report information from the different groups within those populations, such as gender, race/ethnicity and socio-economic status (Hombo, 2003; Johnson, 1992; National Center for Educational Statistics, 2012).

The National Assessment of Educational Progress has never assessed all students, but rather uses a multistage probability sample (Johnson, 1992). The first sampling stage is drawn from selected geographic areas that consist of metropolitan and non-metropolitan areas classified as high minority or not high minority (Johnson, 1992). The second stage of selection is derived from elementary and secondary schools (Johnson, 1992). Schools with higher concentrations of diversity are deliberately sampled to increase the precision from subpopulations (Johnson, 1992). The third stage of sampling involves how the assessment is administered (Johnson, 1992). Some assessment sessions use an audiotape prompt to pace students through the testing booklet, while the other half consists of a print session with no audio pacing (Johnson, 1992). The final stage is a systematic selection of eligible students made from an assigned school (Johnson, 1992).

The Commissioner of Education Statistics in the United States Department of Education is responsible for administering this federally mandated law (Hombo, 2003).
The National Assessment Governing Board develops the assessments that utilize the latest advancements in assessment methodology (Hombo, 2003). This board also sets achievement levels for each assessed subject area (Hombo, 2003). The National Assessment of Educational Progress uses the same assessments across the nation resulting in a uniform way to compare data between states (National Center for Educational Statistics, 2012). Since changes to the tests are rare, the data provide academic progress over time (National Center for Educational Statistics, 2012). The National Assessment Governing Board strongly believes that any minor altering of the assessment can have drastic effects on performance (Johnson, 1992). However, small alterations are necessary for the assessment to remain current (Johnson, 1992). These changes must be met with consensus through a committee of diverse constituents (Johnson, 1992).

In the United States, local educational agencies decide which type of assessment will be used to evaluate student achievement (Hombo, 2003). This inconsistency between schools and districts and within states has made it difficult to make comparisons (Hombo, 2003). The State Assessment of the National Assessment of Educational Progress was created in 1990 to measure achievement in different subject areas, report factors related to achievement, and measure academic trends over time between states (Hombo, 2003; Johnson, 1992; National Center for Educational Statistics, 2012).

Participation in the National Assessment of Educational Progress is totally voluntary for students, and parents are informed that their child is not required to take the assessment (Hombo, 2003). Students who do participate are only required to test for the
allotted amount of time rather than complete the entire assessment (Johnson, 1992). The National Assessment Governing Board feels that requiring students to perform for an extended period of time would have a negative impact on student performance so each child is only assigned and evaluated on a subset of items (Johnson, 1992).

The National Assessment of Educational Progress has never published results from individual students or schools; it is actually forbidden (Hombo, 2003). The National Report Card reports the results of the assessment to the public (National Center for Educational Statistics, 2012). This report describes what students know and can do in the various subject area and compares achievement data between states and various student subpopulations (Hombo, 2003; Johnson, 1992; National Center for Educational Statistics, 2012).

Prior to NCLB the goal of the National Assessment of Educational Progress was to monitor academic trends (Hombo, 2003). Now an additional goal is to evaluate the types of assessments used by states to measure their adequate yearly progress and then to compare academic results between states (Hombo, 2003).

*Proficiency Assessment for Wyoming Students (PAWS).*

In 2005, the state of Wyoming adopted the Proficiency Assessment for Wyoming Students (PAWS) to annually assess reading, writing, and mathematics for student in grades 3 through 8, and 11 (Wyoming Department of Education, 2012). A science test was added in 2008 for grades 4, 8, and 11 (Wyoming Department of Education, 2012).

The purpose of the Proficiency Assessment for Wyoming Students is to assess student performance against the Wyoming Content and Performance Standards as well as
to fulfill the requirements of the federal No Child Left Behind Act (Wyoming Department of Education, 2012). The results of this high-stakes test are used not only to evaluate if individual students are learning but also to determine if the individual schools and the state as a whole are performing and showing growth (Wyoming Department of Education, 2012).


The state uses Advanced, Proficient, Basic and Below Basic to measure student performance levels (Wyoming Department of Education, 2012). A score of proficient indicates that a student is meeting grade level standards (Wyoming Department of Education, 2012). Accommodations for students with Individualized Education Plans, students on a 504 Plan, or English Language Learners are permitted (Wyoming Department of Education, 2012). To be in compliance with the No Child Left Behind Act, results are reported to students, parents, schools, school districts, and the public (Wyoming Department of Education, 2012).

Data collection and its availability to teachers.

Currently in education students are participating in extensive evaluation activities (Nelson & Eddy, 2008). Typically high-stakes test are given in the spring (Sloane & Kelly, 2003). Because tests are given so late, the results are usually not available to
provide teachers with any diagnostic information to utilize in their instruction for the
current year (Heritage, 2007; Sloane & Kelly, 2003).

Since state test data are not available in a timely manner, multiple assessments
need to be given throughout the year to help guide instruction (Heritage, 2007; Nelson &
Eddy, 2008). Some schools have begun to take benchmark tests three to four times a year
(Heritage, 2007). These tests give teachers an indication of how students will perform on
state tests, but they do not provide the information needed to transform both teaching and
learning (Heritage, 2007). Since the concepts tested have already been taught, the
assessments only evaluate whether instructional goals have been met (Stiggins, 2005).

Issues with accessibility of data.

When schools do receive the results of state assessments, these results are not
very useful as a diagnostic tool (Nelson & Eddy, 2008). Not only do the data arrive well
after the students are gone, the aggregate data received are the result of several combined
measurements (Nelson & Eddy, 2008). Although state tests are required to inform policy
makers on the academic progress of students and schools, these assessments do not tell
the complete picture of how a child is learning (Nelson & Eddy, 2008). In order to drive
instruction and provide interventions for improvement, test data need to be disaggregated
to locate specific skills in which students are struggling (Nelson & Eddy, 2008).

Unfortunately, the process of disaggregating the data is time consuming and very
difficult (Nelson & Eddy, 2008; Reeves & Burt, 2006; Wayman & Stringfield, 2006). To
be effective, teachers need to be able to access the data and create reports so they can
adapt their teaching around the strengths and weaknesses of their students and manage
their instructional time more efficiently (Nelson & Eddy, 2008; Wayman & Stringfield, 2006). Making sense of all the data can cause anxiety since it requires the skills of gathering, interpreting, and understanding the information in such a way that will best impact a school or classroom (Earl & Fullan, 2003; Reeves & Burt, 2006). Most educators have not been trained in data interpretation, nor do they have experience in data collection and data management (Earl & Fullan, 2003; Reeves & Burt, 2006).

Assessment results in today’s education serve as a way to hold schools accountable (Earl & Fullan, 2003). Using data to target student achievement is new with the amount of high stake testing and most teachers lack the ability to disaggregate the data efficiently (Reeves & Burt, 2006). Effective schools are able to utilize the information to make informed decisions; however, it is not usual for educators to describe themselves as statistically illiterate when they try to make sense of all the numbers (Earl & Fullan, 2003). Professional development is needed to enhance teachers’ skills (Nelson & Eddy, 2008; Reeves & Burt, 2006; Wayman & Stringfield, 2006). The real question is not whether to integrate the use of data in decision-making, but how (Reeves & Burt, 2006; Wayman & Stringfield, 2006).

For schools that are able to identity a clear pattern in the data and develop an action plan for school improvement, there are other obstacles to overcome such as software tools (Boudett et al., 2005). Many teachers are given crash course instruction on how to use Excel and PowerPoint to help sort the data (Boudett et al., 2005). The technical challenges encountered when using these software tools hinder the data analysis
process (Boudett et al., 2005). Educators need hands on workshops with support experts to learn how to use technology effectively (Boudett et al., 2005).

**Teachers’ use of student test data.**

In an era of data driven instruction, effective teachers test often in order to foster student learning (Heritage, 2007; Sloane & Kelly, 2003; Stiggins, 2005). An assessment given during instruction is referred to as formative assessment (Stiggins, 2005). Formative assessments allow the teacher to make instructional decisions based on the results of the skills evaluated in hopes of moving the learning process forward (Heritage, 2007; Stiggins, 2005).

By administering formative assessments more frequently, teachers can identify struggling learners, monitor student growth, and identify gaps in learning (Heritage, 2007; Nelson & Eddy, 2008; Stiggins, 2005; Stiggins & DuFour, 2009). This type of testing allows teachers the immediate opportunity to directly address the needs of their class if progress is not occurring and then provide the systematic intervention needed to positively impact student learning (Heritage, 2007; Nelson & Eddy, 2008; Stiggins, 2005; Stiggins & DuFour, 2009). Formative assessments can build confidence in the reluctant learner by allowing students to acquire each standard through a scaffolding approach (Stiggins & DuFour, 2009).

Another way teachers use formative assessment is when they use the assessment for learning (Stiggins, 2005). Assessment for learning involves showing the student examples of advanced, proficient, and basic work in order to establish clear academic targets (Stiggins, 2005). With this approach, students use their own personal data to self-
assess and set goals for improvement while instruction is taking place (Heritage, 2007; Stiggins, 2005). Assessment for learning is one of the few times students respond to assessment positively (Stiggins, 2005).

*Effective teacher and school practices.*

Busy teachers need quick effective ways to analyze their student data (Schmoker, 2003). Two important questions to consider to improve student performance are what subjects are students succeeding, and what are their areas of strength or weakness (Schmoker, 2003)? The answers to these questions help teachers focus their efforts, resulting in achievement gains (Schmoker, 2003). Setting goals for improvement can be a powerful tool because they help target areas of low performance (Schmoker, 2003). With extra attention given to areas of need, weaknesses quickly turn into areas of strength (Schmoker, 2003). The short-term results obtained from test data enable teachers to adjust their instruction to meet the individual needs of students (Schmoker, 2003).

In order to make substantial improvements to test scores, schools must drastically change their approach in how they educate their students (Nelson & Eddy, 2008; Reeves & Burt, 2006). To accomplish this task, a strong leader is necessary to facilitate best practices (Nelson & Eddy, 2008; Reeves & Burt, 2006). Successful schools provide professional development and implement data-driven decisions (Nelson & Eddy, 2008). Restructuring the day based on student data, providing block scheduling, and grouping students according to performance level also help to improve scores (Nelson & Eddy, 2008). Grouping students based on ability levels allows teachers to focus on specific subject areas rather than differentiate to a wide range of needs (Nelson & Eddy, 2008).
Tracking student progress frequently not only provides direction to instructional support but also helps accelerate learning (Nelson & Eddy, 2008).

Knowing a child’s strengths and weaknesses is essential when grouping students and developing individualized curriculum (Nelson & Eddy, 2008). Schools are beginning to utilize computer assessment because it organizes data and provides easy access for teachers who need the information the most (Nelson & Eddy, 2008). One current computer assessment on the market and available to schools is from Northwest Evaluation Association called Measures of Academic Progress (NWEA, 2012). This online test interacts with each student individually since during testing the questions are determined by the answer the child selects (NWEA, 2012). The test increases the difficulty for correct responses or follows up with an easier prompt if a question is missed (NWEA, 2012).

Measures of Academic Progress determines precisely which concepts a child has mastered in relation to state and national standards (NWEA, 2012). Identifying strengths and weaknesses allows teachers to eliminate time spent on skills already mastered so instruction can be provided for what is really needed (Nelson & Eddy, 2008). The focus becomes developing plans for action around test scores (Nelson & Eddy, 2008).

Students today are exposed to a variety of tests that produce a great deal of useful data for teachers. Teacher accessibility to the data is sometimes difficult, but educators that are able to use the results to guide instruction have shown increased academic performance.
Conclusion

With the rising demands for improvement and accountability required under the No Child Left Behind Act, schools are beginning to see the value of making data based decisions to dramatically increase student achievement on high stake tests. Unfortunately the information from these end-of-the-year tests often arrives too late in the school year to impact learning. Data from high stakes testing have allowed educators to see the importance of classroom assessment as a vehicle to assess their own work and impact student learning. Schools that have effectively used assessment data have ignited change and achieved positive results. The next chapter will investigate the methodology used in determining which assessment data are most useful to drive instruction and improve student performance.
CHAPTER THREE
METHODOLOGY

Introduction

The current literature repeatedly states that schools today are under extreme pressure to demonstrate academic growth as part of the No Child Left Behind Act. The window for administering these high stakes tests occurs very late in the school year preventing the resultant data to be of any use in the current year. In an effort to make annual yearly progress (AYP), many schools have begun to test more frequently and make data driven decisions to help struggling students.

For almost twenty years, I taught in a small rural school district in California that would make annual yearly progress one year, but never two years in a row. In 2008, a move to the mid-west resulted in my acceptance of a teaching position at a small rural school in Wyoming. My teaching experience at this school afforded me the opportunity to utilize data from Measures of Academic Progress with students over the past five years. Since students also take the Proficiency Assessment of Wyoming Students, I have also witnessed first-hand how decisions based on data have influenced results of this high stakes test. This elementary school has never failed to make annual yearly progress since the enactment of No Child Left Behind of 2001. This phenomenon is rare. In 2012, one hundred eighty schools in the state of Wyoming alone did not make annual yearly progress (Wyoming Department of Education, 2012, September). The ability of this successful school to consecutively make annual yearly progress when so many other
educational agencies were unable to do so prompted me to investigate the factors that may have been influential.

The purpose of this observational case study is to answer the question: What is the impact of on-going assessment data on teacher practices and attitudes in a small rural elementary school? The aims of the research are to reveal if and how teachers use assessment data to make instructional decisions, and what impact, if any, this has on the school’s on-going high performing status.

\textit{Context and Development of Research}

The research that forms the basis of this thesis was conducted during the fall and spring semesters of the 2012-2013 academic year. At that time, West Elementary School, a pseudonym I am using for the purposes of this thesis, was one of five elementary schools in the county serving over 350 students in grades kindergarten through fifth grade with approximately 48 staff members. In the previous three years, the school experienced an increase in male students and those eligible for free/reduced meals, but the number of students qualifying for special education services remained fairly level. Class size in the primary grades was approximately 16 to 20 students and 18 to 20 in grades three through five. Each grade offered three to four sections.

Reading was taught using a balanced literacy approach with leveled books and a basal reader. Writing instruction utilized the Six Traits writing model and sequential writing lessons from Lucy Calkins. The Calkins method provides guidance through all phases of the writing process. For math instruction the school adopted the Bridges
Program that uses direct instruction, structured investigations, and open exploration to master mathematical concepts at each grade level.

In 2012, the Common Core Standards for mathematics were implemented throughout the school. Instruction in music, physical education, and media/library was offered weekly. In addition, the computer lab provided technology skills and individual learning opportunities in the core academics.

One of the strengths of the school was the Building Intervention Team. This team of professionals met weekly with classroom teachers and parents to apply systematic intervention to assist children who were having difficulty learning. In addition, a Home-School Liaison who tracked and supported students considered to be at-risk also supported identified students. These students attended an after-school study hall in which the Home-School Liaison assisted them with homework and spoke daily with parents to discuss intervention strategies.

Professional Learning Communities (PLCs) had been in place for five years. Comprised of teachers and support staff, Professional Learning Communities work collaboratively to enhance their effectiveness with students. Multi-grade level Professional Learning Community teams from the primary grades and from grades three, four, and five met weekly to plan. Thirty minutes twice a week were set aside as a time when students were clustered for remediation and enrichment.

In order to monitor student progress, the school administered the Measures of Academic Progress computer assessment three times annually to its students in all grades.
This standardized growth model assessment was used in planning and implementing instruction based on individual student needs.

Students and staff recited the school challenge every morning: “I challenge myself to be a responsible citizen with high expectations for success.” Effort by teachers and children was evident. In 2001, they were rated as one of the top performing schools in the state of Wyoming. In 2007, the U.S. Department of Education recognized them as a Blue Ribbon School. This accolade recognizes high performing schools or ones who have made significant improvements in closing achievement gaps. This school was one of three in Wyoming to be recognized in 2011 with the Ellbogen Meritorious School Award. This honor is only eligible to schools comprised of at least 20% board certified teachers and who demonstrate excellence in education. West Elementary School made annual yearly progress every year since the enactment of the No Child Left Behind Act. In 2012, they continued to excel academically on the state Proficiency Assessment of Wyoming Student by achieving 95% proficient in both math and reading in grades three through five.

At the start of the research, I had taught third grade at West Elementary School for five years so I had a professional relationship with the research participants and the setting for this research project.

**Sampling Technique & Instrument**

West Elementary School was chosen for this study because of its consistent testing success over time. Each year they reached their projected benchmarks of proficiency while other schools in the district using similar programs struggled. When
Measures of Academic Progress data from the last three years in grades three through five were reviewed, the average growth was consistently higher than national averages.

I selected participants for this study using two criteria: they had to have administered the standardized high stakes test called Proficiency Assessment of Wyoming Students and to have utilized data from the computer assessment called Measures of Academic Progress. All teachers were selected from a single high performing school that had made annual yearly progress since 2001 as measured on standardized state assessments.

I also obtained insight from third through fifth grade teachers and the school’s Professional Learning Community Coordinator using open-ended interview questions. The Professional Learning Community Coordinator was included because of his expertise in presenting Measures of Academic Progress test data and organizing differentiated instruction in grades three through five. For this study, data were obtained through structured one-on-one interviews with pre-developed questions that I gave to the participants prior to their scheduled meeting.

Many of the open-ended questions were directly related to Measures of Academic Progress testing. The interview questions fell into three distinct categories: setting, perspectives, and strategies. In the setting category, I asked general questions such as the number of years the participants had been in education and at the elementary school being studied, as well as the current grade level assignment. Teachers also described their experience with Measures of Academic Progress testing. The purpose of the
information from this section was to inform me of the teachers’ experience and to assist me in evaluating whether the test was being used similarly across grade levels.

The next set of questions encouraged teachers to elaborate on their perceptions of Measures of Academic Progress testing. I specifically asked participants to define how they felt the test was being utilized. In order to elicit personal opinions from respondents, I asked them to state and evaluate both the benefits and drawbacks of the test. These answers were included in an effort to provide a starting point for identifying possible themes. To draw more details from the respondents, I asked them to describe their perception of how often the school tests their students. This was not directly related to Measures of Academic Progress testing but testing in general. Since the literature stated that there is an abundance of test data available to schools, I was hoping to learn from teachers how they perceived the amount of testing taking place.

The third category of questions focused on strategies. The information participants provided in this section could have the potential to influence how other teachers and schools utilize data in order to show academic growth each year. I asked teachers to describe how they disaggregated the Measures of Academic Progress test data, how they modified their teaching practices to increase students’ learning outcomes, and if they perceived a connection between using on-going assessment data and schools being able to make annual yearly progress.

To validate interview questions, face validity was utilized through a small focus group comprised of teachers employed at the school where the research was conducted. Five potential participants received a letter at school detailing the purpose of the study, a
copy of the interview questions, and a consent form. Experience utilizing test data from Measures of Academic Progress was the major criterion for selection of these individuals.

Four teachers returned their consent forms; however, only two of the teacher committed to the scheduled discussion due to summer vacation conflicts. I met with one kindergarten teacher and one second grade teacher at West Elementary. I reminded participants that I would be audiotaping the discussion so I could review their responses, and then I began asking my predetermined questions. (See Appendix A for the content of the questionnaire.) Participants took turns answering each of the questions in a roundtable method. After reviewing the audio recording of the group interview, I felt their responses were descriptive which determined that my questions were valid. Since none of the participants in the focus group administered the standardized high stakes test called Proficiency Assessment of Wyoming Students at their grade level, they were not included in the actual study.

Procedure

To obtain information for my research, I gave nine potential participants a written letter in their mailboxes at school explaining the purpose of the study and a copy of the interview questions to be asked. In order to make this a legally effective informed consent, I included the following key components in my cover letter: the purpose of my research, the form it would take (interview), the amount of time, potential risks and benefits of participation, compensation (none), names and contact information about the research and participants’ rights, voluntary nature of their participation, and ability of participants to decline to answer specific questions or to withdraw from the study.
Seven individuals signed and returned their consent forms indicating they were willing to be included in the study. Confidentiality of the seven participants was of the utmost importance. I created identifiers to code the interviews to shield participants’ identities since some participants could have felt uncomfortable talking about the ways in which they used or did not use the test data in their teaching practice. They could also have had negative perceptions about the move toward data-driven instruction and standardized testing and may have had concerns about revealing those perceptions. For this reason, I aggregated responses and maintained no identifying information.

I interviewed one teacher each week during the summer, audiotaping the interviews to provide an exact transcription of what was said. I began the coding process by reading all of the participants’ responses to questions one through three. Since these were just background questions focusing on experience, I kept track of their years in education, years at West Elementary, and job assignment by making a graph of the information. Starting with question number four, I noted each participant’s response. I highlighted key words or phrases the first time I read the individual responses. After reading all the comments a second time, I was able to write key words or phrases that served as a way to code their responses. To ensure consistency, after the initial coding of each interview, I waited a few days and recoded each interview a second time. By comparing the data across participants, I found that core ideas emerged which captured participants’ perspectives regarding the impact of on-going assessment on teaching practices. To support how I arrived at the different themes, I highlighted quotes taken from participants for each question.
Finally, I used theory triangulation to make the research more credible. Theoretical validity helped to support my research by referencing the different theories presented in the literature review. Test data from the current Measures of Academic Progress available at the school reinforced my theory of the phenomenon being studied at this high performing school. I referenced the school’s Measures of Academic Progress data by grade level to see how well specific grades were performing compared to the national average.

After all of the interviews were completed I accessed Measures of Academic Progress test data for the 2012-2013 school year to validate how well West Elementary students were performing. Results for math and reading by grade level were averaged from fall to winter and fall to spring, and those averages were compared to national averages.

In third grade the average gain for math from fall to winter was 11.6 RIT points compared to 6.4 RIT points for the national average. End of the year achievement at West Elementary was higher than the national average as well. Third graders showed a 16.2 increase on math results from fall to spring where as the national average was only 11.2 RIT points.

Fourth grade performance from fall to winter in the area of math was 6.9 RIT points compared to the national average of only 5.1 RIT points. Fall to spring scores was substantially higher as well. Fourth graders at West Elementary had an average increase of 12.6 points from fall to spring compared to only an 8.9 increase for the national average.
Fifth grade scores showed an increase of 2.5 points from fall to winter for math that was slightly lower than the 3.1 national averages. Fifth grade did improve drastically from fall to spring in math compared to the national average. This grade level increased on average 9.1 RIT points compared to 5.9 nationally.

Reading scores from fall to winter and fall to spring were then averaged for each grade level for reading. In third grade students showed an increase of 7 RIT points from fall to winter compared to 4.7 points for the national average. Fall to spring results showed an average increase for reading of 9.1 points that was slightly lower than the national average of 9.3 RIT points.

Fourth grade test data for reading from fall to winter averaged an increase of 5.4 RIT points compared to national averages of only 3.4. Fall to spring average growth was slightly higher as well. West Elementary fourth graders showed an increase of 7.1 RIT points compared to 6.9 RIT points nationally.

Fifth grade reading scores were slightly higher than national averages from fall to winter. West Elementary students showed an increase of 3.0 from fall to winter above the 2.7 national average. Fall to spring averages for fifth grade, 5.6 was also higher than the 5.2 national average.

Summary

Data for this research were obtained through open-ended interview questions. To validate the questions, I met with a small focus group of teachers who were not in the study and asked my list of predetermined questions. This process allowed me to evaluate if the questions would provide enough detail. Based on the responses shared in the focus
group, I felt confident to proceed with my research. Scheduled personal interviews with each of the participants allowed teachers to share how they use assessment data to make instructional decisions. The impact of on-going assessment data on teacher practices and attitudes was also explained by participants. After transcribing and reading each of the interviews, I began to interpret the narrative data into various themes by looking for patterns. By comparing the data across participants, I found core ideas emerged. The major themes and connections will be the topic of Chapter Four, which contains the results of the research.
INTRODUCTION

This chapter provides an overview of the results of research described in the previous chapter. It included demographic information about the research participants as well as the results of the open-ended interview questions detailing how the participants utilized the results from Measures of Academic Progress (MAP) to improve student performance on end-of-the-year exams. The interview questions fell into three distinct categories: setting, perspectives, and strategies. The themes that emerged will be discussed in each section.

PARTICIPANT DEMOGRAPHICS & SETTING

Seven of the nine teachers from a single high performing elementary school who were invited to participate in the study returned their consent forms. The participants included two teachers from third grade, two from fourth grade, and two from fifth grade, as well as the Professional Learning Community Coordinator. I notified the participants by phone and arranged to meet with them privately at the school site to ask the predetermined interview questions. Each interview lasted approximately 45 minutes.

The first part of the interview asked teachers to provide the number of years they had been in education and at this high performing elementary school. Ninety percent of the participants had taught for more than ten years, with a mean of 18 years of teaching experience. All of the teachers surveyed were employed for an average of nine years of service at the school.
Participants were then asked to describe their experience with Measures of Academic Progress. The teachers all shared that the Measures of Academic Progress assessment was administered three times a year: once in the fall, again in the winter, and finally in the spring. All of the teachers indicated that the district required the reading and math tests and noted that the language arts portion was optional. Only one teacher in the study opted to test language arts and offered this comment to explain why:

I’ll be honest. I did not focus on it (language test) as much, and that’s why my language arts results are not that great. My focus was more on math and reading. Looking at how dismal some of those scores are I’m looking at changing that next year. I know some of the teachers don’t test language arts, but I like the data.

The information from this segment informed me of the teachers’ experience and allowed me to determine that the test was being used similarly across grade levels. In the next section, I will share teachers’ perceptions of Measures of Academic Progress testing.

*Perspectives of Measures of Academic Progress*

The next set of questions allowed participants to express their perspective of Measures of Academic Progress testing based on first hand experience of administering the assessment with their students. The data in the following sections detail the participants’ responses and highlight the different themes that emerged.

*Utility.*

Participants were asked to define how the Measures of Academic Progress assessment was being used at their elementary school. All of the teachers indicated that Measures of Academic Progress is a test designed to show academic growth based on
RIT values not only from year to year, but from segments of the year: fall to winter; winter to spring. One individual stated, “As far as the tools we’ve got, this is the only one that measures growth. For example, this child has grown one year in math. It is statistically pretty accurate because it is norm referenced.”

Some participants expanded on their opinion of the intended use of Measures of Academic Progress testing. Forty-three percent of the participants expressed that they felt the data were not being used as they were intended. Twenty-nine percent perceived that the test was being used inappropriately to assess teacher performance or to publicly compare them. One participant indicated, “It is being used as an administrator tool to evaluate teachers.” Another simply stated, “The test is being used against us in front of our peers during meetings.”

Benefits.

Participants were asked to describe the benefits of Measures of Academic Progress testing. One of the top themes expressed by 86% of the teachers was how quickly results were made available. Teachers expressed that having the data available immediately was very helpful in their work with students. One teacher indicated, “As soon as the kids are done, we get the preliminary results as soon as the child logs off.” Another participant concurred, stating, “(We get results) instantly, that’s what I like about MAP. The report prints as soon as they finish the test. You have the data right then. That’s great because we can use it the next day and make changes which is nice.”

Of the 86% who liked how fast data were available, 43% explained how they prefer to have the results immediately after testing as opposed to waiting on the results
printed by the district office. One participant explained, “I get it as soon as the kids are
done, it has a print out. That is what I use. I don’t wait for the district to get their data
up-loaded, otherwise it would be two weeks later.”

Another theme that appeared in 86% of the responses was the ability of the results
to be broken down into categories for specific skills. These participants expressed how
much they appreciated knowing exactly where a child is struggling or excelling. One
teacher explained, “It identifies very quickly to teachers which students really need
intensive work and which students are going to need some enrichment.”

A third theme that emerged as a benefit was the ability of the data to show growth
in different subjects. Fifty-seven percent of the teachers perceived this as a benefit. One
teacher explained, “The benefits are that growth. You get a chance for a kid to show
growth, especially in those already perceived areas of weakness. If they grow within
areas, that is valuable feedback since it monitors students’ progress.”

One final theme that was reported by twenty-nine percent of the participants as a
benefit was how the test data allow teachers to evaluate their curriculum. During an
interview, one participant offered the following explanation:

This class may be a little more difficult so my results may not be as great, or
maybe I felt this class was more high performing so hopefully I will get higher
results. In a way, we are able to judge the strength of a classroom and the
curriculum and measure that in a pretty stable platform that is a huge benefit.
Drawbacks.

The next question asked participants to reflect on the drawbacks associated with utilizing Measures of Academic Progress testing administered at the elementary school. Negative comments listed by participants fell into two distinct themes: unreliable test data and loss of routine.

Forty-three percent of the teachers in the study felt one of the drawbacks were skewed test results from taking the multiple-choice assessment on a computer. One participant explained:

Being on a computer, because students can guess very easily, it shows their scores higher with good guesses. Sometimes when they are working really hard, they get fatigued because the test is really long. It’s not a true tale of how they are doing.

Another participant related a similar experience:

I’ve had some instances where a child was already identified as struggling and was in remedial reading with interventions in place, and the student got the highest score in the class. What it boiled down to was that she guessed really well.

Students at the elementary school go to the computer lab twice a week for instruction. Three times a year, the computer lab is unavailable for classroom use due to Measures of Academic Progress testing. Seventy-one percent of the teachers felt loss of instruction due to the testing schedule was a drawback. One participant offered this reason:
It is very disruptive. Each time you have a testing session, you basically lose three weeks of routine. Kids and teachers like routine. Even though your kids might only test for a total of four hours during those three weeks, your entire three weeks is disrupted.

Another participant expressed:

It (Measures of Academic Progress testing) messes up nine weeks of school. Even though we only test twice, there are no computers, and it messes up the schedule for three weeks. Nine weeks a year, a whole semester is tied up with testing. You don’t have time to teach.

Amount of testing.

The interview included a section for participants to describe their perceptions of how often the school tests its students. All participants unanimously stated that students were tested too much. This question was not referring to Measures of Academic Progress testing alone but rather to testing in general. Different grade levels repeatedly described various semesters in which numerous assessments needed to be administered. One teacher explained:

The third semester of school I do feel like all I do is test because that is the trimester where we have Measures of Academic Progress testing, Proficiency of Wyoming Students testing, the state test which is three weeks long, the district reading assessment plus all the other district assessments for math. I really don’t look forward to the third trimester because it feels like all I do is test.
Participants provided not only their opinion regarding the amount of testing, but many offered a solution for how to eliminate the amount of testing. Eighty-six percent of the participants specifically stated that they did not feel the winter testing window for Measures of Academic Progress testing was necessary as a progress-monitoring tool. Many described other tests that were given during that time frame which prevented new concepts from being introduced in the classroom. One teacher stated, “The most growth I typically see on a Measures of Academic Progress test is from fall to winter. Winter to spring growth is not as consistent because you are spending more time on testing rather than academics.”

In this section, teachers elaborated on their perceptions of Measures of Academic Progress testing by specifically defining how they felt the test was being utilized. Participants also stated and evaluated both the benefits and drawbacks of the test and described their perception of how often the school administers testing. These narrative data revealed patterns that provided a starting point for identifying possible themes. The next section provides the instructional strategies participants suggested.

*Instructional Strategies*

To learn more how this successful school continued to make annual yearly progress as measured on standardized assessments, I asked teachers to reflect and share some of the strategies they utilized based on the test data. This next section offers an overview of the different strategies being implemented in the classroom to improve academic growth based on the test results from Measures of Academic Progress.
Disaggregating data.

Because the Measures of Academic Progress assessment provides numerous reports of test data, the interview included a section designed to shed light on how the educators in this study made sense of all their test data. Participants were asked how they disaggregate the data, and 100% said they initially look at the data to determine who needs help and who needs enrichment. One teacher explained, “First of all I see who is at grade level and who is not. The ones that are below grade level I see in which categories. The ones above grade level I can see by how far.”

A second theme that emerged with 29% of the participants was the connection between the data and the efficacy of the curriculum in the classroom. When the data show that a number of students are struggling with the same concept, then possibly the teacher needs to address how the lesson can be taught more effectively. However, if the same concept is difficult for the entire grade level, the curriculum might need to be evaluated. As one teacher shared, “I look across at the different categories and see if there are common gaping holes.” Another participant expressed a similar comment, stating:

When I sort it, it is by highs and lows. If I see an overall low, then it might be an area where our curriculum is lacking. More than likely we are doing a good job, but students coming into the building, they have definite deficits. After a couple of years in the building, the only difference is teacher strengths in how they teach and how they use their instruction.
After teachers explained how they disaggregate their Measures of Academic Progress data, I then probed them to describe specific strategies they utilized from their data. Fifty-seven percent stated that they grouped their students according to ability for math and reading based on the results. One teacher said, “Well, in our grade level, we get together, and we group the students below grade level, at grade level, and above grade level. Then we teach lessons based on the needs of each of the different groups.”

In addition to evaluating the curriculum and providing ability groups, 29% of the participants mentioned how the data help tell a story by grade level or by individual teachers. One teacher explained:

Right now we are in the process of using the five years of data to show what the average gain is for second grade math or second grade reading over a five-year period. Even though classroom strengths may be different from one classroom to another, we should have a pretty good idea of what the average gain is for them as a grade level and even be able to break it down for the average gain for that teacher so they know. They can compare their gains with their colleagues in that grade level so they can tell who is stronger in reading or math. We are encouraging them to look. If someone is coming out every year two or three points higher, then you need to observe what that teacher is doing because obviously they are doing something just a little different in their instruction, but year after year they are getting higher gains every year in that area. In fourth grade we know that one teacher is stronger in reading, and the other two are stronger in math, and if they could collaborate better than they could probably all
have stronger gains. It would benefit the students, but that takes time, and time is the enemy.

*Modifications to teaching practices.*

The themes in the following section are examples of how Measures of Academic Progress test data have been beneficial in terms of modifying teaching practices to increase students’ learning outcomes. Fifty-seven percent of the teachers stated that they review the test data and use ability grouping as a modification to their teaching. One participant described the classroom modification as follows:

I identify my target goal strands, and then I look for areas of weakness within each goal strand and look to how to help those students advance along that strand. Of course small group instruction is probably how I’ve done it the best by putting those kids with similar needs together.

In addition to ability groups, 29% of the participants mentioned the use of Compass Learning folders to allow students to work at their individual skill level. Compass Learning is a software program designed by Northwest Evaluation Assessment which imports test data from Measures of Academic Progress and creates folders on the computer for students that address their identified learning needs. One teacher described how differentiation works in their room:

It goes back to a resource that I use in the classroom for teaching that allows me to differentiate which is Compass Learning. The way I differentiate is I take their MAP score in the different areas, and I have them start working in their folders based on the data.
Twenty-nine percent of the participants utilized a teaching modification called the DesCarte binders provided by Northwest Evaluation Assessment. The purpose of the DesCarte binder is to help guide target instruction for students based on reports from Measures of Academic Progress assessment. This is how one participant explained it:

I look at the DesCarte that gives you the general activities you can do with kids in a certain RIT band, and this last year I organized kids on Friday into some type of math game or fun activity that had to do with a hole or spot they needed filled.

A final teaching modification that one participant used based on the identified student needs from the Measures of Academic Progress data was more of a classroom management technique. This teacher stated:

I really press that a good student is one who gives consistent effort, effort is a choice, and effort creates ability. I push that all year long. In class I call on them (student who is identified as low performing) more. I may call on them in class, then call on another child, and then come right back to them. I keep them on their toes a little bit more than some other students. If I keep their attention where they never know when I am going to call on them, then they pay attention better in class that will improve their test scores.

*Improvements to Measures of Academic Progress.*

The school relies heavily on the test data from Measures of Academic Progress to guide instruction. In an effort to improve data based decisions, I thought it might be helpful to learn what the people who use this testing program perceive as ways to make the data more user-friendly which could result in improved teaching strategies.
Several themes emerged.

Forty-three of the participants were unhappy with the software capabilities of showing growth rates for students. Participants felt this could partly be an issue with the type of computer one uses since some of the useful reports available can only be run on a PC, and the elementary school in the study uses Macintosh. One person explained:

Well I saw a report that I had never seen. This report can only be created on a PC, and I think it is really meant for principals, but in reality I think teachers need to have it. It is a report that shows a typical growth point from wherever they start.

Another participant suggested being able to disaggregate the data in one report by the entire grade level rather than just by individual teachers in order to show growth or areas of weakness. One teacher suggested,

If we want to see RIT gains for vocabulary and comprehension from fall to winter and be able to isolate that and do that not only for individual teachers, but for the grade levels so teachers would have an overall picture of what the grade level was looking like in specific areas.

Currently, there is no way to create reports for students across a specific grade level or for one student year to year. For example, one teacher noted, “If you want to measure summer slump, you have to go back and load the kids alphabetically and then subtract (fall scores) manually since the kids are no longer in the same class.”
A second theme that emerged regarding how to make the data user-friendly was the idea of lesson links. Twenty-nine percent of participants wanted to see specific lessons tied to RIT band scores. One person suggested the following:

I wish that as soon as you got the scores you could go somewhere and see specifically what they need. These are the areas the child needs to work on, this is what can be done at home, this is what I can do at school, independently. Rather than trying to search through all your toolboxes, it is so time consuming to find those activities or to figure out how you’re going to target it or exactly what it means. Sometimes those RIT bands are confusing. What is a poly diagram? So (it would help) if there were some examples of the top ten things to try.

A final theme suggested by 29% of the participants was improvement from the district. The amount of time it takes to upload the data on the district site so teachers can sort the data in different ways was very long. Having test results available immediately is critical when making instructional decisions, as one teacher expressed:

If we are truly going to use MAP data to guide instruction, then we can’t wait, in some cases, weeks to get the data uploaded onto Milepost so we can sort it. The quicker we can look at the different goal strands the better. I think this is more at the district level.

Another suggestion for the district would be to provide on-going in-service for teachers each year to show how to run the numerous reports available. One teacher explained:

It’s taken me forever to understand it (MAP data). Just what the RIT bands are. No one explained what this test was, how to use the data. There is no training on
how to use the data. How do I make use of it? The first year I just looked and
said, “That’s cool.” The next year I looked at those individual scores and could
see where my kids were low; that was a little better. This year my goal was to do
a better job. It’s hard to understand all the data because of the time it takes to sit
down and look at it, make some graphs and figure out what you are looking at.

We need a little explanation each year.

*Connections between ongoing assessment data and annual yearly progress.*

Schools each year must show annual yearly progress on high stakes assessments.
One of the last questions in this study asked participants to share whether they perceived
a connection between ongoing assessment data and yearly progress.

One hundred percent of the participants felt ongoing assessment data were useful
to reach growth targets. Teachers described that they are constantly looking at how their
students are doing compared to year-end goals and making adjustments since the data
help identify starting points and determine individual needs. The data, whether they are
from Measures of Academic Progress or informal observations, help teachers make the
incremental steps toward achieving academic growth. One participant observed:

Yes, I think there is a definite connection. A teacher can either go through the
curriculum and teach it and hope the kids got it, or they can look at the data and
see they didn’t get it so let’s redo that or fix this or do this. That’s really what
needs to be happening.

Another said:
Schools and teachers who use the data to see a child who is truly struggling, those schools are going to get all the students where they need to be by the next year. Sometimes they won’t make it, but the next year’s teacher will have the data to see where to start and move forward.

Participants’ final thoughts on assessment.

The final question of the interview asked participants if they had anything else they would like to add about assessment. Responses had some commonalities.

Twenty-nine percent of the teachers expressed the need for accountability but felt assessment is currently being misused. Assessments have become high stakes because they not only evaluate student performance but teachers and schools as well. One teacher expressed:

My biggest concern is that it is misused. That’s the problem. If you want to use it to see what kids have learned so you can go back and help them then I think it is a great tool. If you want to browbeat the teachers, then it is a horrible tool. If you want to use it to determine if schools are successful, it is a horrible tool. There are two sides to it. It depends on who is using the data and what they are using it for.

Forty-three percent of the participants expressed a need for more of a balance in terms of testing. Currently, students are being tested multiple times in a week which takes up valuable instructional time needed in order to teach the concepts students on which are to be evaluated. One teacher explained:
I think these types of tools are necessary evils. As a classroom teacher you lose valuable instruction time. We work really hard to get all those foundational pieces in before we get to the point where we are hitting some of the high stakes assessments but again at the expense of being able to teach. When you look at winter data compared to spring data that is actually our biggest problem. I don’t think you are going to see a lot of growth from winter to spring because we have done a majority of the instruction up until that time period and then you lose out on classroom time to teach, to push them a little further until after PAWS and after the state writing assessment but then you are overwhelmed with multiple district assessments in conjunction with the state assessments. Maybe (we should) do away with winter MAP testing since the state doesn’t require it.

Another teacher concurred by stating:

I think we have to strike a balance, what’s beneficial and necessary for our data processes to know and how to benefit our students and our curriculum and what is overkill. Once we understand that and keep a good positive approach to our assessment process then I think we are going to come out on the good side.

Twenty-nine percent of the participants felt that assessment was an excellent way to drive instruction in order to increase student achievement on standardized assessments. One teacher expressed:

At the beginning of year we worked on multiplication and fact fluency because we knew our kids had to have that to do well with common core so we started that right away, and it helped a lot. We broke the kids apart based on their fact
fluency, but it wasn’t based on MAP data, it was from a formative assessment, still data driven.

One final thought on assessment that one participant shared was the idea of teachers’ attitude towards assessment in terms of having either a positive or negative impact on test results:

If teachers see the assessment as overwhelming, overdone, there’s too much of it, I think if they start to develop an attitude that says it’s not going to be beneficial we’re being overworked, we’re being over-tested and they see it as a negative then I think our results are going to reflect that.

Teachers all had points to make as the interview came to an end. The ways in which assessments are used, the need for more of a balance in testing, the ability of test data to drive instruction, and the ways in which teachers’ attitudes can affect test performance were the main points expressed by the participants.

Summary

The responses from the participants yielded many insights as to how this high achieving school has managed to make annual yearly progress year after year. Teachers expressed that one of the benefits of Measures of Academic Progress is how quickly the data are available. Not only is the information available immediately, the data can also be disaggregated into specific skills so teachers can decide who needs intervention and who needs enrichment. Since the test is administered three times a year, teachers felt they could monitor the growth of their students easily and viewed this as strength.

Participants also revealed some drawbacks of Measures of Academic Testing.
Since testing takes place on the computer, some teachers felt this could result in skewed data. Many of the participants viewed the three-week testing schedule and amount of testing in general as unfavorable as well.

Teachers shared specific instructional strategies based on test data being implemented in the classroom to improve academic growth. Disaggregating the data to precisely determine who needs assistance and who needs enrichment was one strategy. Participants felt that by examining the disaggregated data, holes in the curriculum could be detected more easily. Knowing students’ strengths and weaknesses based on the data provided an easy way of grouping students based on individual needs.

The study revealed many modifications currently in use that could potentially improve student achievement. These included ability groups, the utilization of the Compass Learning folders and the DesCarte binder, as well as specific classroom management techniques. Teachers also suggested ideas to make Measures of Academic Progress more user-friendly that may result in improved student learning in the classroom. Participants expressed the need for reports that show typical growth rates from where students start, the ability to sort data across the grade level rather than just by teacher, the need of lesson links not only for the classroom but for home connections, and more in-service to effectively sort the data by running different reports. The following chapter is an analysis of the results presented in Chapter Four.
CHAPTER FIVE
ANALYSIS

Introduction

The question that guided this study was: What is the impact of on-going assessment data on teacher practices and attitudes in a small rural elementary school? This chapter will provide an analysis of the data collected by open-ended interview questions broken down into the following categories: setting, perspectives, strategies, and final thoughts regarding assessment.

Setting

The responses from the open-ended interview questions revealed that teachers at West Elementary utilized test data on a regular basis. Participants in this study included teachers from grades three, four and five, as well as the Professional Learning Community Coordinator. Because the vast majority of the participants had been teaching more than ten years (with a mean of 18), it is clear that all those who participated in the research were very experienced educators who had a long history of service at the school. Such longevity in the profession and as a group of colleagues at the school would provide a highly stable faculty with long-term relationships that could translate into greater intra-grade cooperation and communication. Student outcomes could be stronger in a stable setting with highly experienced educators who have worked together over a number of years than in one in which the faculty and administration experience a high degree of transiency, leading to a possible lesser ability to communicate and collaborate over time.
I believe the experience of each teacher contributed to the details conveyed in the responses to the interview questions. Despite the differences in teaching assignment, many commonalities existed in their responses. The next section will provide an analysis of how the teachers utilized the Measures of Academic Progress.

_Perspectives._

The participants in this study administered Measures of Academic Progress three times a year to their students. Their perceptions of how it was used as well as the benefits and drawbacks came from first-hand knowledge.

_Utility._

Analysis of the results of participants’ perceptions of the Measures of Academic Progress assessment revealed two distinct issues. First, participants all perceived that the assessment was designed to show academic growth and did indeed meet that goal. However, some were uncomfortable with the underlying use of test results. I posit that teachers feel less enthusiastic and supportive of student assessments when they perceive that administrators will use those assessments to judge teacher efficacy. The use of student test scores does not take into account the many reasons for students’ performance on such assessments, and utilization of the assessments to judge teacher performance can increase teachers’ mistrust of administrators, foster a competitive rather than collaborative atmosphere among teachers, and undermine the creation of sense of community at the school. It may also increase teachers’ focus on teaching toward the test since Measure of Academic Progress has become another high stakes assessment.
Benefits.

Speed in which results were made available, ability of test data to be disaggregated for specific skills, and the capacity of the data to show student growth in subject areas and to evaluate curriculum were the themes expressed by participants as the benefits of Measures of Academic Progress testing.

Eighty-six percent of the participants expressed how much they appreciated receiving the test data so quickly. The ability to use the results to make immediate decisions concerning a child has the potential of making a large impact on that child’s learning. At the time of this study, high-stakes test were generally given late in the year, and the results were usually not available to provide teachers with any diagnostic information to utilize in their instruction for the current year (Heritage, 2007; Sloane & Kelly, 2003). Teachers’ responses concurred with the literature, and their comments made it clear that empowering educators with current student data may facilitate teachers in developing action plans to improve student achievement. Teachers need to receive assessment data early in the year to enable them to modify instruction for their students, thus fostering greater student success. Assessments completed late in the year do not afford teachers the time to analyze the resultant data and make necessary curricular adjustments, both for specific students and groups as a whole. Thus, teachers often see preparing students for tests as a waste of valuable instructional time. Ensuring that assessments are completed early enough in the school year that teachers are able to use the data in planning instruction for their students may lead to less teacher resistance and more teacher buy-in to such assessments.
Two additional benefits stated by participants were the ability to disaggregate the data (86%) and to show student growth (57%). As part of the No Child Left Behind Act, schools are supposed to make incremental achievement targets each year to ensure one hundred percent proficiency in 2014. As the teachers noted, Measures of Academic Progress helps monitor students’ progress in a given year. If schools are allowed to administer Measures of Academic Progress, identify the achievement gaps for each student, and then remediate or provide enrichment when necessary, students should show academic growth. When students are assessed throughout the year with Measures of Academic Progress, effective teachers can make steady gains to reach individualized achievement targets.

A final benefit to Measures of Academic Progress stated by 29% of the participants is the way in which the test data can help evaluate classroom curriculum. One of the features of Measures of Academic Progress is the ability to disaggregate the skills assessed. During grade level collaboration, teachers are able to identify common areas in which a majority of the students in each classroom have struggled to grasp a specific concept. Measures of Academic Progress data allow teachers to evaluate their curriculum by looking at trends across grade levels in order to evaluate the effectiveness of their instructional materials. Further, if a school has multiple years of Measures of Academic Progress test data, strengths and weaknesses may reveal trends for individual teachers or grade levels. If a particular teacher’s students in a single grade level are consistently showing gains over those of other teachers at the same grade level, then the opportunity for peer coaching arises. Schools could use their test data to find the teachers
whose students are consistently a little bit higher in math and reading scores and evaluate what they are doing differently from their colleagues to increase achievement scores. Given the discomfort of some teachers in this study with the perception that assessment data could be used to assess teacher performance, it seems essential that this kind of analysis of test data and call for peer coaching be accomplished in a collaborative culture rather than a punitive one.

**Drawbacks.**

Analysis of perceived drawbacks of the Measures of Academic Progress assessment revealed three concerns: unreliable test data, loss of routine, and excessive testing. Forty-three percent of the participants felt the test results were unreliable because the students could click their way through the test without even reading the questions. Multiple-choice tests are designed to give a quick snap shot of where a child is struggling or excelling. In order to assess higher order thinking that would eliminate pure guessing, the test format would have to be altered. Constructed response or essay format for both math and reading would require deeper understanding; however, these types of test would take longer to perform, and teachers expressed that they were already upset with the amount of time it takes to perform the multiple choice test given by Measures of Academic Progress. Evidence of students’ skills can be pulled from the test currently being administered. Some scores will not be accurate, but this is true for any test. Correlations between amount of time taken on test and final score can be reviewed to potentially provide teachers with information on the likelihood of students guessing rather than authentically attempting to answer the test questions. Data from the existing
assessment can provide teachers with valuable information for individual students. Informal assessments in the classroom can either support the test data or contradict them. Teachers who receive information that it is likely students guessed on the test may decide to focus more on in-class work than on the test results.

Teachers indicated that the scheduling of Measures of Academic Progress eliminates use of the computer lab for three weeks each semester. Since teachers and students do better with consistency, it could be useful to schedule testing when students come into the computer lab each week rather than creating an entirely new schedule for testing. If time constraints due to students taking the test prevent this from happening in the computer lab, the assessment could be taken in the rooms of individual teachers using the mobile lab. Nine weeks a year to perform a relatively short test seem excessive, and the loss of the computer lab during fall, winter, and spring trimesters, especially given the rising centrality of technology to teaching and learning, seems a legitimate and real concern.

All the participants in the study felt there is too much testing in general. Students are administered Measures of Academic Progress, District Reading Assessment, and District Writing Assessments three times a year. Benchmarks for every math and reading skill on the progress report are assessed by both formal and summative assessments. Students are frequently taking some type of test, and these tests take time to administer, time that is needed to teach the new skills required at each grade level. While assessment in the fall is seen as being crucial for creating a base line for every student, teachers indicated there is an abundance of district assessments and state assessments that take
place in March. Eliminating the winter testing of Measures of Academic Progress might allow for time in the classroom to instruct those students who are still struggling. It is possible that having teachers at all grade levels document on a weekly basis the type of test students are given as well as the time the tests take may provide the administration a clearer picture of the impact of assessment and accountability measures at the classroom level.

*Instructional Strategies*

The aims of the research were to reveal if and how teachers use assessment data to make instructional decisions. The next section highlights some of the techniques this high performing school has utilized that may impact student learning.

The ability of the Measures of Academic Progress to be broken down by specific skills is very advantageous. All of the participants indicated that they first use the disaggregated data to determine who needs help and who needs enrichment. The current literature indicates that in order to drive instruction and provide interventions for improvement, test data need to be disaggregated to locate specific skills in which students are struggling (Nelson & Eddy, 2008). Teachers are busy professionals, and it is evident from their responses that they rely heavily on disaggregated data to allow them to effectively manage their instructional time. Rather than concentrating on concepts students have already mastered, teachers at West Elementary quickly reviewed the disaggregated data and adapted their teaching around the strengths and weaknesses of their students.
When participants were asked how they disaggregate the data, 29% stated that they use it to evaluate the efficacy of the curriculum in the classroom. This was also mentioned earlier in the interviews as a benefit of Measures of Academic Progress. Teachers at West Elementary seemed very concerned if their students did not achieve academic benchmarks and were constantly searching for ways to help them succeed. Test data are a major source for providing those strategies. Educators in this study were consistently looking at data to drive instruction and to evaluate if the curriculum and their teaching strategies achieved the necessary learning outcomes for their students to be successful.

West Elementary has a proven history of on-going high performing status. Since the purpose of this study was to reveal if and how teachers use assessment data to make instructional decisions, I specifically asked the participants to describe how they modify their teaching practices based on the data.

**Teaching modifications.**

Meeting the needs of diverse abilities can be a challenge. In order to address multiple academic levels, 57% of the participants stated that they ability grouped their students based on the test data. This differentiation allowed target instruction for groups of children with similar needs. Providing differentiation was one of the strengths of the school, and it was performed on a regular basis not only by individual teachers but also through Professional Learning Communities. As part of the process, West Elementary was able to restructure the day to utilize all support staff for intense remediation or enrichment as directed by the classroom teachers two times a week. This strategy not
only served to improve academic achievement but also promoted cooperation among colleagues to help students acquire the necessary skills to be successful.

In order to specifically address all academic levels, 29% of the teachers indicated using Compass Learning folders. Compass Learning folders consist of test data imported from Measures of Academic Progress that provide students with additional instruction and activities on the computer. Students are allowed to work through each of the folders at their own rate as they master each concept. Teachers who used Compass Learning felt that it was the best way to differentiate since it is so personalized. It is possible that the reason more teachers did not utilize this resource is because they were intimidated by using computers in the classroom. Having the computer knowledge to fix random technological glitches may hinder learning causing some to avoid this computer resource altogether.

Another modification utilized by 29% of the participants was the DesCarte binder. This resource explains the concepts covered in each RIT band and helps guide instruction. Perhaps more teachers would utilize the DesCarte binder if there were links to actual lessons for each skill. Most teachers know what their students are lacking; sometimes they just do not have the curriculum to teach it effectively. The time it takes to find a lesson for a specific skill is time teachers do not have. Providing lesson ideas for each of the skills might encourage more teachers to utilize the binder more frequently.

An additional strategy based on student needs from Measures of Academic Progress test data was a classroom management technique. This method involves keeping students identified as low performing on task. Classroom behavior is very
important in the learning environment. Students who are consistently attentive and are able to complete their schoolwork demonstrate greater academic gains. For some students, the teacher may need to enforce those good work habits that create ability.

**Improvements to Measures of Academic Progress.**

Teachers at West Elementary consistently utilized test data to make instructional decisions. In an effort to improve teaching strategies, participants in the study felt modifying software capabilities, expanding ways to generate reports, creating lesson links, increasing efficiency at the district, and providing teacher in-service would lead to better-informed instructional decisions.

Forty-three participants were unhappy with the current software capabilities of Measures of Academic Progress. The program is primarily designed for a PC computer, and some of the growth reports can only be run on a PC. Many elementary schools utilized the Macintosh computer. It would seem advantageous for Measures of Academic Progress to be capable of performing the same reports on dual platforms. This would allow Macintosh users the same opportunity to utilize valuable reports in their decision process.

Another suggestion to improve Measures of Academic Progress involved expanding the types of reports available. At the time of this study, test data were only available for individual classrooms. Participants expressed the need to combine test data across a particular grade level or by individual students in addition to the assigned teacher reports in order to show growth or areas of weakness. By isolating individual students over multiple years of testing, teachers may see trends in which the student makes large
academic gains or losses. In order to obtain this information, the school must create their own reports in which student data are imported manually. This is not a good use of time, yet this elementary school felt it was necessary in order to capture the overall picture. Perhaps this limitation is for confidentiality reasons. If so, designating administrative privileges would save valuable time when creating these informative reports.

Participants felt the data could be user-friendlier if there were specific lessons tied to the RIT band scores. Northwest Evaluation Association provided Compass Learning folders that allowed students to complete activities for each of the RIT bands. The drawback to this resource is that it was done entirely on the computer individually. Since the activities were computer based, teachers may have had a difficult time assigning supplemental work for their students who needed additional practice at home. In some cases, there were no substitutes for paper pencil exercises. Additionally, some teachers may not have felt comfortable teaching students on computers due to a lack of computer knowledge. These factors, along with monitoring students as they self-paced though the different activities, may have prevented teachers from utilizing it in their classroom.

The final improvements for Measures of Academic Progress were both suggestions for the District. These included quicker access to test data and more in-service. When students completed their Measures of Academic Progress test, their overall score was immediately printed. In order to obtain the disaggregated data, teachers had to wait for the district to upload the information. Since the entire district took the test at the same time, this could take time to accomplish. The ability to use the test data to make immediate decisions concerning a child has the potential of making a huge impact
on learning. Access to this information is what teachers regarded as one of the biggest benefits. Perhaps hiring additional personal at peak testing periods would help remedy this issue and eliminate teacher frustration.

Teachers also expressed a need for additional in-service demonstrating how to run the reports offered by Measures of Academic Progress. Education has changed drastically in the last ten years. In the past, teacher preparation did not include analyzing test data. Measures of Academic Progress testing has made it easier for schools to understand assessment results by disaggregating the data by skills. Unfortunately, many educators do not have the abilities needed to interpret the data in such a way that will best impact students. Sometimes the way the information is presented helps highlight different strengths and weaknesses. Providing in-service, especially after test data are available, could prove to be extremely beneficial. This opportunity would allow assistance with computer issues that teachers may encounter when running reports, as well as a chance to explore the data in order to make informed decisions.

*Connections between ongoing assessment data and annual yearly progress.*

West Elementary has never failed to make annual yearly progress since the enactment of No Child Left Behind of 2001. Because this school has been so successful, I asked participants to share whether they perceived a connection between ongoing assessment data and yearly progress. All of the participants felt ongoing assessment data were useful to reach growth targets. Measures of Academic Progress test data allowed teachers to identify areas of weakness, make instructional adjustments, and then retest to check for understanding before high stakes tests were administered at the end of the year.
Access to test data early in the year may empower teachers with the necessary information to plan instruction that will impact learning. If teachers see their students making academic gains based on data driven decisions they have made, than educators may see the value in all of the testing being administered.

*Participants’ final thoughts on assessment.*

Participants were invited to share any final comments in regards to assessment. The ways in which assessments are used, the need for more of a balance in testing, the ability of test data to drive instruction, and the way in which teachers’ attitudes can affect test performance were the points expressed.

Teachers in the study stressed their unhappiness in regard to how assessments were being used. Many feel it was unfair to evaluate a teacher or a school based on how well a child performs on a single test. Rather than focusing on the final outcome, more attention may need to be given to the process in order to eliminate some of the resentment felt by teachers. If educators were evaluated on how much growth students made in a given year, then possibly educators would not feel they were being treated unfairly. It appears that teachers are not against testing but rather against the idea that all students will be at a certain level at a specified time. Many students will reach the designated target, but if students are continually showing growth and need the gift of time, their growth should be celebrated. Sometimes the child who needs to make the most growth is showing progress, but just not at the same rate as other students.

Forty-three percent of the participants expressed a need for a balance in terms of testing. While many of the teachers agree that testing is an essential part of the learning
process since it gives teachers an idea of what the students have mastered, many of the participants felt more time is spent on assessment than instruction. School districts may need to evaluate how much testing is currently taking place in the classroom and decide which tests are the most valuable. If not, teachers will be forced to teach only what is on the test, eliminating some of the productive and essential activities that offer solid creative thinking and educational opportunities and that motivate students to want to learn.

Twenty-nine percent of participants shared that assessment is an excellent way to drive instruction in order to increase student achievement on standardized assessments. Making decisions based on test data and then progress monitoring allows teachers to assess students’ academic performance and determine if they are receiving the appropriate instruction to accelerate learning. Effective teachers who assess regularly can make informed instructional decisions to guide students on their journey to attaining academic proficiency. Documenting interventions and individual student growth may be a better way of holding schools and teachers accountable.

The final thought on assessment the participants shared was the idea of teachers’ attitude towards assessment in terms of having either a positive or negative impact on test results. Educators know that a student’s attitude in class is very important. The influence a teacher has over a child is powerful. If students feel their teacher has a negative attitude, that attitude could have detrimental consequences on test results. Enthusiasm toward instruction could not only help students learn more but help them retain information longer. This could be a potential benefit when tests are administered as well.
Summary

The experienced teachers who participated in this research regularly administered Measures of Academic Progress to their students in order to show academic growth. Many perceive the data are being unfairly used to evaluate not only the students but the teachers as well. Participants expressed how beneficial Measures of Academic Progress is in regards to instruction since the disaggregated results were quickly made available. Major drawbacks consisted of the loss of scheduled computer lab time due to Measures of Academic Progress testing and the possibility of skewed test results due to guessing. Despite these negative factors, the test information was viewed as useful since it allowed teachers to make instructional decisions that accelerate learning. Ability grouping, Compass Learning folders, the DesCarte binder, and classroom management techniques were all mentioned as modifications to increase students’ learning. Improvements to Measures of Academic Progress included creation of dual platforms for access to data, more flexibility when running reports, and lesson links tied to RIT band scores. Participants felt the district should be able to upload the data quicker and provide in-service on running the available reports necessary to drive instruction. Participants perceived a connection between ongoing assessment data and annual yearly progress since test data allow teachers to make instructional adjustments that enhance learning. Assessment is needed, but teachers in this study believed students are generally tested too much, almost to the point that there is more time spent on assessment than actual instruction. In an era of excessive testing, teachers’ attitudes may influence test performance. The next chapter offers an examination of the limitations of the study,
conclusions, and implications for future research.
This qualitative study was developed out of an interest to help empower other teachers and schools regarding the best practices to demonstrate academic growth based on test data in terms of a way to make annual yearly progress. The school chosen for this study had consistent testing success. Each year, West Elementary reached their projected benchmarks of proficiency when other schools in the district using similar programs struggled. When Measures of Academic Progress data from the last three years in grades 3 through 5 were reviewed, the average growth was consistently higher than national averages. The ability of this successful school to consecutively make annual yearly progress when so many other educational agencies were unable to do so prompted me to investigate the factors that may have been influential.

Conclusions

Participants in this study currently use data to drive instruction in their classrooms. Unlike end-of-the-year state assessments, Measures of Academic Progress provides disaggregated feedback very quickly, which empowers teachers to recognize individual strengths and weaknesses. Knowing where a child is struggling or even excelling allows educators the ability to make instructional decisions that accelerate learning. It is recommended that teachers regularly evaluate test data to determine the best course of action to enhance learning. Ability grouping according to needs seems to be the action of choice at West Elementary. Some of the teachers utilize Compass Learning folders and the DesCarte binders. Both are valuable resources that help
teachers concentrate on what their students need. Since some of the teachers did not indicate using these two resources and still achieve high-test scores, I believe the resources one uses are not as important as the instructional method. All of the teachers at West Elementary identify strengths and weaknesses from test data and provide target instruction. In order to provide differentiation, these teachers ability-group their students to meet individual needs. These professionals also monitor students’ progress frequently to ensure that their students are showing academic growth and are on the path to educational proficiency.

Limitations of the Study

This study was conducted during the summer of 2013. The number of teachers meeting the requirements of administering the Proficiency Assessment for Wyoming Students and Measures of Academic Progress testing at West Elementary limited the participants to just the third, fourth, and fifth grade teachers as well as the school’s Professional Learning Community Coordinator. Unfortunately, only seven of the ten participants who were invited returned their consent forms. Others who did not participate conveyed that they were unavailable for an interview due to scheduled summer vacations, family commitments, or enrollment in extended education courses. The data that were utilized in this study were taken from experienced teachers; however, a larger sample size drawn from other successful schools would enable researchers to make more generalizations.

An additional limitation has to do with the relationship of the researcher and the participants from West Elementary. Since the researcher had a professional relationship
with the participants, it is possible that participants could have felt uncomfortable talking about the ways in which they used or did not use the test data in their teaching practice with a colleague. They could also have had negative perceptions about the move toward data-driven instruction and standardized testing and may have had concerns about revealing those perceptions.

A final limitation deals with the qualitative nature of the study. Interviews were transcribed and coded solely by the researcher. In future studies, potential biases could be eliminated with the help of additional researchers. Allowing each person to code the responses separately and then compare the work would yield stronger results.

**Implications for Future Research**

Two issues that arose in this study that I feel need further investigation are centered on classroom behavior and teachers’ attitudes toward testing. During one of the interviews, a participant noted how he continually calls on identified students with low-test scores in order to increase their attentiveness, the rationale being that the more attentive students are, the more likely they are to learn new skills. It might be interesting to study the relationship between the amounts of time on task during classroom instruction compared to test scores. Effort certainly can create ability, and this is just one way teachers at West Elementary use their experience in the classroom to improve their test scores.

When participants were allowed to add final comments, one individual shared the idea of teachers’ attitude towards assessment in terms of having either a positive or negative impact on test results. The influence a teacher has on students is powerful.
Given the increased amount of assessments that teachers currently administer and the various attitudes regarding those assessments, it would be worthwhile to know if teachers’ negative attitude towards administering an assessment impacts students’ results.
APPENDIX A

Interview Questions

1. How many years have you been in education?
2. How many years have you been employed at Lizard Elementary School?
3. In what capacity are you employed at Lizard Elementary School?
4. What are your experiences with toward MAP testing?
5. How do you feel about the utility of MAP testing, and why?
6. What do you see as the benefits of MAP testing, and why?
7. What do you see as the drawbacks of MAP testing, and why?
8. What are your perceptions of how often the school tests its students?
9. On what do you base those perceptions?
10. How quickly do you receive the MAP data?
11. How do you disaggregate the MAP data?
12. Specifically, how do you use the MAP data?
13. Are the MAP data beneficial in terms of modifying your teaching practices to better your students’ learning outcomes? If yes, please give some examples of modifications you have made that have resulted in positive learning outcomes. If not, please explain what keeps the data from being beneficial to you.
14. In what ways could the MAP data be more user-friendly and useful for you?
15. Do you perceive a connection between teachers using on-going assessment data and schools being able to meet AYP? If so, what is the connection, and if not, what keeps there from being a connection?

16. Do you have anything else you would like to add about assessment?
REFERENCES


