

THE RELATIONSHIPS AMONG MIDDLE LEVEL TEACHERS' ASSESSMENT
PRACTICES, INSTRUCTIONAL LEADERSHIP, AND STUDENT ACHIEVEMENT

A Dissertation

presented to

the Faculty of the Graduate School

University of Missouri – Columbia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

by

KRISTIN W. MATTHEWS

Dr. Jerry Valentine, Dissertation Supervisor

MAY 2007

© by Kristin W. Matthews 2007

All Rights Reserved

The undersigned, appointed by the Dean of the Graduate Faculty, have examined
a dissertation entitled:

**THE RELATIONSHIPS AMONG MIDDLE LEVEL TEACHERS' ASSESSMENT
PRACTICES, INSTRUCTIONAL LEADERSHIP, AND STUDENT
ACHIEVEMENT**

presented by Kristin W. Matthews, a candidate for the degree of

DOCTOR OF PHILOSOPHY and hereby certify that in their opinion it is worthy of
acceptance.

Dr. Jerry Valentine, Advisor
Educational Leadership & Policy Analysis

Dr. Margaret Grogan
Educational Leadership & Policy Analysis

Dr. Juanita Simmons
Educational Leadership & Policy Analysis

Dr. Betty Porter Walls
Educational Leadership & Policy Analysis

Dr. Richard Robinson
Learning, Teaching and Curriculum

Acknowledgements

To Dr. Jerry Valentine, who as my advisor, employer, mentor, teacher and friend, provided for me wonderful opportunities to learn about leadership, research, statistics, middle schools, assessments and how to survive on daily sub sandwiches through my work with the Middle Level Leadership Center. The time you spent working with me through the dissertation process was above and beyond and I truly appreciate all that I gained from you through the experience. You are a wonderful teacher in so many ways, and I will always think of you as I begin my career as an educational leader.

To Dr. Margaret Grogan, Dr. Juanita Simmons, and Dr. Betty Walls, for being role models for me as exemplary women in educational leadership. Dr. Grogan, thank you for pushing me to reach for high standards of excellence in all of my leadership work. Dr. Simmons, I thank you for the encouragement and words of advice. Dr. Walls, I am glad I had the opportunity to work with you through the St. Louis New Leaders Project. I learned so much about urban leadership and it has enhanced and shaped my educational philosophy.

To Dr. Richard Robinson, for teaching me about the history of literacy. I learned from you that throughout history, scholars have written about and known what is right for children in education. It is up to future educators to continue to bring to light how children best learn.

To my husband, Chris, for being so supportive, encouraging, and brave as we experienced this major change in our lives. I am proud of you for all that you have

sacrificed and accomplished as we went through this journey together. We will be forever changed because of this process and I can't wait to see where the future takes us.

To my children, Brooke and Zach, for your patience during the times when I was not around. You are both wonderful children, and I love you more than you will ever know.

To all of my colleagues and fellow graduate students at the University of Missouri, especially Bernard Solomon and Greg Mees from the Middle Level Leadership Center, and Candy Abadir and Martille Elias from the Literacy Department, for setting the bar for excellence in graduate students. I appreciate the time we spent laughing, working hard, and finding ways to avoid working hard. I wish you all good luck in your future endeavors.

THE IMPACT OF MIDDLE LEVEL TEACHERS' ASSESSMENT PRACTICES AND INSTRUCTIONAL LEADERSHIP ON STUDENT ACHIEVEMENT

Kristin Weiser Matthews

Dr. Jerry Valentine, Dissertation Supervisor

ABSTRACT

Purpose of the study. The purpose of this study was to analyze the relationships among teachers' assessment practices, instructional leadership, and student achievement in the middle level schools in Missouri. The method of analysis was primarily quantitative, with survey data being used to determine, (a) what is descriptive of the demographics of Missouri middle level schools, (b) what is descriptive of the assessment practices of teachers in Missouri middle level schools, (c) what is descriptive of the instructional leadership practices of the school leader, and (d) if any relationships exist between the teachers' assessment practices, the principal as instructional leader and student achievement. Supplementing the quantitative data were data obtained through written responses to open-ended questions.

Findings. The study found significant relationships among teachers' assessment practices and instructional leadership as measured through instructional improvement and curriculum improvement. One assessment practice that revealed significance was collaboration through assessment, which is the process when teachers and administrators come together and analyze assessment data to plan for instruction and discuss student achievement. Other practices were assessment during instruction, which describes the strategies that teachers employ before and during assessment, like planning for classroom

discussions, providing feedback, and adjusting instruction to meet all students' needs, and the use of summative assessments, which describes the analyses of final summative data at the building level, the teacher and classroom level, and the individual student level as a basis for changes in curriculum, instructional practices and programs. A final variable, overall collective assessment practices, described teachers' perceptions of the collective ability of the faculty of their school to use assessment data to improve instruction and student achievement.

There were significant differences in middle level teachers' assessment practices for the factors and scale, Application of Formative Assessment, Providing Feedback to Students and Parents, and Use of Summative Assessments when schools were grouped according to student achievement scores on the Missouri Assessment Program. A model was developed illustrating the relationships among the instructional leadership factors and the teachers' assessment practices scales and factors. Implications for research and practice were discussed.

Table of Contents

Acknowledgements	ii
Abstract.....	iv
Table of Contents.....	vi
List of Tables	x
List of Figures.....	xiv
Chapter 1: Background of Study.....	1
Introduction.....	1
Statement of the Problem.....	4
Purpose of the Study.....	7
Research Questions.....	7
Hypotheses.....	8
Limitations.....	9
Definitions.....	9
Outline of the Study.....	11
Chapter 2: Review of Related Literature.....	13
Introduction.....	13
Historical Conceptions about Assessment.....	13
The Taylorism of the American Educational System.....	14
The Institutionalization of Testing in American Schools.....	18
The Modern Assessment Movement.....	21
Contemporary Conceptions about Assessment.....	23
Assessment in Education.....	24

Classroom Assessment and Student Achievement.....	26
Teachers' Assessment Literacy.....	31
Teacher Strategies for Classroom Assessments.....	37
Instructional Leadership.....	50
Instructional Leader Attributes and Job Functions.....	52
Instructional Leadership Knowledge.....	55
Teacher Development and Intellectual Stimulation.....	58
Teachers' Assessment Practices and Leadership.....	60
Conclusion.....	61
Chapter 3: Research Design and Methodology.....	64
Introduction.....	64
Purpose of the Study.....	65
Research Questions.....	66
Hypotheses.....	66
Population.....	67
Procedure.....	69
Instrumentation.....	70
Middle Level Teacher Assessment Practices Survey.....	71
Principal Instructional Leadership	73
Data Collection.....	74
Data Analysis.....	75
Chapter 4: Presentation and Analysis of Data.....	77
Introduction.....	77

Study Design.....	78
Research Questions.....	80
Hypotheses.....	81
Descriptive Findings.....	82
Demographic Data.....	82
Teacher Demographics.....	87
Leader Demographics.....	89
Factor Analysis.....	90
Hypothesis Testing.....	96
Hypothesis One.....	96
Hypothesis Two.....	112
Hypothesis Three.....	116
Hypothesis Four.....	121
Supplemental Data.....	136
Summary of Results.....	140
Chapter 5: Discussion of Findings.....	149
Introduction.....	149
Overview of Study.....	149
Research Questions.....	151
Hypotheses.....	152
Summary of Findings.....	152
Descriptive Results.....	152
Hypothesis Testing.....	153

Discussion of Findings.....	155
Implications.....	173
Implications for Practice.....	173
Implications for Research.....	176
Closing Thoughts.....	178
References.....	179
Appendix A: Principal Cover Letter.....	197
Appendix B: Principal Demographic Survey.....	200
Appendix C: Principal Consent Form.....	204
Appendix D: Secretary Letter.....	207
Appendix E: Teacher Cover Letter (Short Version).....	210
Appendix F: Teacher Cover Letter (Long Version).....	213
Appendix G: Middle Level Teachers' Assessment Practices Survey.....	217
Appendix H: Audit of Principal Effectiveness.....	222
Appendix I: Consent Form.....	229
Appendix J: Campus Institutional Review Board Approval.....	232
Vita.....	233

List of Tables

1.	Instructional Leadership responsibilities and their correlations with student academic achievement.....	54
2.	Middle Level Schools with Grades Five through Nine in Missouri.....	68
3.	Grade Level of Respondents and Statewide Comparisons.....	83
4.	Missouri Assessment Program Student Achievement Test, 2006 Communication Arts.....	84
5.	Missouri Assessment Program Student Achievement Test, 2006 Mathematics.....	85
6.	Total Averages of Study and Statewide Schools.....	85
7.	Comparisons of Respondents to State of Missouri, Selected Demographic Variables.....	86
8.	Teachers' Gender.....	87
9.	Assessment course taken for teacher certification.....	88
10.	Teaching Area.....	88
11.	Teaching Assignment.....	89
12.	Leadership Demographics.....	90
13.	Items Clustered by Scale.....	91
14.	Factors Identified and Loadings.....	92
15.	Factor Correlations with Free/Reduced Lunch Control.....	95
16.	Correlations: Middle Level Teachers' Assessment Practices Items for Instructional Preparation Practices with Audit of Principal Effectiveness.....	97

17.	Correlations: Middle Level Teachers' Assessment Practices Items for Assessment during Instruction with Audit of Principal Effectiveness.....	98
18.	Correlations: Middle Level Teachers' Assessment Practices Items for Adaptation of Instruction per Assessment Data with Audit of Principal Effectiveness.....	99
19.	Correlations: Middle Level Teachers' Assessment Practices Items for Feedback to Students and Parents with Audit of Principal Effectiveness.....	100
20.	Correlations: Middle Level Teachers' Assessment Practices Items for Development of Self and Peer Assessment with Audit of Principal Effectiveness.....	102
21.	Correlations: Middle Level Teachers' Assessment Practices Items for Use of Summative Assessment Data with Audit of Principal Effectiveness...	104
22.	Correlations: Middle Level Teachers' Assessment Practices Items for Overall Assessment Practices with Audit of Principal Effectiveness.....	105
23.	Correlations: Middle Level Teachers' Assessment Practices Factors with Audit of Principal Effectiveness.....	106
24.	Correlations: Middle Level Teachers' Assessment Practices Scales with Audit of Principal Effectiveness.....	107
25.	Correlations: Educational Improvement Factors from the Audit of Principal Effectiveness with Student Achievement.....	108
26.	Correlations: Middle Level Teachers' Assessment Practices Factors with Student Achievement on Communication Arts.....	110
27.	Correlations: Middle Level Teachers' Assessment Practices Factors with Student	

	Achievement on Communication Arts.....	111
28.	Linear Regression: Instructional Leadership Factors with Collaboration through Assessment (Factors 4 -MLTAPS).....	113
29.	Linear Regression: Instructional Leadership Factors with Assessment during Instruction (Scale 2 – MLTAPS).....	114
30.	Linear Regression: Instructional Leadership Factors with Use of Summative Assessment Data (Scale 6 – MLTAPS).....	115
31.	Linear Regression: Educational Improvement Factors with Overall Assessment Practices	116
32.	Linear Regression: Instructional Leadership Factors and Middle Level Teachers Assessment Practices with Student Achievement in Communication Arts	119
33.	Linear Regression: Instructional Leadership Factors and Middle Level Teachers Assessment Practices with Student Achievement in Mathematics	121
34.	Test of Differences: Teachers Assessment Practices Factors with Communication Arts (Quartiles: 4 High – 1 Low).....	123
35.	Test of Differences: Teachers’ Assessment Practices Factors with Mathematics (Quartiles: 4 High – 1 Low).....	125
36.	Test of Differences: Teachers’ Assessment Practices Scales and Communication Arts (Quartiles: 4 High – 1 Low).....	127
37.	Test of Differences: Teachers’ Assessment Practices Scales and Mathematics (Quartiles: 4 High – 1 Low).....	130

38.	Test of Differences: Educational Improvement Factors with Audit of Principal Effectiveness Survey and Communication Arts.....	133
39.	Test of Differences: Educational Improvement Factors from Audit of Principal Effectiveness Survey and Mathematics.....	135
40.	Average Percentage of Students Receiving Free or Reduced Price Lunch by Achievement Quartile.....	136
41.	Summary of Partial Correlation Findings.....	141
42.	Summary of the Influence of Independent on Dependent Variables from the Multiple Regression Findings.....	143
43.	Differences among the Student Achievement Quartiles for the MLTAPS Factors.....	144
44.	Correlations for Collaboration through Assessment, Use of Summative Assessment Data, Assessment during Instruction and Overall Assessment Practices controlling for Free/Reduced Lunch.....	146
45.	Linear Regression: Middle Level Teachers' Assessment Practices – Collaboration through Assessment, Assessment during Instruction and Use of Summative Assessment with Overall Assessment Practices.....	147
46.	Differences among Student Achievement Quartiles for MLTAPS Factors.....	168

List of Figures

1.	An historical overview illustrating how changing conceptions of curriculum, learning theory, and measurement explain the current incompatibility between new views of instruction and traditional views of testing.....	3
2.	Shared principles of curriculum theories, psychological theories and assessment theory characterizing an emergent, constructivist paradigm.....	4
3.	Teachers' classroom assessment decision making.....	36
4.	Correlations and Regressions between Instructional Leadership Factors, Instructional Improvement and Curriculum Improvement, and Middle Level Assessment Practices Scales and Factors, Collaboration through Assessment, Assessment during Instruction, Use of Summative Assessment Data and Overall Beliefs about Assessment Practices.....	145
5.	Correlations and Regressions for Instructional Leadership and Assessment Practices on Overall Beliefs about Assessment Practices.....	148
6.	Correlations and Regressions between Instructional Leadership Factors, Instructional Improvement and Curriculum Improvement, and Middle Level Assessment Practices Scales and Factors, Collaboration through Assessment, Assessment during Instruction, Use of Summative Assessment Data and Overall Beliefs about Assessment Practices.....	163
7.	Correlations and Regressions for Instructional Leadership and Assessment Practices on Overall Beliefs about Assessment	

Practices.....	165
----------------	-----

CHAPTER 1

BACKGROUND OF THE STUDY

Introduction

In the era of No Child Left Behind (2002), educators are being held accountable on many levels. Nearly every state has established standards that determine what students should know and be able to do. Also, most states have developed official standardized tests aligned to those standards (Popham, 2001). Tensions arise among teachers, students, parents, and policymakers because the tests are used for differing goals: to check if students learned important content standards, and to know how schools and students rank locally, statewide, and nationally (Sloane & Kelly, 2003).

As a result of the conflicting goals of standardized testing, the increasing focus on school and student ranking has lead to many external stresses beyond the classroom. For instance, teacher and administrative salaries have been tied to official test scores (Solmon & Podgursky, 2000). At the district level, school boards have demanded increases in official test scores (Popham, 2001). At the state and national level, politicians, such as California's ex-governor, Gray Davis, have staked their careers on the improvement of test scores (Fienberg, 2002). In each of these instances, teachers and students become the focal point for improvement with fear of public failure used as the incentive (Stiggins, 1999). This use of "scoreboard-induced motivation" (Popham, p. 12) in schools to increase student achievement seems to detract focus from the other and, perhaps, most important goal of testing – checking for evidence of student learning.

Many well-intentioned principals and teachers have emphasized standardized testing strategies due to the pressure to achieve high test scores; however, research suggests that teachers will often skew their efforts in the direction of activities that would lead to increases in these highly publicized scores (Earl, et al., 2003). Furthermore, when instructional time is only focused on standardized testing skills, increases in scores are not sustained and tend to drop when students move to the next grade or class (Darling-Hammond & Wise, 1985).

Despite the external pressures that accompany standardized testing, many researchers have encouraged educators to embrace standardized testing and align curriculum to ensure that tested objectives are taught but with an emphasis on teaching the child, not the test (Stiggins, 2004; Darling-Hammond & Wise, 1985). Stiggins (2002) calls this “assessment for learning” in which teachers use classroom assessment practices to guide instruction and work to include students in the assessment process. Examples of classroom assessment practices include using assessment data to frame performance goals, providing learning criteria to students in advance, providing feedback early and often, encouraging student self- and peer-assessment, and allowing new evidence of achievement to replace old evidence (McTighe & O’Connor, 2005).

The conflicting purposes for educational testing are not a new phenomenon in education. Beliefs from the early 1900s about learning testing are deeply rooted in the American educational system and have influenced traditions and practices that are still seen in twenty-first century schools. Early ideas were centered on the notion that learning is tightly sequential and hierarchical and tests were isomorphic with learning (Shepard, 2000a). Modern learning theorists have changed their description of learning to an active

process of mental construction and sense making that is developed through socially supported interactions (Shepard, 2000b), yet testing systems remain firmly entrenched in a stimulus-response framework. In fact, Graue (1993) pointed out that classroom testing and instruction have become completely disconnected and separate causing testing practices to become a barrier in the implementation of more constructivist approaches to education (see Figure 1).

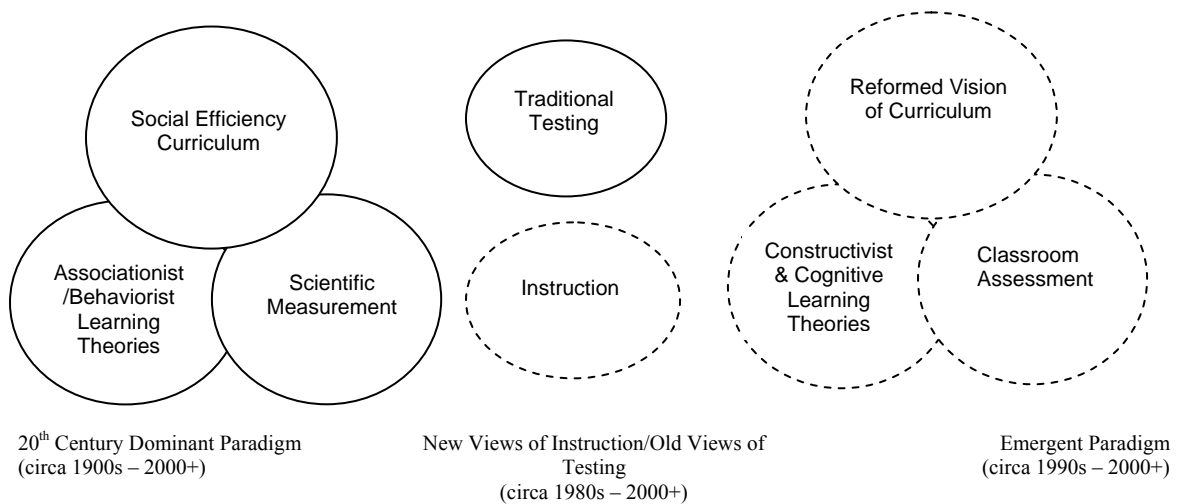


FIGURE 1. *An historical overview illustrating how changing conceptions of curriculum, learning theory, and measurement explain the current incompatibility between new views of instruction and traditional views of testing.* (Shepard, 2000a, p. 5).

Shepard’s Emergent Paradigm (2000a) from Figure 1 represents the framework for this study on the impact of teachers’ classroom assessment practices and instructional leadership on student achievement in which learning is believed to be constructed in a social context (Vygotsky, 1978), shaped by prior knowledge (Resnick & Klopfer, 1989), and involves “metacognition” or self monitoring of learning and thinking (Brown, 1994). Additionally, Shepard (2000b) proposed that curriculum and assessments should be

overlapping and combined to provide thoughtful and authentic learning experiences.

Figure 2 depicts an expanded description of the Emergent Paradigm.

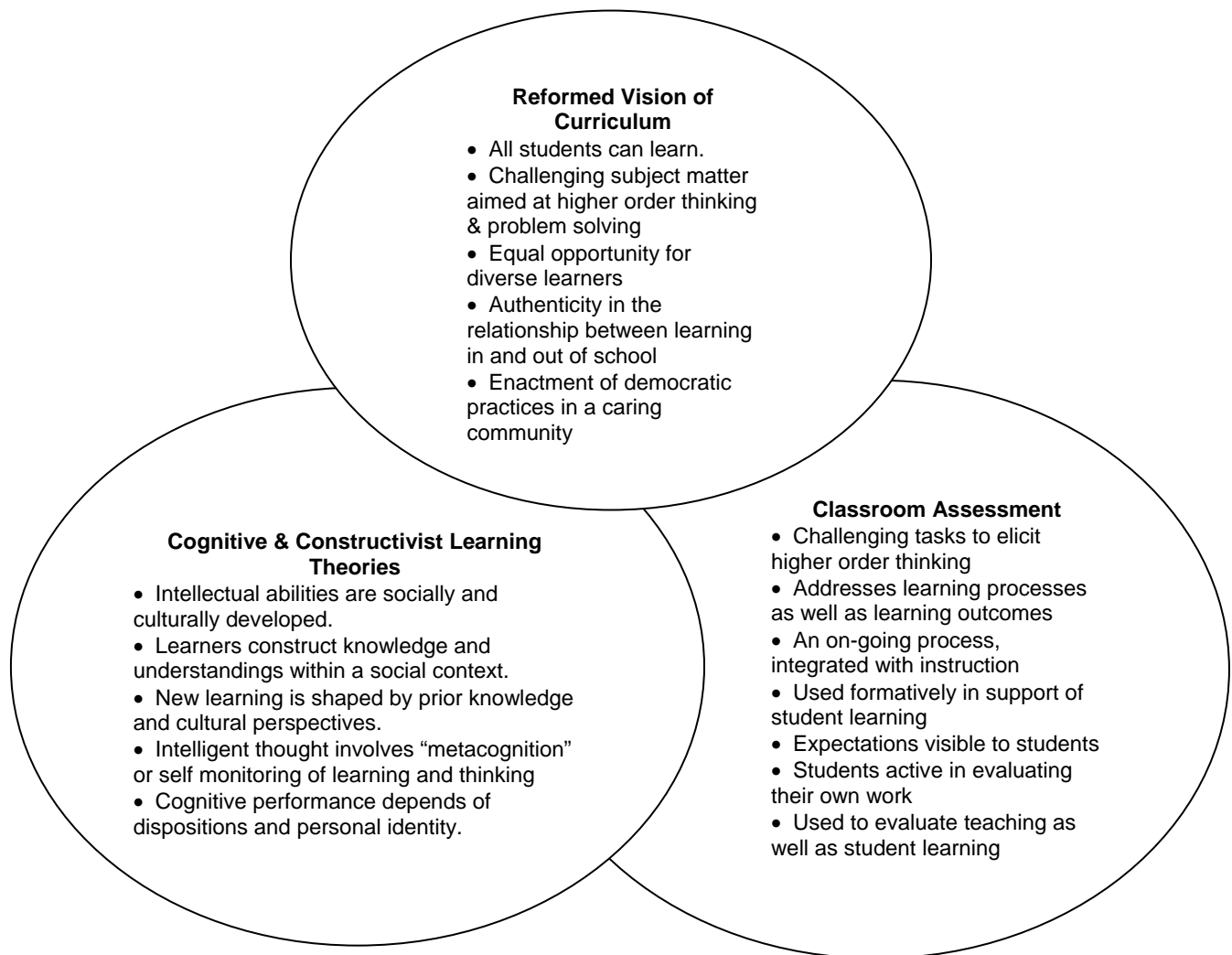


FIGURE 2. *Shared principles of curriculum theories, psychological theories and assessment theory characterizing an emergent, constructivist paradigm.* (Shepard, 2000a, p. 8).

Statement of the Problem

Traditionally, official standardized testing and classroom assessments have been considered as separate entities (Noonan & Renihan, 2006). More recent studies have begun to consider whether the two notions should be more integrated as a means to: (a)

improve teachers' assessment literacy - or levels of knowledge, appreciation and skills concerning assessment processes, alternatives, and uses (McMillan, 2003), (b) increase instructional leadership among administrators (Noonan & Renihan), and (c) boost student achievement as measured by standardized test scores (Black, et al., 2004).

It is important to understand the difference between 'assessments' and 'testing' (Herbert, 1997). The word 'assessment' is a more far-reaching term than the word 'test'. Assessment is defined as a judgment of the measure of a student's achievements. Achievements can be seen as gaining control of knowledge, skills or understanding of any piece of work. Achievements can be shown at any time, proving that assessment does not mean the same as a test. A test is the product that is administered at the end of a unit or course of study, such as statewide standardized tests. Types of assessments include a variety of tasks, questioning, tests, activities, and observations by teachers and pupil's peers, as well as the student (Herbert).

Why look to classroom assessment practices as a means to improve student achievement as measured by official standardized test scores? Students spend vastly greater amounts of time engaged in classroom assessment activities than in standardized testing; therefore, a greater impact is made through classroom assessments (Crooks, 1988; Stiggins & Bridgeford, 1985). In fact, teachers spend at least one third of their professional time on assessment activities that inform a wide variety of decisions made daily and those decisions directly influence students' learning experiences (Stiggins & Conklin, 1992).

Stiggins (1999) emphasized the two main reasons for using classroom assessment practices were to "keep students from losing confidence in themselves as learners and to

rekindle confidence among those students who have lost that confidence” (p.6). He continued to explain that academic failures are not to be blamed as self-concept issues, such as low self-esteem, academic failures are classroom assessment issues. Students must first experience academic success in the classroom, and then increased confidence will follow (p. 8).

Numerous studies on classroom assessment have demonstrated tremendous gains in student achievement as measured by teacher/researcher observations, summative assessment data, like the Third International Mathematics and Science Study (TIMSS) and student self-perception survey scores (e.g. Rodriguez, 2004; Black & Wiliam, 1998b; Bloom, 1984). If classroom assessments predict growth in various forms of student achievement, can classroom assessment practices predict growth in student achievement as measured by official standardized tests? If so, what particular aspects of the classroom assessment process are the best predictors? Do teachers understand the relationship between classroom assessments and standardized testing? Can classroom assessment strategies be mandated by external forces such as state and local school administrators and have positive results? Can the principal make a lasting impact on teachers’ assessment practices which in turn increases student achievement on statewide assessments?

Black and Wiliam (1998a) claimed educators and policy makers treat the classroom as a ‘black box’. Certain inputs — students, teachers, standards, tests, parental worries, and so on — are placed the box. Outputs are expected like higher test scores, students who are more knowledgeable and better teachers. Black and Wiliam’s extensive literature review focused on the ‘black box’, in other words, the classroom. Similarly, this study is to trying to pursue “what is happening inside the black box” relating to teachers’

classroom assessment practices, the principal's influence on teachers' assessment practices, and the resulting impact on student achievement as measured by official standardized tests.

Purpose of the Study

The purpose of this study was to investigate the relationships among teachers' classroom assessment practices, principals' instructional leadership practices, and student achievement as measured by official standardized tests. Teachers' classroom assessment practices were measured by items, scales and factors describing existing assessment practices. Instructional leadership practices were measured by two factors describing the degree of instructional and curriculum leadership of the principal.

Research Questions

To accomplish the purpose of this study, the following research questions were developed:

1. What are the types and levels of assessment practices used in Missouri's middle level schools according to teachers' perceptions?
2. What are the levels of instructional leadership for principals in Missouri's middle level schools according to teachers' perceptions?
3. Are there relationships between assessment practices and principal instructional leadership in Missouri's middle level schools according to teachers' perceptions?
4. Are there relationships between the types and levels of assessment practices and the levels of principal instructional leadership on student achievement in communication arts and mathematics according to teachers' perceptions?

5. Are there differences in the levels of principal instructional leadership and the types and levels of assessment practices between schools with high levels of achievement and schools with low levels of achievement in communication arts and mathematics?

Hypotheses

The following hypotheses were tested in this study:

H_{01} : There are no significant correlational relationships among the items, scales, and factors measuring teachers' perceptions of assessment practices, the factors measuring teachers' perceptions of principal instructional leadership and student achievement as measured by the Communication Arts and Mathematics portions of the Missouri Assessment Program.

H_{02} : There are no significant predictive linear relationships for the factors measuring teachers' perceptions of principal instructional leadership, and the scales and factors measuring assessment practices.

H_{03} : There are no predictive linear relationships among the scales and factors measuring teachers' perceptions of assessment practices and the factors measuring teachers' perceptions of principal instructional leadership on student achievement in communication arts and mathematics.

H_{04} : There are no significant differences for the scales and factors measuring teachers' perceptions of assessment practices and the factors measuring teachers' perceptions of principal instructional leadership between schools with high levels of achievement and schools with low levels of achievement in communication arts and mathematics.

Limitations

The following are the limitations of the study:

1. The findings of this study are limited to the validity and reliability of the instrument used to collect data.
2. The findings of this study are limited to the accuracy and perception of the teachers who completed the instruments.
3. The findings of the study are subject to the limitations associated with survey data collection.
4. The population was limited to middle schools in the state of Missouri that had a grade configuration between grades five through nine, included at least two grade levels and one of those grade levels had to be either grades seven or eight.
5. The concept of student achievement is broader than academic achievement. In this study, however, the measures of student achievement are solely academic measures, including student achievement on state-wide assessments in communication arts and mathematics.

Definitions

Assessment: The process of gathering, interpreting, recording, and using information about pupils' responses to educational tasks (Lambert and Lines, 2000, p. 4).

Assessment Literacy: Teachers with strong assessment literacy possess a working knowledge of when and how to design, develop, use and value a wide variety of methods for assessing student achievement (Stiggins & Conklin, 1992).

Classroom Assessment: Formative and summative assessments that occur in the classroom environment (Stiggins & Conklin, 1992).

Communication Arts: The standards associated with the subject that are formally assessed on the Missouri Assessment Program (MAP) are: a) speaking and writing standard English; b) reading fiction, poetry, and drama; c) reading and evaluating nonfiction works and material; and d) writing formally and informally (Missouri Department of Elementary and Secondary Education, 2006).

Curricular Improvement: The principal promotes an articulated, outcome-based curriculum through diagnosis of student needs and systematic program review and change (Valentine & Bowman, 1984).

Formative Assessment: Formative assessment refers to all those activities undertaken by teachers, and by their students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged (Scriven, 1967).

Feedback: Teachers must provide feedback to students to inform their level of performance on assessment tasks. Feedback is defined as having three elements: redefinition of the desired goal, evidence about the student's present position, and some understanding of a way to close the gap between the two. All must be understood to some degree before the student can take action to improve learning. (Black & Wiliam, 1998a; Sadler, 1998).

Instructional Improvement: The principal affects instructional skills positively through effective clinical supervision, knowledge of effective schooling, and commitment to quality instruction (Valentine & Bowman, 1984).

Mathematics: The standards associated with the subject that are formally assessed on the Missouri Assessment Program (MAP) are: a) addition, subtraction, multiplication

and division; other number sense, including numeration and estimation; and the application of these operations and concepts in the workplace and other situations; b) geometric and spatial sense involving measurement, trigonometry, and similarity and transformations of shapes; c) data analysis, probability and statistics; d) patterns and relationships within and among functions and algebraic, geometric and trigonometric concepts; e) mathematical systems; and f) discrete mathematics (Missouri Department of Elementary and Secondary Education, 2006).

Official Assessment: Official assessments are formal and systematic assessments that are required by the school bureaucracy for purposes such as pupil testing, grading, and placement (Airasian, 2000).

Standardized test: A test designed to be given under specified, standard conditions to obtain a sample of learner behavior that can be used to make inferences about the learner's ability. Standardized testing allows results to be compared statistically to a standard such as a norm or criteria. If the test is not administered according to the standard conditions, the results are invalid (SABES, 2006).

Summative Assessment: Summative assessments are used to evaluate the outcomes of instruction to determine whether learning occurred (Scriven, 1967) and take the form of tests, projects, term papers, and final exams (Airasian, 2000).

Outline of the Study

This chapter included background information and identifies the need for the study. Additionally, research questions, hypotheses, limitations and definitions related to the study are presented in this chapter. Chapter Two is a review of the literature relevant

to the history of assessment in the United States, assessment, classroom assessment, and instructional leadership.

Methods and procedures are presented in Chapter Three. Included is information about the population, instrumentation, data collection, and data analysis.

The results and statistical analysis of the study are presented in Chapter Four. Chapter Five includes a summary of the findings, conclusions, implications for practice, and recommendations for further research.

CHAPTER 2

REVIEW OF RELATED LITERATURE

Introduction

To understand the progression of assessment practices and beliefs, this review of literature was organized into four main sections: (1) historical conceptions about assessment, (2) current conceptions about assessment, (3) strategies for classroom assessment, and (4) instructional leadership. Section one, historical conceptions about assessment, begins with the beginnings of educational assessment and concludes with modern assessment practices in the United States. Section two, current conceptions about assessment, focuses on the research on assessments, classroom assessments and assessment literacy. Section three, strategies for classroom assessment, takes the previous literature on classroom assessment and narrows in on specific strategies listed as best practices for classroom assessment. Section four, instructional leadership, centers on the impact that the principal can have on teachers' assessment practices.

Historical Conceptions about Assessment

Assessment in education has a long tradition in the United States. Actually, there has never been in a time in the history of American public schools without some form of assessment in place (Cuban, 2004). During colonial times, oral examinations were administered to verify that children learned the prescribed material (Congress, 1992). In 1845, Horace Mann convinced the Boston School District to replace oral quizzes with written tests (Gerberich, 1963).

The tests were not intended to be used for comparisons among children and their schools; however, they were quickly utilized for that purpose. Long before the Civil War,

schools used mandated written examinations to assess student progress and to aid in policy decisions (Congress, 1992). In 1894, J. M. Rice demonstrated that tests of performance could be important for assessing instruction, curriculum, and student learning. He administered spelling tests in different school systems, looking to see if added instructional time improved student performance (Resnick, 1980; Gerberich, 1963).

The Taylorism of the American Education System

The Industrial Revolution brought social efficiency and scientific measurement to education (Shepard, 2000b; Resnick, 1980). Schools played a critical role in response to the changing job market brought on by the Industrial Revolution. More white-collar opportunities and rising standards created more accountability on the part of schools. In addition, schools had to answer to the acculturation of the largest wave of immigrants in the nation's history. With the growing population and the changing job market, public schools saw a huge increase in the 14 to 17 year age group in schools (Resnick). Researchers were influenced by public concerns about education that were shaped by industrialization, fears of loss of community, and the need to absorb and "Americanize" large groups of immigrants (Shepard, 2000a).

The organizational theorist, Frederick Taylor, developed his theories of scientific management during this time period. He believed that in order to achieve optimum efficiency, tasks were split into many separate elements that were to be observed and evaluated. These models of efficiency greatly influenced the industries on the United States (Morgan, 2006).

Tayloristic influences found their way into the American education system especially through the work of E. L. Thorndike. Thorndike was an experimental psychologist teaching at Columbia University (Teachers College) and he became actively involved in educational work (Resnick & Hall, 1998). Considered the “father of scientific measurement” in education, E. L. Thorndike fostered the development and dominance of the objective achievement tests, curriculum of social efficiency and associationist learning theory.

In 1904, Thorndike and his students published some of the first standardized achievement tests in handwriting and arithmetic reasoning (Gerberich, 1963). Around World War I, multiple achievement tests were produced and commercially marketed across the United States. The achievement tests were not completely used to diagnose student learning problems. Their main purpose was to permit comparisons of class and student performance in an effort to study education in a more scientific manner (Resnick, 1980; Shepard, 2000b). For instance, the need to remedy inconsistencies in teachers’ grading practices became important because it was thought to be highly inefficient. In a classic study that illustrated this point, Starch and Elliot (1913) distributed the same geometry paper to 116 high school mathematics teachers and received grades ranging from 28 to 92. Similarly, other assessments were created to find out why children were leaving school at eleven and twelve years old. In economic terms, twelve-year-old dropouts represented a “waste of scarce resources” (Cuban, 2004, p. 19).

Along with achievement testing, Thorndike also promoted a curriculum that focused on acquiring one skill at a time (Shepard, 2000a; Gerberich, 1963). Shepard (2000b) explained the movement:

This meant taking Taylor's example of a detailed analysis of the movements performed by expert bricklayers and applying similar analyses to every vocation for which students were being prepared. Every step would be taught specifically and precise standards of measurement would be needed to assess every step of the learning process. Since it was not possible to teach every student the skills of every vocation, scientific measures of ability were also needed to predict one's future role in life and thereby determine who was best suited for each vocation (p. 6).

Thorndike's view was that a core curriculum, concentrated in a few academic disciplines, made no sense for schools, especially at the secondary level where students were close to adulthood. Instead schools needed a vastly expanded array of curriculum options, differentiated both by student abilities and by projected future occupation and focused on the specific knowledge and skills that the student needed. Essentially, education became vocational training (Larabee, 2005).

Thorndike developed the associationist learning theory - which was a forerunner of Skinner's behaviorist learning theory (Shepard, 2000a). According to associationists, knowledge consists of a "collection of bonds: links between pairs of mental entities or between external stimulus and an internal mental response" (Resnick & Hall, 1998, p. 93). Learning involves strengthening the correct bond and decreasing the incorrect bonds. Bonds are strengthened through rewards and weakened through punishment or absence of rewards (Resnick & Hall).

The instructional practices that grew out of associationism called for instruction to be broken into singular minute steps, breaking educational processes into many steps that were reinforced through punishers or rewards. It also called for frequent testing, in which tests were made of many separate items of information based on the notion of knowledge as multiple bonds (Resnick & Hall, 1998). Textbooks adopted were composed of minimally connected bits of information with workbooks to support (Resnick & Hall).

John Bobbitt, a leader of the social efficiency movement, reflected the efficient beliefs of the day. He felt that the primary goal of curriculum design was the elimination of waste, and it was wasteful to teach people things they would never use. He felt that individuals should be educated according to his or her capabilities; therefore, testing became the tool of scientific measurement because it was necessary to determine students' abilities and intelligence (Larabee, 2005; Shepard, 2000b).

Just as we still see the effects of Taylor's work in American industries, the proponents of the objective achievement tests, curriculum of social efficiency and associationist learning theory created lasting effects in education. To this day, objective achievement tests are the typical method of delivery for testing. In fact, Senk, Beckmann, and Thompson (1997) found in a study of 19 mathematics classes 77% of students' grades were determined by singular math objective tests. The curriculum of social efficiency is still seen today through vocational schools and ability tracking of students. Finally, the stimulus-response beliefs that accompany associationist theories are definitely alive and well in modern education. Teachers use workbooks and textbooks that are published to disseminate information in fragmented sections of the curriculum.

The effects of this early era of American education have been extremely profound and lasting.

The Institutionalization of Testing in American Schools

The development of intelligence testing early in the century helped in the expansion of systematic testing in the United States. As the decades progressed, testing for various purposes became an institutionalized part of education.

The development of intelligence quotient (I.Q.) tests resulted in differing applications and beliefs between the United States and France. The French neurologist Alfred Binet had a strong influence in the development of the first I.Q. tests. He believed that intelligence was not a measurable trait, like height or weight, but that it could be improved through ‘mental orthopedics’ (Shepard, 2000b, p. 7). His research centered on the notion that intelligence was not delivered at birth, but that it varied with the age of the individual being observed (Congress, 1992; Shepard).

In 1917, the United States army asked Stanford professor Lewis Terman to produce intelligence tests and a group intelligence scale. Terman revised Binet’s I.Q. test, but with the philosophy that the I.Q. results were an exact measure of a fixed trait that couldn’t be raised or lowered throughout a person’s life (Congress, 1992; Shepard, 2000b). This later became the Alpha scale used by the Army to efficiently determine where to place recruits in jobs. The Alpha test was used for the normal and the Beta test was administered to the subnormal. Intelligence theorists poured over the data, proclaiming that a substantial proportion of American soldiers were “morons” (Congress). The idea to classify people in education was bolstered by this claim. After World War I, educators adopted the testing system under the guise of protecting the

“slow witted from the embarrassment of failure while allowing the gifted to rise to their rightful levels of achievement” (Congress, p. 110).

Because measured differences were taken to be innate, the only way to manage the differences was through a highly differentiated curriculum and tracking. By the 1920s, tracking students by their I.Q. was firmly in place in American schools (Shepard, 2000b). Educational administrators and researchers urged that students, most of them from minority racial groups, be segregated in special classes and taught in a concrete and practical manner that would make them efficient workers (Shepard).

In 1929, the University of Iowa initiated the first major statewide testing program for students. The Iowa Test of Basic Skills and the Iowa Test of Educational Development results were used to evaluate both students and schools. By the late 1930s, Iowa tests were being made available to other states (Congress, 1992). It was during this period that multiple choice and true/false questions were created for testing. The formats were quickly adapted to student tests and used in classroom assessments. The creators of the Iowa tests invented scoring machines that would make possible the “streamlined achievement testing of millions of students” (Congress). By the 1930s, multiple-choice tests were a part of public education.

Also in the 1930s, college admissions tests, like the Scholastic Aptitude Test (SAT), were implemented. The goal of this testing activity was to improve decision making about career choices and college entrance (Resnick, 1980). The tests quickly became national measures of school accountability, too. If the test scores were up, then the education system was considered to be doing well. If the scores dropped, then school

quality was considered to be low. During this period, higher test scores became equated with more productive schools (Stiggins, 1999).

By the 1930s, the vocational influence of the social efficiency movement had faded and most students took traditional academic courses, but there were lasting effects. Differentiation of the curriculum and the segregation of studies by gender and social class were well established in the United States educational system (Larabee, 2005; Shepard, 2000b).

In the 1940s, tests and test batteries were introduced. These tests were designed to measure general educational development rather than knowledge of a specific content subject (Gerberich, 1963). After World War II, the administration of standardized test batteries on a regular basis throughout the student's school career was a common educational practice (Resnick, 1980).

The launch of the Russian satellite, Sputnik, was the catalyst for President Dwight D. Eisenhower's signing of the National Defense Education Act, Title V-A in 1958. This contributed to the expansion and the development of testing programs in many schools through the country (Gerberich, 1963; Resnick, 1980; Cuban, 2004). Many reforms resulted, including raised graduation requirements in math and science, added programs for the gifted and advanced placement high school courses (Cuban). Unlike the educational reforms that came after World War II, which were mostly demographic, the post-Sputnik concerns were curricular, focusing on what was being taught and how, rather than who was being taught. "The finger of blame was pointed directly at the schools" (Rutherford, F., 1997, p. 2).

The Modern Assessment Movement

The sixties were the beginning of a new reform movement in educational testing (Resnick, 1980). According to James Popham (2001), there was a change in the public's perception of education. Stories were starting to surface about students graduating from high school without the ability to read or write and this influenced the public's discontent. This resulted in "mutterings" about requiring testing to ensure that learning occurred (p. 4-7). The district-wide standardized tests that we see today were created during the 1960s. The tests were commercially developed and norm-referenced in an effort to achieve local accountability (Stiggins, 1999). At the federal level, the Economic Opportunity Act, which established Head Start, and the Elementary and Secondary Educational Act of 1965, or Title I (Resnick, 1980) were passed in an attempt to address gaps in educational attainment (Guskey, 2005). These laws were required to be monitored and standardized achievement tests became the primary means for doing so (Koretz, 2002).

In the 1970s, accountability measures expanded to a statewide level. During the seventies and early eighties, the minimum competency testing movement spread rapidly (Linn, 2000; Resnick, 1980). The decade began with three statewide assessments and ended with nearly 40 (Koretz, 2002; Stiggins, 1999). Minimum competency tests were most often relatively easy multiple-choice tests used as a requirement for high school graduation (Koretz). The tests were established due to legislation that was passed with the objective of supplying parents with a "limited warranty" that a child who passed a competency test had at least mastered the fairly modest set of basic skills these tests measured. In addition, the policymakers who installed the competency tests weren't

focusing on students; they were actually displaying their doubts about public school educators (Popham, 2001, p. 12).

The 1980s brought in a transition to national testing (Stiggins, 1999). *A Nation at Risk* (National Commission on Excellence in Education, 1983) suggested that American students were outperformed on international academic tests by students from other industrial societies (Thattai, 2001). As Koretz (2002) attested, “The growing use of standardized tests for accountability was the core component of educational reform movement of the decade. New state-mandated tests were implemented, some tests were made more difficult, and the consequences of scores for students, educators, and administrators were often increased” (p. 760). Near the end of the decade, support for testing by researchers waned because of a growing awareness that coaching for these tests inflated scores and instruction was centered on the tests (Koretz, 2002; Shepard, 2000a), however, testing remained in place in educational systems.

During the 1990s, the United States became deeply involved in international testing programs (Stiggins, 1999; Stiggins, 2002). The media perpetuated the need by reporting math and science results in comparison with other countries’ scores. Concerns about the results led to worries that educators “had better raise our standing among the nations of the world or risk social and economic decay” (Stiggins, 1999, p. 1). These concerns were addressed at President George H.W. Bush’s educational summit, *America 2000*, attended by all of the nation’s governors. The program led to the passing of the *GOALS 2000: Educate America Act* in 1994 under President Bill Clinton (*GOALS 2000*). A federal presence was now established in educational assessment in the United States.

The 2000s could be called the “learning through standards and accountability” era in American education (Sloane & Kelly, 2003). This idea was formalized when President George W. Bush signed into law the No Child Left Behind (NCLB) Act (2002). This measure mandated standardized testing of every student in the United States in math and reading (Karp, 2004). The significance of standardized tests led many to believe that the United States was “once again revealing our faith in assessment of learning as a school improvement tool” (Chappuis, et al., 2004, p. 17).

In short, it is evident that some form of assessment has always existed in public education. Teachers in their classrooms have always used tests to measure achievement, ranging from oral examinations to portfolio assessments, as part of the process of student assessment. At the same time, standardized examinations have been used since the early 1800s to keep the public informed about the general quality of schools and schooling. Throughout the history of United States education, the majority of the focus has been on the latter. Stiggins (2002) urged, “If we wish to maximize student achievement in the U.S., we must pay far greater attention to the improvement of classroom assessment.” (p. 758). Also, the influence of educational theorists from the early 1900s, like Thorndike, has had lasting and deeply engrained effects in the American educational system.

Contemporary Conceptions about Assessment

This section of the literature review is organized around a holistic notion of assessments. First, general assessments are discussed, then in an attempt to categorize, formative, summative, and official assessments are defined. The next section will include major studies and literature reviews on the impact of classroom assessments and student achievement. Finally, teachers’ assessment literacy will be defined and explored.

Assessment in Education

Using Airasian's (2000) descriptions, there are three different types of assessments: formative, summative and official. Formative assessments are used to change or improve ongoing classroom processes while learning is still in progress (p. 94). Summative assessments are used to evaluate the outcomes of instruction and take the form of tests, projects, term papers, and final exams. Formative and summative assessments take place in the classroom; therefore, they are also called classroom assessments. Official assessments are formal and systematic tests that are required by the school bureaucracy for purposes such as pupil testing, grading, and placement (p. 95).

Lambert and Lines (2000) defined assessment as "the process of gathering, interpreting, recording, and using information about pupils' responses to educational tasks" (p. 4). They suggested that the four purposes of assessment are: a) to provide feedback to teachers and students about progress to support future learning, b) to provide information about the level of pupils' achievement at points during and at the end of school, c) to provide the means for selecting by qualification, and d) to contribute to the information on which judgments are made concerning the effectiveness or quality of individuals and institutions in the system as a whole.

The purposes for assessing vary considerably across many groups of people within the educational community. Nagy (2000) proclaimed the three roles of assessment as gatekeeping, accountability, and instructional diagnosis. For example, policymakers use assessments to monitor the quality of education and to formulate policies. Administrators and principals identify program strengths and weaknesses to plan and improve programs. Teachers use assessments to perform individual diagnosis, monitor

student progress, carry out curriculum evaluation, and determine grades. Finally, parents and students use assessments to assess student strengths and weaknesses, determine school accountability, and make informed educational and career decisions (NCREL, 1991).

Assessing for different purposes and for different groups of people can result in intense stress for all involved. Barksdale-Ladd and Thomas (2000) interviewed 59 teachers and 20 parents in two large states that have standards, benchmarks, and standardized tests to assess students on the standards. They found that teachers and parents were unanimous about the intense stress, the undermining of meaningful instruction and learning, and the high stakes involved. Interestingly, teachers did not want to see a total dismissal of assessment practices. They encouraged best practices in assessment, including: a) providing feedback to help students improve their learning; b) making assessment a part of a student's work, which can go into a working portfolio; c) providing flexibility without dominating curriculum; d) using data to inform instruction to help teachers improve instruction; and e) using more than one type of measurement for assessing students' learning (Barksdale-Ladd & Thomas, p. 395). Additionally, Shellard (2005) suggested that teachers use frequent assessments of student performance. The data produced from these assessments can be used to determine how well students are doing and identify areas where intervention or changes in instruction are needed.

To summarize, assessments are a judgment of a pupil's achievements. There are different purposes and types of assessments, which are used for varying reasons to inform different groups of people. Although assessments can cause stress to students, teachers,

parents, and all others involved, teachers conveyed the benefits of assessment when administered using best practices.

Classroom Assessment and Student Achievement

While many studies have focused on large-scale standardized testing programs, classroom assessments, both formative and summative, have received increased attention (Popham, 2002; Airasian, 2000, 1991; Black & Wiliam, 1998a; Stiggins, 2001, 1997; Stiggins & Conklin, 1992; Crooks, 1988). Students spend vastly greater amounts of time engaged in classroom assessment activities than in standardized testing; therefore, a greater impact is made through classroom assessment (Crooks, 1988; Stiggins & Bridgeford, 1985). In fact, teachers spend at least one third of their professional time on assessment activities that inform a wide variety of decisions made daily and those decisions directly influence students' learning experiences (Stiggins & Conklin, 1992). Reviews of the research by Natriello (1987), Crooks (1988) and, more recently, Black and Wiliam (1998b) have demonstrated that substantial learning gains are possible when teachers introduce classroom assessment strategies, particularly formative assessment strategies.

Natriello (1987) provided a conceptual framework in which the steps of classroom assessment process were broken into eight steps: a) establishing purpose for evaluating students; b) assigning tasks to students; c) setting criteria for student performance; d) setting standards for student performance; e) sampling information on student performance; f) appraising student performance; g) providing feedback to student performers; and h) monitoring outcomes of the evaluation of students (p. 156). Each of the stages of the model suggested features that teachers must attend to because they may

have an impact on students. Perhaps his most significant point was that the key purposes for assessment are conflated due to the multiple uses by various stakeholders. Until educators can come to a single purpose for assessing students, then the research will continue to be irrelevant (Natriello).

Crooks' (1988) literature review from 14 specific fields of research focused on relationships between classroom assessment practices and student outcomes. The primary conclusion was that classroom assessment has powerful direct and indirect impacts. For instance, a classroom assessment guides students' judgment about what is important to learn, while it also affects students' motivation and self-perceptions of competence. Crooks also concluded that grades, considered a summative function of assessment, have been too dominant and that more emphasis should be given to using classroom assessments formatively to assist in learning. The feedback given to students should focus on the task, should be given regularly and while still relevant, and should be specific to the learning task.

Black and Wiliam (1998b) used Natriello's and Crooks reviews of the literature as a baseline for their seminal review in an effort to build on previous work, as well as include more current literature to produce a review of 250 publications. In their synthesis of studies on classroom assessment, typical effect sizes of formative classroom assessment experiments were between 0.4 and 0.7 with larger gains made in low-achieving students. They stressed that no reform or policy aimed at increasing student achievement through official testing will be successful because "learning is driven by what teachers and pupils do in classrooms" (Black & Wiliam, 1998a, p. 140). Suggestions for classroom assessment practices included enhancing feedback, actively

involving students in their learning, adjusting instruction and re-teaching, and engaging students in self and peer-assessment activities. Another key point of the review revealed what they called a “poverty of practice” (p. 141) on the part of teachers. The primary difficulties teachers had with classroom assessments centered on three issues – effective learning, the negative impact of assessments, and the managerial role of assessments. Effective learning from assessments diminished because teachers encouraged rote and superficial learning and delivery practices, such as questioning and classroom discussions have not been critically reviewed in relation to classroom assessments. Assessments typically left a negative impact due to the approaches used by teachers. They reiterated Crooks (1988) thoughts on grading; stating that it was overemphasized, while useful feedback to students was underemphasized. In addition, teachers tended to create a competitive atmosphere through comparing students with one another causing low-achieving students to believe that they cannot learn. Teachers used assessments to fulfill a managerial role and to fill a grade book, rather than to diagnose student learning. Finally, classroom assessments typically looked like official standardized tests and the results are not addressed once they were recorded.

The reviews of literature were beneficial in synthesizing the characteristics of classroom assessments. However, numerous studies, both qualitative and quantitative, have indicated that classroom assessment practices have a tremendous impact on student achievement.

In an earlier study, Benjamin Bloom and his students (1984) set out to determine what teaching style would net student achievement gains as highly as one-on-one tutoring. The control class consisted of 30 students with one teacher who taught the

content and tested periodically with percentage scores given per test. The first experimental group consisted of the same classroom demographics and testing procedures, however, after tests were administered, feedback was provided followed by corrective procedures and parallel formative tests to determine the extent to which the students had mastered the subject matter. A final experimental group consisted of one student per one tutor that provided instruction that was followed by periodic formative tests, feedback and corrective procedures, and parallel formative tests. In a final achievement measure, the tutored group scored 98% higher than the control group, but more importantly, the first experimental group that consisted of 30 students that received feedback and corrective procedures scored 84% higher than the control group.

Brookhart (1997) developed a framework that measures classroom assessment and student achievement, but also includes student self-efficacy and student effort. She tested this framework in two third grade classrooms in a small urban district (Brookhart & DeVoge, 1999). Significant relationships were found among perceptions of task, self-efficacy, effort, and achievement. This study demonstrated that the use of classroom assessment strategies will not only increase student achievement, but also students' self-perceptions about learning.

Rodriguez (2004) used Brookhart's (1997) framework to study the interrelationships of teacher assessment practices, student self-efficacy, student effort, and achievement performance. He used the United States portion of the Third International Math and Science Study (TIMSS) to estimate the relationships. At the classroom level, teacher assessment practices had significant relationships to classroom performance. He reported gains of over one and a half standard deviations on math

performance arising from the effective management of classroom assessment.

Comparable to Bloom's (1984) findings, Rodriguez predicted that expected achievement score gains will rival in their impact on student achievement the implementation of one-to-one tutoring instruction, with the largest gains being realized by the lowest achievers, thus reducing achievement gaps.

Similar to Rodriguez (2004), Meisel, Atkins-Burnett, Xue, and Bickel (2003) found gains made by low achievers. In their study that utilized data collected over three years, Meisel, et al. focused on curriculum-embedded classroom assessments and its impact on scores on the Iowa Tests of Basic Skills. The study took place in Pittsburgh schools in which the clientele was 71% African American, 90% received free or reduced lunch and had 9.8% mobility. The comparison group was chosen to match schools as closely as possible on race, income, mobility, school size, and number of parents in the home. Students that were in classrooms that used curriculum-embedded classroom assessments displayed significant growth in reading and mathematics. Perhaps even more important, results of above and below average students were examined separately and they were able to demonstrate growth in both groups and across the entire study group.

In another study on classroom assessment and student achievement, teacher training was emphasized. Wiliam, Lee, Harrison, and Black (2004) wanted to determine if increased classroom assessment strategies improved student achievement on official assessments, in this case, England's school-leaving examination (the General Certificate of Secondary Education, or GCSE). Rather than simply direct teachers to use classroom assessment strategies, the researchers established training for 24 teachers over a six month span of time in exploring and planning their approach to classroom assessments.

Then the teachers put the plans into action with selected classes. Results indicated that improvements equivalent to approximately one-half of a GCSE grade per student per subject were achieved.

Researchers have presented clear evidence that there is a positive relationship between teachers using classroom assessment strategies and the impact on student achievement through various measures. Knowledge and proficiency are important among classroom teachers. This concept is generally referred to as assessment literacy.

Teachers' Assessment Literacy

An increase in recent literature has called for teachers to become assessment literate. Popham (2004) referred to a lack of assessment literacy as “professional suicide” (p.82) and proclaimed that teachers are obligated to invest time toward gaining knowledge in this area. Assessment literacy refers to a teacher’s knowledge about the basic principles of sound assessment practice, including terminology, the development and use of assessment methodologies and techniques, and a familiarity with standards of quality in assessment. Increasingly, this knowledge includes familiarity with alternatives to traditional measurements of learning (Hearne, 2001).

Assessment formats

Assessment literate teachers choose appropriate formats to assess different achievement targets and clearly understand the strengths and weaknesses of each of these formats (Stiggins & Conklin, 1992). Examples of teacher-created assessments are teacher-developed tests and quizzes, text-embedded tests and quizzes, classroom discussions, questioning, homework, and seatwork. Additionally, assessment literate teachers match items with course objectives and instruction to ensure content validity

(Airasian, 1991; Black and Wiliam, 1998b). Also, teachers provide opportunities for students to express their understanding through authentic assessments, because they know this will ensure the interaction necessary for learning to occur (Black and Wiliam, 1998a).

Stiggins and Conklin (1992) administered questionnaires to teachers about their levels of concern about quality of teacher-made assessments. They found that the quality of classroom assessments varied with grade levels and slightly with subject areas. There was an increased concern among teachers about the improvement of teacher-made objective tests at higher-grade levels. Also, math and science teachers were more concerned about the quality of the tests they produced than were writing teachers.

Questions

Another suggestion for teacher-made assessments is to tap into higher-order thinking skills only after teachers have presented a solid base knowledge of what they are teaching. Stiggins (1997) originally encouraged teachers to use higher order questions in classroom assessment. For example, in a study on measuring thinking skills in classroom assessment, Stiggins, Griswold, and Wikelund (1989) analyzed writing assessments written by a group of 36 teachers from grades two through 12. Utilizing Quellmalz Taxonomy (Stiggins, Rubel & Quellmalz, 1986) – recall, analysis, comparison, inference and evaluation - the researchers found that across the grades questions of recall dominated in classroom assessments, while comparison and evaluation questions were rare. In a more recent writing, Stiggins (2001) declared that the foundation of academic competence rests on knowledge and understanding. Teachers cannot write higher-level questions without establishing a foundational knowledge base.

Classroom assessment design

Teachers create a majority of their assessments, or they adapt assessments from a text-based format. It is imperative that they have the knowledge to design, develop, use and value methods for assessing students. Teachers must also know how to match teaching objectives with assessment items, as well as use the appropriate assessment format (Black & Wiliam, 1998b). When assessing, teachers must be sure to establish a solid knowledge base but it is also important to push students into higher-order thinking. An easy way to accomplish higher levels of thinking is through performance assessments (Black & Wiliam, 1998a).

Perceptions about classroom assessments

While there are many of resources that identify strategies for improving assessment literacy, ample research has illustrated weaknesses among practicing teachers in this area. Perhaps the most telling in the lack of assessment literacy in practicing teachers were studies that focused on teachers' perceptions of classroom assessment. Brown (2004) administered a 50-item survey to 525 New Zealand primary school teachers and principals. The four factors on the survey were improvement of teaching and learning, school accountability, student accountability and treating assessment as irrelevant. The participants agreed with the improvement conceptions and the school accountability conception, while they rejected the view that assessment was irrelevant. However, the respondents disagreed that assessment was for student accountability. Interestingly, no statistically significant differences were found in mean scale scores for each conception regardless of teacher demographics, like age, gender, or role, or for school demographics, like size, location, or socio-economic variables.

Another study of teachers' perceptions about assessment revealed teachers' underlying beliefs about assessment. The findings may help explain the struggles in changing assessment practices. Hargreaves (2005) surveyed 83 teachers asking them for their definitions of "assessment" and "learning". Through coding, the two categories that emerged from the definitions of "assessment" were assessment-as-measurement, meaning assessment used as a final judgment of students' work, and assessment-as-inquiry, meaning assessment was used to help students discover what they know about content. The first category, assessment-as-measure, was the prevalent definition by teachers in the study. Similarly, two related conceptions about "learning" were learning-as-obtaining-objectives and learning-as-construction-of-knowledge. The first conception, learning-as-obtaining-objectives, was predominant, once again demonstrating teachers' beliefs that assessments should be used as an end-product only.

Professional Development

In addition to teachers' beliefs about assessment, a lack of professional development and training may explain teachers' lack of assessment literacy. For instance, Plake, et al. (1993) developed a survey based on the Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, NEA, 1990). The standards addressed seven broad areas in classroom assessment:

- Choosing assessment methods appropriate for instructional decisions
- Developing assessment methods appropriate for instructional decisions
- Administering, scoring, and interpreting the results of both externally produced and teacher-produced assessment methods

- Using assessment results when making decisions about individual students, planning teaching, developing curriculum, and improving schools
- Developing valid pupil grading procedures
- Communicating assessment results to students, parents, other lay audiences, and other educators
- Recognizing unethical, illegal, and other inappropriate methods and uses of assessment information.

Plake, et al. (1993) administered the survey to a national sample of teachers and administrators, representing large and small school districts in rural, suburban, and urban areas. Returns were received from 42 out of the 50 United States. Overall, the highest performance area was on the subscale measuring teacher knowledge in the areas of administering, scoring, and interpreting test results. The poorest performance was on items measuring the teachers' knowledge about communicating test results.

Building on Plake's (1993) work, Metler (2004) developed the Classroom Assessment Literacy Inventory. The survey was administered to 67 undergraduate secondary education students and 101 in-service teachers in an attempt to measure teachers' assessment literacy. In-service teachers scored highest on administering, scoring, and interpreting results of assessment and lowest on developing valid grading procedures. Pre-service teachers scored higher on choosing appropriate assessment methods and lowest, like the in-service teachers, on developing valid grading procedures.

Numerous studies focused on the severe lack of training that teachers received, both as pre-service teachers and as practicing in-service teachers (Black & Wiliam, 1998b; Cizek & Fitzgerald, 1996; Schafer, 1993). In addition, studies revealed that

teachers who had taught for several years still lacked in assessment practices, thus teachers do not tend to acquire skills in this area through “on the job” training (Cizek & Fitzgerald, p. 170). Nevertheless, Cizek and Fitzgerald also found that teachers believed that it was important to do whatever would help students succeed. McMillan (2003) agreed with Cizek and Fitzgerald, but demonstrated that external factors - like accountability testing and district policies, and classroom realities - like absenteeism and disruptive behavior, created tensions that influenced teachers’ decision-making in their assessment practices. The author’s illustration displays this notion in Figure 3.

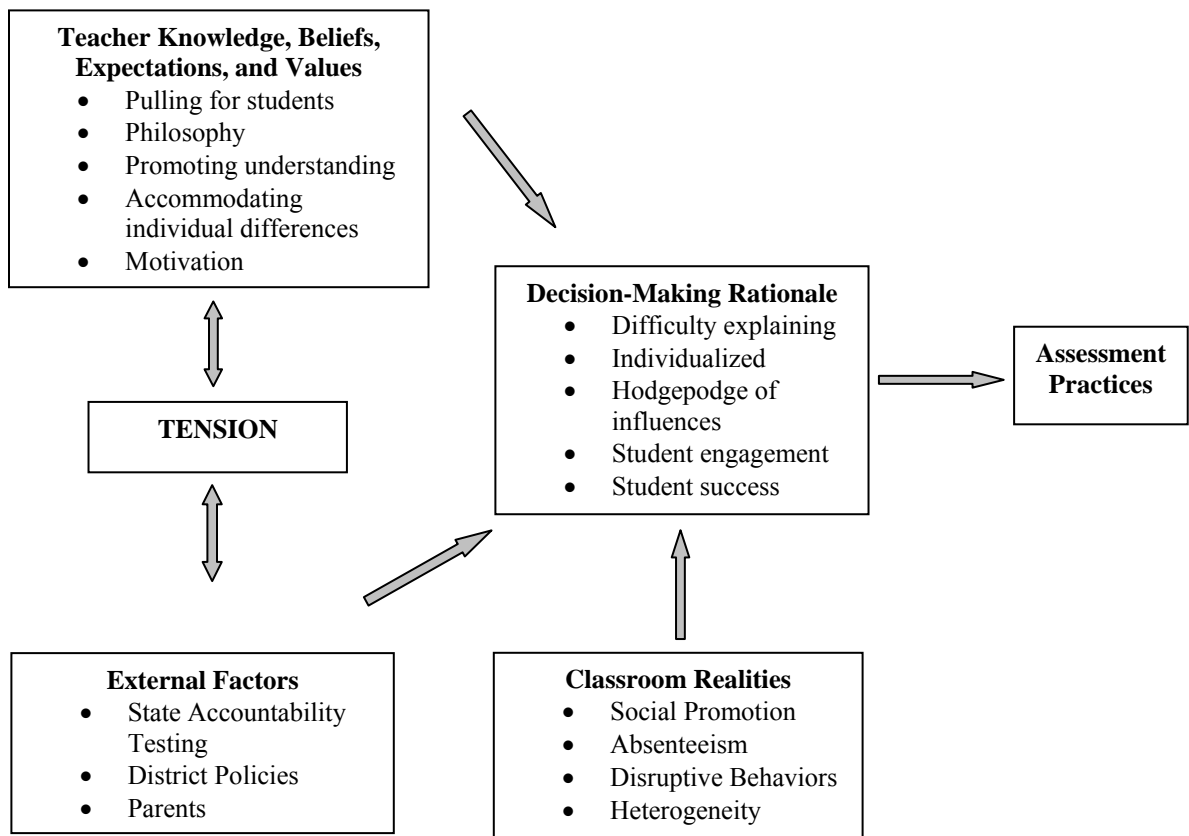


FIGURE 3. *Teachers’ classroom assessment decision making.* (McMillan, 2003, p. 36)

Recent studies have described training for teachers with indications of success. For instance, Lukin et al. (2004) launched reform efforts in teacher training in assessment through the University of Nebraska at Lincoln and the Lincoln Public School District. Two programs were designed to provide training to experienced teachers and two programs were developed for pre-service teachers. Evidence suggested that all of the training models had positive impacts on teacher confidence, knowledge, and skills in key areas of assessment.

Teachers' assessment literacy is a vital key in classroom assessment practices. Teachers with strong assessment literacy possess a working knowledge of when and how to design, develop, use and value a wide variety of methods for assessing student achievement (Stiggins & Conklin, 1992). A majority of the research illustrated that teachers lacked assessment literacy skills, however, there is growth in teacher training and professional development in this area.

Teacher Strategies for Classroom Assessments

Based on the extensive literature reviews (Crooks, 1988; Natriello, 1987; Black & Wiliam, 1998b) and the studies on classroom assessment and student achievement (Rogriguez, 2004; Wiliam, et al., 2004; Meisel, et al., 2003; Bloom, 1984), numerous strategies are listed as non-negotiable in classroom assessment. In this review, the strategies are broadly addressed as (1) assessment strategies used during instruction, (2) teacher adaptation of instruction per assessment data, (3) teacher feedback per assessment data, (4) teacher development of peer and self-assessment strategies, and (5) the formative use of summative assessments.

Assessment strategies used during instruction

Teachers should think of assessment as a process rather than an end product. Assessments occur throughout lessons. Talk and questioning are productive forms of assessing during instruction (Black, 1998). Talk has potential to transform learning (Gilles, 1995). Through talk, students can give the reasoning behind strategies, tactics, and interpretations that may not be clear from the final product. The Appalachia Educational Laboratory (AEL) (2005) developed a framework that presents questioning as a five-stage process: a) question preparation; b) presentation of questions; c) prompting; d) processing of student responses; and e) reflection on questioning practice. A study was conducted testing the framework in 28 fifth- and sixth-grade classrooms. Teachers were trained in questioning strategies and then videotaped leading a class discussion. Questions were coded based on the framework criteria. Growth in teachers' questioning strategies was shown in all areas.

Wolf, et al. (2006) examined the relationship between the quality of classroom talk and academic rigor in reading comprehension lessons. The data from the study included 21 reading comprehension lessons in several elementary and middle schools from three urban school districts. Quantitative analyses revealed that students' ratings on talk providing knowledge and thinking had strong, positive relationships with academic rigor. Similarly, Myers' (2005) action research study in a kindergarten classroom showed that questioning strategies improved comprehension in very young nonreaders. In a three-month study using interviews, rubrics and anecdotal records, findings revealed that even younger students were able to reflect on their learning and to self-monitor their comprehension through classroom discussions.

While questioning and talk are good methods for developing comprehension during classroom assessments, it is vital that teachers record results during the assessment (Stiggins & Bridgeford, 1985). Examples for recording can be checklists, performance rating scales, or anecdotal records (Stiggins, 1997).

“One of the challenges in teaching is designing, and to be a good designer you have to think about what you're trying to accomplish and craft a combination of the content and the instructional methods, but also the assessment” (Wiggins, 2002).

Performance assessments are an effective form of classroom assessment that requires good design. A performance assessment can take the form of many activities, but it is defined as the observation and rating of student behavior and products in contexts where students actually demonstrate proficiency (Stiggins & Bridgeford, 1985). Typically, students show the teacher what they know through an activity of some sort. According to Black (1998), one unifying idea among different types of performance assessments is that the activities are “direct models of reality to be assessed rather than disconnected fragments or surrogates” (p. 87). It is generally agreed that performance assessments must be designed well in advance with observable and clearly defined performance tasks (Black; Stiggins, 1997; Airasian, 1991).

Shepard, et al. (1996) wanted to study if performance assessments directly impacted student achievement. In a yearlong project, thirteen third-grade classrooms used performance assessments as a part of instruction in reading and math. Control classrooms were used for comparison. Interestingly, the achievement scores in the focus classes were not higher than the control classes, however gains were found in lower-achieving

students. Sheperd, et al. pointed out that teacher training and professional development was necessary and the lack of training influenced the study.

Detailed scoring protocols must be provided before the assessment. In a detailed review of literature, Soles (2001) presented the benefits of utilizing scoring guides throughout lessons. By sharing a scoring guide at the beginning of lessons, anxiety about success is reduced and a relationship is established between teaching and assessing. Clear, shared scoring guides reduced confusion about expectations of success. Finally, scoring guides empower students because they encouraged students to become active participants in the learning.

Assessing throughout lessons helps the teacher get to know students on an individual level. Talk and questioning provide a means for obtain assessment data. Performance assessments are a powerful type of assessment in which data are collected throughout the task. Lastly, scoring guides provide a powerful means for students to become a part of the assessment process.

Teacher Adaptation of Instruction per Assessment Data

As Black and Wiliam (1998) attest, the key part of assessment is the formative part, in which assessments are used to change or improve ongoing classroom processes while learning is still in progress (Airasian, 2000, p. 94). Therefore, teachers must use classroom assessment results to inform their teaching. If students do not exhibit success, re-teaching and possibly adjusting instruction are necessary to complete the classroom assessment cycle.

A study by Bergan, et al. (1991) on early acquisition of basic skills illustrated that re-teaching is a necessary step in cognitive development. The study involved 838

students drawn mainly from disadvantaged home backgrounds in six different regions of the United States. The teachers in the experimental group were trained to implement a measurement and planning system which required an initial assessment input to inform their teaching at the individual level. Progress was checked after two weeks, and a new assessment was given and new decisions about students' needs were made based on the results of the assessment. The cycle continued for eight weeks. The teachers mainly used observations of skills to assess progress, and worked with open-style activities which enabled them to differentiate tasks within each activity in order to match to the needs of the individual child. The experimental group achieved significantly higher scores in reading, mathematics and science than the control group. It is also important to note that the final assessments were traditional multiple choice tests and were not adapted to match the formative teaching and learning styles of the classroom.

Teacher Feedback per Assessment Data

One of the most commonly cited necessities of classroom assessment is communicating assessment results, or feedback (Sadler, 1989; Black & Wiliam, 1998a; Cronbach, 1977; Sloane & Kelly, 2003; Bloom, B. S., Hastings, J.T., & Madaus, G., 1971). Feedback on tests, seatwork, and homework should give students guidance on how to improve, and each student must be given help and an opportunity to work on improvement. Feedback should be nonjudgmental (Brookhart, 1997) and it should focus on particular qualities of students' work, with advice on what he or she can do to improve, and should avoid comparisons with other students (Black & Wiliam, 1998a). It should be frequent and descriptive, providing students with information about their

strengths as well as areas for improvement (Chappuis, Stiggins, Arter, and Chappuis, 2003).

Grading as feedback.

A tremendous amount of research has focused on the administration of grades as a form of feedback. Since grade-based decisions may have lasting academic and social consequences (Messick, 1989; Popham, 2001), teachers should weigh assessment elements according to instructional emphasis (Airasian, 1991; Stiggins, Frisbie & Griswold, 1989). For example, if the main point of a writing lesson was proper nouns, then scoring should focus on proper nouns. In addition, teachers need to communicate grade criteria in advance (Stiggins, Frisbie, & Griswold).

Grades should be based on achievement factors only (Cizek, & Fitzgerald, 1996; Stiggins, Frisbie & Griswold, 1989). Brookhart (1997) found that teachers make value judgments when assigning grades. Eighty-four teachers responded to a multiple-choice questionnaire about teachers' choices when assigning grades. The multiple-choice questions were followed by an open-ended question asking why they made that particular choice. Teachers indicated that they were concerned with the consequences of grade use, especially for developing student self-esteem and good attitudes toward future schoolwork. In another study on teachers' grading practices, McMillan (2001) utilized an instrument designed to measure factors teachers consider when assigning grades. Academic achievement was the most prevalent factor, but use of extra credit, academic enablers (effort, ability, improvement, and participation) and homework and zeroes were also considered. Non-achievement factors such as effort and ability should not be incorporated into subject-matter grades (Stiggins, Frisbie & Griswold, 1989).

A negative relationship was found between feedback provided only through grades and student achievement in a study of 48 Israeli students selected from 12 classes across four schools (Butler, 1988). The students were given one of three types of feedback on assigned tasks. One-third of the students were given individually composed comments on how well their work matched evaluation criteria. A second group was given only grades, derived from the scores on the preceding session's work. The third group was given both grades and comments. In a measure of achievement, the group that received only comments increased about 33% by the second session and remained high throughout the rest of the sessions. The group that received grades and comments showed significant decline across the sessions, while the group that only received grades scored low and remained that way throughout the sessions. Tests of students' interest about the task reflected the achievement scores with the group that received comments only reflecting high interest and the group that received grades only indicated low interest on the task.

Grading is an important part of classroom assessment because it is traditionally the way that teachers communicate results to students and parents (Popham, 2002). It is important to remember that grading is done appropriately by making sure that scores match the emphasis in a lesson (Airasian, 1991). It is also essential that grading criteria are presented in advance and value judgments are not considered into a score (Brookhart, 1997). Grades do not motivate all students, particularly those who feel that high grades are out of their reach (Sadler, 1989). Also, grades tend to be holistic judgments rather than descriptions of strengths and weaknesses (Cronbach, 1977; Sloane & Kelly, 2003). Through written feedback, teachers can help students "internalize quality criteria by

translating them from latent to manifest and back to latent again until these criteria become so obviously taken for granted that they need no longer be stated explicitly” (Shepard, 2005, p. 68).

Feedback loop.

An important feature about feedback is that it should be used to close a gap, referring to the level in which students are performing and the desired level of achievement. If information is simply “recorded, passed to a third party who lacks either the knowledge or the power to change the outcome, or is too deeply coded (for example, a summary grade given by a teacher) to lead to appropriate action, the control loop cannot be closed, and ‘dangling data’ is substituted for effective feedback” (Sadler, 1989, p. 125).

Feedback is defined as having three elements: redefinition of the desired goal, evidence about the student’s present position, and some understanding of a way to close the gap between the two. All must be understood to some degree before the student can take action to improve learning. (Black & Wiliam, 1998a; Sadler, 1998). According to Black and Wiliam (1998b), for an assessment to be formative the feedback provided has to be used by the student to complete the feedback loop. Students need to be trained in how to interpret feedback and how to use the feedback to improve their work (Sadler).

Using Bandura’s social cognition theory, Ross and his colleagues (2002) studied how students in grades two, four, and six processed feedback. Data were collected from 71 students in Toronto, Canada, equally divided among three grades and two genders. The students reported that assessment feedback told them what they were good at, what they needed to improve on, whether they needed to work harder and clarified teachers’

expectations. Virtually all of the students reported that they used evaluation data to plan future actions; however they found that older students processed feedback more thoughtfully than the younger students. They also discovered that parents' opinions about assessment feedback influenced student's thoughts on assessment and their performance.

Teachers need to be trained in writing feedback with a focus on keeping the feedback descriptive rather than evaluative (Chappius, et al., 2004). A study by Elawar and Corno (1985), focused on training teachers to focus their written feedback on deep learning rather than superficial. In this study, 18 mathematics teachers received training before working with over 500 students in three schools. A control group followed the normal practice of grading homework without comments. All students were given a pre-test and post-test. Analysis of variance of the results showed a large effect associated with the feedback treatment, which accounted for 24% of the variance in the final achievement score.

Feedback should be more than giving a letter grade. It should be given frequently in a nonjudgmental manner. The purpose of feedback is to close the gap between where a student should be and where they actually are performing. By providing feedback, students are invited to think about their learning along with teachers. This leads to student involved classroom assessments.

Teacher Development of Peer and Self-Assessments

Involving students is an essential component of classroom assessment (Black & Wiliam, 1998a). When students become involved there is a shift from passive receptivity to active involvement in learning (Shepard, 2005; Ellis, 2001). Students are also provided the opportunity to develop skills that are essential in their professional lives, such as the

ability to function as independent learners, to exercise judgment and to transfer learning from one context to another (Ellis).

Ways to involve students are student-involved record keeping, student-involved communication, and student and peer assessment (Stiggins, 2001). Some specific record keeping and communication activities are students creating visual displays of important performance criteria for bulletin boards, involving students in the process of transforming performance criteria into checklists, rating scales, and other recording methods, and having students reflect and write about their own growth over time with respect to specific criteria (Stiggins, 1997). Other examples are students helping in the development of assessment exercises, creating scoring guides, and applying the scoring guide to their own work (Stiggins, 2001).

Fortana and Ferrandes (1994) presented an example of the use of self-assessment methods and its positive impact on student achievement. They studied 25 Portuguese math teachers whom they had trained in self-assessment methods during a 20-week part-time course. They put the methods into practice with 246 students ages eight and nine and with 108 older students with ages between ten and 14. Twenty other math teachers, who were taking another course in math education, were used as the control group. Through pre- and post-testing, both groups of students made gains in math achievement, but the experimental group's mean was about twice that of the control group.

While involving students makes assessment more democratic and provides students opportunities to debate and exercise collaborative decision-making (Ellis, 2001), it has been shown to generate stress on students, particularly females. Pope (2005) looked at the effect of stress on students involved in self - and peer-assessment. Results indicated

that while self- and peer-assessment increased stress, it also led to improved student performance on summative tasks.

On the other hand, Schunk (1996) asserted that providing students the opportunity to use self-assessment raised self-efficacy, skill, motivation and task orientation. Two groups of elementary students worked on the introduction of fractions. The group that used self-assessment strategies showed significant effects on the analysis of covariance (ANCOVA) in self-efficacy due to self-evaluation (13.85) at an alpha of .01 and also the interaction between goal attainment and self-evaluation (7.10) at an alpha level of .05.

True self- and peer-assessment is much deeper than asking students to trade papers to grade. Students take on a partnership role in their learning (Stiggins, 2001). Eventually, students come to understand how assessment and self-assessment affect their own academic success (Stiggins, 2001). Higgins, et al. (1994) worked with first and second grade teachers and students in their development of assessment skills through peer evaluation. The children generated their own criteria and the quality of the criteria rose throughout the study. An example of the criteria requirements was that “it must relate to what was learned” (p. 321). This showed sophistication in the students’ knowledge of the connection between learning and assessment.

Students can assess themselves only when they have a sufficiently clear picture of the targets that their learning is meant to attain (Black and Wiliam, 1998b). A study from a group of Stanford researchers presented findings that clearly illustrated this fact. Two groups of students were asked to work on group projects, but only one group was given evaluation criteria. This groups’ mean on the final product was significantly higher than

the group without criteria, while the group without criteria spent almost 20% of their time off-task (Cohen, et al., 2002).

Students learn how to self-assess through the teacher modeling the act of providing feedback. Then students use this knowledge to self-critique and check for those elements in their work (Shepard, 2005). Teachers have to know, or receive training in how to administer a lesson that involves self- and peer-assessment (Stiggins, 2001), but they also have to realize that student-involved assessment strategies are only a part of a larger multi-faceted assessment process that may affect student learning processes and outcomes (Marshall & Weinstein, 1984). Student-involved activities, like self- and peer-assessments, can affect the motivation and self-esteem of students (Black & Wiliam, 1998a). These activities can help increase students' control of their own welfare and will remove the mystery that too often surrounds the meaning of success in the classroom (Stiggins, 1997).

The Formative Use of Summative Data

To enhance learning, researchers have proposed using an assessment system that integrates summative and formative assessments that will improve both the quality of learning and the quality of assessments (Stiggins, 2002; Wiliam, 2000; Shepard, 2005). Wiliam claimed that improving the quality of teachers' day-to-day classroom assessment practices has a substantial enough effect to take an average country, like the United States or England, up into the top five countries on the international rankings on student achievement (p. 112).

When thinking of integration between formative and summative assessments, it is important to remember the meanings of the terms. Bloom, et al. (1971) were the first to

define the terms. They defined summative evaluations as assessments given at the end of units, mid-term, and at the end of a course, which are designed to judge the extent of students' learning in a course (p. 117). They contrasted these with formative evaluations, which elicit evidence that yields interpretations that form the basis for successful action in improving performance (Black & Wiliam, 1996). To integrate formative and summative assessment practices, teachers take summative assessment data and use it to make instructional decisions to improve student performance (Black & Wiliam, 1998a).

Black and Wiliam (1996) stressed that the terms formative and summative not be applied to the assessments, but to the functions that they served. Shepard (2005) stressed that “knowledge is constructed, and that learning and development are culturally embedded, socially supported processes” (p. 66). This learning theory provides the big-picture understandings needed to change teaching and assessment practices.

An example of a school that worked to integrate formative and summative assessment was described by Many and Jakicic (2006). In the qualitative study, teachers focused on assessments, and they realized that they were inconsistent in their classroom assessment practices with a wide range between highly summative and loosely formative assessments. They worked to develop a continuum of assessments throughout the middle school years constant to all of the teachers. They created a balanced assessment system that looked at four categories of assessments: classroom assessments, common assessments, district-level assessments, and official assessments. Each of the assessments were analyzed and discussed. Teachers discovered that the conversations about the assessments affected their teaching and what they felt was important. Additionally, nearly 90% of the school's students improved on official assessments.

In another study, Henning (2006) studied 24 elementary and middle school teachers as they analyzed standardized achievement test scores in a Teacher Leadership Program at the University of Northern Iowa. Findings indicated that teachers were adept at analyzing and disaggregating data after they were provided with training. Findings also showed that the teachers found numerous ways to analyze the data creating a wide variety on information.

Integration of summative and formative assessments is perhaps the key factor to the successful use of classroom assessments. Teachers take summative data but rather than use it to make a final judgment about student performance it is used formatively to provide students with the information they need to improve achievement. The extent to which instruction is guided by assessment data depends on the leadership of the principal. Specifically, it depends on the principal's ability to model tools and strategies for using assessment data to improve instruction; help the staff acquire the requisite skills to use assessment data for instructional decision-making; and establish a school-wide norm that instruction will change based on assessment data (Fox, 2003).

Instructional Leadership

Since the beginnings of the principalship in American education, a distinctive definition for the position never emerged. Due to the technical and bureaucratic nature of running a building, principals were originally thought of in a managerial role. However, in every decade some attention has been paid to the principal's impact on teaching and learning processes (Beck & Murphy, 1993).

Traditionally, principals were expected to be the manager of the building. During the eighties, there was a shift in attention to educational roles with both teachers and

principals receiving emphasis (Murphy, 1985). A prevailing assumption of this era was that the principal should become directly involved with the teaching and learning processes of their building (Beck & Murphy, 1993). Principals were to “intervene” to ensure that teachers focused on the central mission of the school (Beck & Murphy). More than ever, principals were expected to become directly involved with instruction rather than the more traditional role as managerial leader.

Research in the nineties “pushed the principalship in new directions, focusing on the larger external forces that were shaping the future of schools” (Beck & Murphy, 1993). The focus expanded from instruction to principal as servant, organizational architect, social architect, educator, moral agent, and a person in the community (Beck and Murphy).

Later, attention on instructional leadership shifted from a focus from instructional leadership as direct involvement in classrooms to a multi-dimensional role in which the administrator fosters an environment promoting the improvement of instruction through non-managerial approaches (Macmillan, Meyer & Sherman, 2001). Administrators need to know about instructional leadership to: a) create coherence in improvement efforts; b) finely balance mandate and empowerment; 3) model learning – openness to new ideas, willingness to be driven by results, and persistence in the face of difficulty (Lashway, 2002).

Leithwood and Duke (1998) described instructional leadership as focusing on the behaviors of teachers as they engage in activities directly affecting the growth of students. At the heart of this definition is the conception that instructional leaders directly impact student achievement (Hallinger & Heck, 2000; Smith & Andrews, 1989;

Andrews, Basom, & Basom, 1991). Sheppard (1996) referred to two views of instructional leadership, “narrow” and “broad.” The narrow view of principal instructional leadership identifies it as a separate component of the principal’s responsibility, excluding behaviors that focused on school climate and teacher development (Sheppard). The broad perspective involves all activities that affect student learning (Murphy, 1988). Taking a broad view of instructional leadership, the common themes that emerged from the literature focused on instructional leader attributes and job functions, instructional leader knowledge and teacher development.

Instructional Leader Attributes and Job Functions

While Sergiovanni (1991) suggested that research on instructional leadership attributes and functions needed to be “situationally specific” (p. 90), many researchers determined characteristics that suggest strong instructional leadership. Smith and Andrews (1989) compiled a list of characteristics from existing literature that suggest strong instructional leadership. This list includes “high energy, assertiveness, ability to assume the initiative, openness to new ideas, tolerance for ambiguity, a sense of humor, analytic ability, and a practical stance toward life” (p. 8). In addition, referent power, strong motivation, and high self-esteem were influential attributes (p. 39).

In their review of related literature, Andrews, Basom and Basom (2001) found three inherent attributes common in strong instructional leaders: vision, the ability to communicate that vision, and the ability to create trust in the workplace. Building on the notion of trust, Beatty and Brew (2004) studied principals’ work in establishing trust in their schools. Using an emotional epistemologies theoretical framework, the researchers interviewed and analyzed written responses given by 42 principals in an educational

leadership program. They found that effective principals who were comfortable creating a high amount of trust established a “Vygotskian zone of proximal development” (p. 351) with their teachers when introducing instructional reform. By establishing trust, instructional leadership was more effective (Beatty and Brew).

Some discussions on the functions of an instructional leader focused on the nature of the job, stating that instructional leaders set “high academic expectations, review lesson plans, supervise classroom instruction, and monitor curriculum” (Lashway, 2003, p.4). Smith and Andrews (1989) explained instructional leadership as a combination of several tasks, such as supervision of classroom instruction, staff development, and curriculum development.

Common descriptions of the functions of instructional leaders included the necessity to provide constructive support and obtain the resources and materials necessary while keeping abreast of the latest developments in teaching, learning, motivation, classroom management, and assessment (Hoy and Hoy, 2003). For example, Smith and Andrews (1989) found in their review of the literature that principals who were instructional leaders demonstrated skills as a resource provider, instructional resource, communicator, and were a visible presence in the school.

Similarly, Hallinger and Murphy (1985) studied the instructional leadership behaviors of ten elementary principals. They administered a questionnaire, The Principal Instructional Management Rating Scale, to teachers, principals and the principals’ supervisors. The measure was divided into three main functions of instructional leadership. The three functions were defines the school mission, manages instructional programs, and promotes school culture. They also analyzed principals’ documents such

as teacher evaluation reports, newsletters, and school bulletins. They found the principals in this study frequently engaged in instructional leadership behaviors. They also noted that the high instructional leadership behaviors might have been influenced by the introduction of assistant elementary principals to the district. They speculated that by adding assistant principals, principals had more time for instructional leadership activities.

Specific leadership behaviors and functions were the focus of the Marzano, Waters, and McNulty (2005) meta-analysis of 69 other studies conducted that focused on specific leadership behaviors and their impact on student academic achievement. Table 1 illustrates the responsibilities that Marzano, et al. found to influence instructional leadership practices and their correlations with student academic achievement.

Table 1

Instructional Leadership responsibilities and their correlations (r) with student academic achievement.

Responsibility	The Extent to Which the Principal.....	Average <i>r</i>
Intellectual stimulation	Ensures faculty and staff are aware of the most current theories and practices and makes the discussion of these a regular aspect of the school's culture	.24
Involvement in Curriculum, Instruction, and Assessment	Is directly involved in the design and implementation of curriculum, instruction, and assessment practices	.20
Knowledge of Curriculum, Instruction, and Assessment	Is knowledgeable about current curriculum, instruction, and assessment practices	.25
Monitoring/Evaluating	Monitors the effectiveness of school practices and their impact on student learning	.27

Marzano, Waters and McNulty (2005) also indicated that a critical function of leadership is the knowledge to “do the right work” (p. 76) in their schools. Strong instructional leaders know the direction to lead their schools when implementing comprehensive school reforms that lead to increased student achievement. “There is a significant difference between the existence of programs commonly recommended for practice in middle level literature and a thorough understanding of that literature coupled with the effective implementation of the programs” (Valentine, Clark, Hackmann, & Petzko, 2004, p. 108). Instructional leaders must have a working knowledge of best practices.

Instructional Leader Knowledge

Efficient, thoughtful and purposeful daily and ongoing functions of a principal are essential for a school’s success; however a principal also has to have a deep comprehension of the theories and knowledge necessary to maintain a successful instructional environment (Fisk and Resnick, 1999). Fullan (1997) emphasized, “when it comes to learning, effective leaders need to be greedy” (p. 45).

In a national study, Valentine, et. al (2004) determined that leadership knowledge was a characteristic of highly successful middle schools:

...what was impressive about the (highly successful) schools...was the degree of knowledge and understanding that principals and teachers had about the research and literature on best programmatic practices in the middle level. They expressed strong support for middle level programs such as interdisciplinary teaming, exploratory courses, and advisory opportunities; they implemented those programs very effectively; and they

articulated why they implemented the programs and how the programs were effective for their students (p. 109).

In a case study of a problems-based instructional leader development course, Hallinger and McCary (1990) argued that research on instructional leadership needed to address the knowledge and thinking that underlies the exercise of leadership. Utilizing a computer simulation program, principal trainees were provided with \$30,000 and 2,000 hours of staff energy to address the problem of low test scores. Simply investing in the most cost effective program was not enough for success. The principal trainees found that knowledge of instructional methods was necessary for school improvement. One participant stated, “Efforts to use new instructional strategies are unlikely to bring about lasting changes in a teacher’s behaviors if peers and the principal do not understand or appreciate the significance of the new strategy” (p. 100). Hallinger and McCary concluded that good instructional leadership involved proficient planning and management. Additionally, instructional leaders must have the ability to use forethought before acting, as well as the ability to strategically respond in regards to a specific purpose.

Ruff and Shoho’s (2005) collective case study also focused on the cognitive levels of instructional leadership of three elementary principals at varying stages of their careers. They utilized Brewer’s definition of mental models as a means of communicating the tacit assumptions of instructional leadership. They defined mental models as the “specific knowledge structures that are constructed to represent a situation through the use of generic knowledge” (p. 557). The three participants were principals from highly successful urban schools in San Antonio, Texas. The first principal was a

first year principal, the second was in her sixth year, and the third was a nationally recognized principal with over 20 years of experience. Findings supported their beliefs that principals use different cognitive levels depending on their stage in their career. The first year principal's mental models were appropriate, but surface level. The sixth year principal focused on how involved she should become in the process of instructional leadership. They found the veteran principal had the ability to adjust and shift her perceptions of what was important to meet the needs of a particular problem. These findings supported their beliefs that principals use different cognitive levels depending on their stage in their career.

In a study of administrators' knowledge about assessments, or assessment literacy, Impara and Plake (1995) found administrators to be more knowledgeable than teachers. Strength areas for administrators included an understanding for selecting assessment strategies. In addition, they performed well on communicating assessment results, but were less proficient in the interpretation of standardized test results.

In a case study on New York City's District Two, Fisk and Resnick (1999) utilized cognitive apprentice theory in an attempt to understand the development of knowledge in principals as instructional leaders.

Building an effective community of principals is about both things—the craft of teaching and learning and the building of strong interpersonal relationships. I believe no effective learning can go on without very strong personal relationships. But relationships can't substitute for deep knowledge. The challenge is to build those relationships around studying

teaching and improving instruction for kids and a belief system about learning (Fisk and Resnick, 1999, p. 56).

Instructional leaders must possess a working knowledge of the curriculum and instructional strategies, but they are not alone. Part of an instructional leader's job is intellectual stimulation of teachers through teacher development. In fact, Marzano, et al. (2005) found a .24 correlation in importance of intellectual stimulation of teachers. This refers to the extent to which the school leader ensures that teachers and staff are presented the most current theories and practices regarding best instructional practices and creates an open environment where discussions of the practices are a regular aspect of the school's culture.

Teacher Development and Intellectual Stimulation

Teacher development is an important facet of instructional leadership (Blase and Blase, 1998). In successful middle schools, continual learning permeates every aspect of the building starting with the principal creating a culture of learning for teachers and teachers modeling learning for students through attendance of graduate school, readings, and small and whole group studies (Valentine, et. al, 2004). The principal's role is to cultivate teachers' intrinsic motivation through frequent classroom visits and attendance of team meetings to continually work to improve their teaching abilities while remaining focused on the school's vision (Jackson & Davis, 2000).

Darling-Hammond, Aneess, and Falk (1995) described the importance of leadership and teacher development in successful schools:

Each of the schools is collaboratively managed and in each case the principal encourages a culture of inquiry that provides a context for adult

engagement. School leaders have encouraged multiple opportunities for professional growth in various forms and shapes, inside and outside the school community, for teachers at all states of development. Each understands the ebb and the flow of change, respect swells of enthusiasm and the undertow of resistance, and tolerates the directional cross currents, eddies, and still waters that characterize their faculty's engagement in change. All the while, their commitment to their vision remains steadfast (p. 265).

In a study by Leithwood, et al. (2002), findings indicated that school leaders needed to realize the potential use of official testing policies as a means to improve teaching and learning and principals must recognize such outcomes might significantly increase the likelihood of teachers using official assessment initiatives for their own purposes, rather than dismissing them out of hand. Moreover, Blase and Blase (1998) supported the idea by declaring that everyone in a school is a learner; therefore a school is comprised of a "community of learners" (p. 127). They also emphasized that it is essential that all members act as colleagues and coaches; thus they all learn from one another. Through open discussions on curriculum, instruction, philosophy, research and all things related to learning, instructional leadership occurs (Blase and Blase).

Illustrating the need for teacher development when looking at instructional leadership, Marks and Printy (2003) proposed that school leadership needed to be analyzed through a lens of "integrated leadership" (p. 377), or the combination of transformational leadership and instructional leadership. Transformational leaders motivate followers by "raising their consciousness about the importance of organizational

goals and by inspiring them to transcend their own self-interest for the sake of the organization” (p. 375). The researchers wanted to thoroughly investigate instructional leadership in 24 schools that were deemed exceptional on a national level by the Center for Organization and Restructuring of Schools. They found that strong transformational leadership by the principal was essential in supporting teacher commitment; therefore, when teacher commitment increased, so did instructional leadership. Thus, they proposed that instructional leadership could be transformational. They also proposed that active principal and teacher collaboration around instructional matters enhanced the quality of teaching and student performance.

Conceptions about instructional leadership have changed throughout the decades. Prior conceptions about principals, particularly as managers, are important to understand to fully appreciate how important the work was from the eighties, as well as current notions on instructional leadership. The demands of standardized testing have created a need for principals to become instructional leaders (Quinn, 2002).

Teachers’ Assessment Practices and Leadership

The educational reforms initiated by No Child Left Behind have created a change in the principal’s role as an instructional leader. New forms of assessment, school reform initiatives, the growing number of mandated assessments, and a host of individuals demanding accountability have converged on our schools. Dealing with these and other assessment issues requires a type of leadership that Cizek (1995) coined as “assessment leadership.” Assessment leadership requires four characteristics that school and district-level administrators must possess: a) an intimate knowledge of what occurs in classrooms, b) a clear focus on the desirable educational outcomes targeted by the

curricular programs in place, c) familiarity with the purposes of a given assessment and audiences to which results are to be reported, and d) some understanding of fundamental assessment concepts (Cizek, p. 247).

Considering the growing number of mandated assessments and growing number of groups interested in various assessment outcomes, the principal must have a vision regarding how the various assessment parts fit together (Cizek, 1995). When establishing the assessment vision and goals in a school, it is essential that the principal have an appreciation of the integral relationships among teaching, learning and assessment and the resulting influence on student achievement (Noonan & Renihan, 2006; Girvin, 2005).

It is also essential to include the teachers in the assessment leadership of the school. The act of bringing together teachers to focus on assessment data lends itself to the research on professional learning communities (DuFour & Eaker, 1998). The common tenets of professional learning communities are: a) ensuring that students learn b) creating a culture of collaboration, c) focusing on results, and d) hard work and commitment (DuFour, Eaker, & DuFour, 2005). Classroom assessment data provides the stimulus for each of the tenets of a professional learning community (Stiggins, 2004).

Conclusion

The history of assessment practices in the United States revealed an institution steeped in beliefs that were primarily based on the notion that assessments were used as an end product designed as a measure of learning. Traditions in education that were established during the Industrial Revolution have had lasting impacts in current education.

The accountability movement has increased the purposes for assessment, thus, there has been amplified stress that comes with assessments due to the varying involved parties like policymakers, principals, teachers, students and parents. This review of literature explored the use of classroom assessments as a possible solution to decreasing the focus on official assessments and using assessments to increase student achievement.

The literature on current conceptions about classroom assessments indicated that they are an important part of the assessment process. Classroom assessments are not only “one of our indicators of educational outcomes, but classroom assessments are also part of the very instructional treatments that produce the desired outcomes” (Stiggins & Conklin, 1992). Numerous literature reviews and studies have bolstered this claim. Three points are noteworthy about the existing literature. First, teacher knowledge of classroom assessment, or assessment literacy, is essential (Chappuis, Stiggins, Arter & Chappuis, 2004). Many teachers did not receive assessment training during teacher certification (Stiggins, 2001; Black & Wiliam, 1998a; Shepard, et al. 1996), so professional development is essential (Chappuis, et. al). Second, classroom assessments involve a broad range of strategies. Teachers may be involved in some activities more than in others due to the nature of the grade levels and content areas they are required to teach. The existing literature has suggested that grade levels and subject areas may account for some variations in classroom assessment (Stiggins & Conklin, 1992).

Ideas about instructional leadership have evolved throughout the years. In today’s climate of accountability, there is a need for instructional leaders to transcend classroom management and evaluation. Today’s instructional leaders need to work to create a community of learners in which teachers see assessment as a tool that can be used to

improve overall student achievement. This focus on assessment calls for principals to develop strong “assessment leadership” (Cizek, 1995). Assessment leadership requires the principal to comprehend the “big picture” of new forms of assessment, school reform initiatives, and the growing number of mandated assessments to create a vision for the school (Cizek, p. 247). The assessment vision is the catalyst for teacher collaboration where teachers and administrators come together to concentrate on assessment data results with a focus on improving student learning (DuFour & Eaton, 1998).

CHAPTER 3

RESEARCH DESIGN AND METHODOLOGY

Introduction

In recent years, assessment programs in the American education system have been the source of tremendous conflict for all involved parties, like teachers, parents, and policy makers. Standardized tests used to fulfill the requirements of No Child Left Behind (NCLB) (2002) are necessary for policy and programmatic decision makers (Stiggins, 1999); however, the tests are used for reasons other than as a method of accountability thus, creating tension.

According to Nagy (2000), the three roles of assessment are gatekeeping, accountability, and instructional diagnosis. The multiple roles for assessment illustrate the need to look beyond standardized test scores and take a closer look at classroom assessments, particularly for the purpose of instructional diagnosis (Black & Wiliam, 1998a). Senator Edward Kennedy echoed the need to focus on instructional diagnosis when asked about the reauthorization of NCLB:

(Kennedy's) priorities include ensuring that tests are used to improve instruction, not just for labeling schools, and ensuring that students struggling to meet high standards have the best teachers and expanded learning opportunities (Karhuse, 2007, p. 1).

Assessment is defined as “the process of gathering, interpreting, recording, and using information about pupils’ responses to educational tasks” (Lambert & Lines, 2000, p. 4). The three types of assessment are described as a) official, like standardized tests used for NCLB, b) summative, and c) formative, both used by the classroom teacher to

inform instruction and evaluation (Airison, 2000). While many studies have focused on large-scale standardized testing programs, classroom assessments have received increased attention (Popham, 2002; Airisian, 2000; 1991; Black & Wiliam, 1998a; Stiggins, 2001, 1997; Stiggins & Conklin, 1992; Crooks, 1988).

Also influential in assessment practices, research has shown that principals as instructional leaders directly impact student achievement (Hallinger & Heck, 2000; Smith & Andrews, 1989; Andrews, Basom, & Basom, 1991). The review of literature revealed that instructional leader attributes and job functions, instructional leader knowledge, and the ability of instructional leaders to provide teacher development and intellectual stimulation have direct relationships with student achievement (Marzano, Waters & McNulty, 2005; Valentine, et. al, 2004; Smith & Andrews, 1989).

There is a need to take a more focused look at classroom assessment practices and student achievement as measured by official standardized tests. Additionally, it would be beneficial to examine the relationships between assessment practices and the levels of principal instructional leadership, particularly those related to student achievement as measured by official standardized tests.

Purpose of the Study

The purpose of this study was to investigate the relationships among teachers' classroom assessment practices, instructional leadership practices, and student achievement as measured by official standardized tests. Teachers' assessment was measured by items, scales and factors describing existing assessment practices. Instructional leadership practices were measured by two factors describing the degree of instructional and curriculum leadership of the principal.

Research Questions

To accomplish the purpose of this study, the following research questions were developed:

1. What are the types and levels of assessment practices used in Missouri's middle level schools according to teachers' perceptions?
2. What are the levels of instructional leadership for principals in Missouri's middle level schools?
3. Are there relationships between assessment practices and principal instructional leadership in Missouri's middle level schools?
4. Are there relationships between the types and levels of assessment practices and the levels of principal instructional leadership on student achievement in communication arts and mathematics according to teachers' perceptions?
5. Are there differences in the levels of principal instructional leadership and the types and levels of assessment practices between schools with high levels of achievement and schools with low levels of achievement in communication arts and mathematics?

Hypotheses

The following hypotheses were tested in this study:

H_{01} : There are no significant correlational relationships among the items, scales, and factors measuring assessment practices, the factors measuring principal instructional leadership and student achievement as measured by the Communication Arts and Mathematics portions of the Missouri Assessment Program.

H_{02} : There are no significant predictive linear relationships for the factors measuring principal instructional leadership, and the scales and factors measuring assessment practices.

H_{03} : There are no predictive linear relationships among the scales and factors measuring assessment practices and the factors measuring principal instructional leadership on student achievement in communication arts and mathematics.

H_{04} : There are no significant differences for the scales and factors measuring assessment practices and the factors measuring principal instructional leadership between schools with high levels of achievement and schools with low levels of achievement in communication arts and mathematics.

Population

This study examined teachers' assessment practices and instructional leadership present at the middle level and determined the relative influence each had on the other and on student achievement. For the 2006-2007 school year, the number of middle level Missouri schools with any combination of grades between grades five through nine was 371. This study included only middle level schools in Missouri that had at least two grade levels with either a grade seven or a grade eight. The distribution of middle level schools, by grade level configuration, and the percent of the population for schools serving middle level students in the state of Missouri are provided in Table 2.

Table 2
Middle Level Schools with Grades 5 through 9 in Missouri, 2006-2007

Grade Configuration	Number	Percent
5	1	0.27
5-6	21	5.66
5-6-7	2	0.54
5-6-7-8	48	12.94
5-6-7-8-9	1	0.27
6	4	1.08
6-7	9	2.43
6-7-8	188	50.67
6-7-8-9	1	0.27
7	0	0.00
7-8	70	18.87
7-8-9	7	1.88
8	2	0.54
8-9	9	2.43
9	8	2.15
Total Missouri Middle Level Schools	371	100.00
Total meeting criteria to participate in this study	336	90.84

(Missouri Department of Elementary and Secondary Education, 2006)

In the fall of 2006, the Middle Level Leadership Center began a statewide study of teachers' assessment and instructional leadership practices in Missouri middle level schools. All Missouri middle level schools serving either grade seven or eight with more

than one grade level were asked to participate in the study. Teachers from the 336 middle level schools in the state were asked to respond to an extensive survey about their assessment practices and their beliefs about their school leader's instructional leadership. For the purposes of this study the unit of analysis was the school level.

As a graduate student research assistant at the Center, this researcher participated in the design of the study, completed the research and literature review for the portion of the larger study represented in this study, designed portions of the instrumentation for this study, and participated in all aspects of the data collection and analysis process for this portion of the larger study. The letter of request for permission to have access to the data is provided in Appendix A.

Procedure

To accomplish the purpose of this study, the following procedures were followed. For all statistical tests, significance was set at $\alpha = .05$.

1. Use of a cross-sectional design. This study examined the influence of leadership and assessment practices by using a cross-sectional design. A cross-sectional design is a study in which data are collected from many different kinds of people from multiple groups or types of people in a relatively short period of time (Johnson & Christensen, 2000). A cross-sectional design was appropriate for this study due to the nature of the inquiry about teachers' assessment practices, leadership and student achievement.
2. Correlation, regression, and tests of differences were used to test null hypotheses and identify the statistical relationships and differences for the study variables. All analyses studied the selected variables of instructional

leadership and assessment practices and student achievement as measured by the Missouri Assessment Program (MAP) Test for Communication Arts and Mathematics.

3. All analyses were made while controlling for the effect of socioeconomic status using the free/reduced lunch data as the controlling variable.

Instrumentation

Student achievement data were analyzed using reported test results from the Missouri Assessment Program (MAP). The MAP test is a performance-based assessment system used to measure student achievement, which is administered annually to all students in all middle level grades in Missouri Public schools. Missouri's State Board of Education (1997) designated the purposes of the assessment program as:

- improving students' acquisition of important knowledge, skills, and competencies
- monitoring the performance of Missouri's educational system
- empowering students and their families to improve their educational prospects
- supporting the teaching and learning process.

Students in all middle level schools are administered a Communication Arts and Mathematics assessment. Test items fall in three categories: multiple-choice, constructed-response, and a performance event. Multiple-choice questions require selecting the correct answer while short-answer, constructed response questions ask the student to supply the appropriate response. The performance event requires students to focus on an issue with problem solving as the focus for assessment (Missouri Department of Elementary and Secondary Education, 2006).

The dependent data, student achievement, were obtained from the web site of the Missouri Department of Elementary and Secondary Education. The achievement data used for this study were from the state assessment administered in the spring of 2006.

Middle Level Teacher Assessment Practices Survey

The independent data for this survey were collected through survey format. The Middle Level Teacher Assessment Practices Survey (TAPS-ML), developed at the Middle Level Leadership Center by the researcher and Center Director, was used to collect the data from middle level teachers about their self-perceptions of their classroom assessment practices. The instrument was developed from the review of literature. The items of the TAPS-ML were developed into five scales. Subsequent analysis of the items produced four factors. A copy of the instrument and its items, grouped by scale, are in Appendix B. The scales of the Middle Level Teachers' Assessment Practices Survey were:

Assessment during Instruction: Strategies teachers employ before and during assessments. Planning is essential along with the development of a means of communication for students to understand what is expected of them prior to an assessment task, such as providing detailed scoring guides. Talk and questioning are productive forms of assessing during instruction (Black, 1998).

Teacher adaptation of instruction per assessment data: Teachers use assessment data to determine if they need to reteach, or possibly adjust instruction.

Assessments are used to change or improve ongoing classroom processes while learning is still in progress (Airasian, 2000, p. 94).

Teacher feedback per assessment data: Teacher provides information to students; in which the information is used to close a gap, referring to the level in which students are performing and the desired level of achievement (Sadler, 1989).

Teacher development of peer and self-assessment strategies: Students are trained to judge their own level of performance on an assessment task, as well as their peers. When students become involved there is a shift from passive receptivity to active involvement in learning (Sheperd, 2005; Ellis, 2001).

Teachers use of summative assessments: Teachers take summative data, but rather than use it to make a final judgment about student performance, it is used formatively to provide students with the information they need to improve achievement (Fox, 2003).

Four factors were developed during the course of this study. The following definitions were constructed by the instrument authors for the purpose of defining the major concepts of the factors of the Middle Level Teachers Assessment Practices Survey.

Preparation for Formative Assessment: This factor describes the actions teachers take when planning for classroom assessments. These may include providing a scoring guide, providing previous work examples to guide students thinking about developing a successful product, and careful planning of questions and discussions so responses provide insight into student learning. The focus of this preparation work is to ensure that students clearly understand the desired goal of a classroom assessment and how to achieve the desired goal.

Application of Formative Assessment: This factor describes the actions teachers take while in the process of instruction. The focus is on re-teaching and adjusting

instruction to encourage mastery of content, providing feedback to students and parents through written comments and conversation, and engaging students in the process of self-assessment.

Impact of Summative Assessment: This factor describes the analyses of final summative data at the building level, the teacher and classroom level, and the individual student level. Strengths and weaknesses are established forming the basis for change at the classroom and school-wide levels. Changes may occur in instructional practices, curriculum, or programs.

Collaboration through Assessment: This factor describes the process of teachers coming together with the principal and colleagues to discuss and analyze classroom assessment data. Planning for instructional and student achievement are the focus for discussions.

Principal Instructional Leadership

Principal instructional leadership was measured using the factors of the Educational Program domain portion of the Audit of Principal Effectiveness (Valentine & Bowman, 1988). The measure was completed by teachers and was based on teachers' perceptions of the principal's instructional leadership behavior. A copy of the instrument and its items, grouped by factor, are in Appendix C. The Educational Program domain contains 15 items, and it measures teachers' perceptions of the following two factors that underlie the construct of instructional leadership:

Instructional Improvement: the principal influences positively the instructional skills present in the school through clinical supervision, knowledge of effective schooling, and commitment to quality instruction. This factor contains eight

items, and it has a reported reliability coefficient (Chronbach's alpha) of .85 (Bowman & Valentine, 1984).

Curricular Improvement: the principal promotes an articulated, outcome-based curriculum through diagnosis of student needs and systematic program review and change. This factor contains seven items, and it has a reported reliability coefficient (Chronbach's alpha) of .84 (Bowman & Valentine, 1984).

Data Collection

Data collection packets were mailed to each of the 336 Missouri middle level schools in January, 2007. Each survey box included a survey packet for the principal, the secretary and each of the classroom teachers in the school. The principal's packet included a cover letter of invitation (Appendix A), a brief principal and school demographic survey (Appendix B), and a written consent form (Appendix C). The secretary's packet included directions to distribute, collect, and return the completed surveys and an explanation of the secretary incentive process for recognition of the schools with the highest percentage of completed returns (Appendix D). Each teacher packet included a "short-explanation" cover letter and a longer explanatory letter about the study, directions for completing and returning the survey, a written consent form, and a copy of the teacher survey (Appendix E and F). The surveys used in this study were the Middle Level Teachers' Assessment Practices Survey and the two factors of the Educational Domain of the Audit of Principal Effectiveness (Appendix G and H).

School secretaries were asked to return completed surveys in one return envelope within two weeks or whenever about half of the teachers had completed their surveys. The secretary was then asked to return all surveys within a month and no later than 40

days after the boxes were mailed to the school. Upon receipt of the survey packets, staff at Middle Level Leadership Center (MLLC) entered respondent data into excel spreadsheets, analyzed the data for errors, and then transferred the data into Statistical Package for the Social Sciences (SPSS), version 13. All written consent forms were filed and secured to maintain confidentiality. All data were encrypted and secured to protect respondent confidentiality. All teacher responses were compiled by school to create a school variable for each factor and all analyses were made with the school as the unit of analysis.

Data Analysis

Data for this study were quantitative. Correlations, regressions, and tests of differences were used to analyze the data for the hypothesis in this study. The level of significance for all statistical tests was set at $\alpha = .05$.

Hypothesis one was analyzed using Pearson product moment zero-order correlations and partial correlations using free and reduced lunch percentages as the control variable for socioeconomic status. Hypothesis two was analyzed using stepwise linear regression of the independent variables, the factors measuring principal instructional leadership and the free and reduced lunch percentages for each school, against the dependent variables for Teachers' Assessment Practices, specifically, Collaboration through Assessment, Assessment during Instruction, Use of Summative Assessment Data and the rating of Overall Collective Assessment Practices. The free and reduced lunch percentages for each school were entered as independent variables to account for the effect of socioeconomic status (SES) in the equation. Hypothesis three was analyzed using stepwise linear regression of the independent variables, the scales and

factors measuring assessment practices, the factors measuring principal instructional leadership and the free and reduced lunch percentages for each school, against the dependent variables, student achievement. Again, free and reduced lunch percentages for each school were entered as independent variables to account for the effect of socioeconomic status (SES) in the equation. The regressions for hypothesis two and three were run for both communication arts and mathematics student achievement. The fourth hypothesis was analyzed using analysis of variance. To conduct this analysis, the schools in the study were sorted into quartiles, based upon the variables of mathematics and communication arts student achievement.

The achievement data used in this study for each participating school was computed by identifying the percent of students at each grade level who mastered the state assessment and thus creating a school average of those who reached master by averaging the grade level percentages. This process established a composite school achievement score each school. All data analyses were performed using SPSS version 14.0. Following the analysis of all hypotheses, an explanatory model was developed to describe the relative impact of teachers' assessment practices and instructional leadership on student achievement in mathematics and communication arts.

CHAPTER 4

PRESENTATION AND ANALYSIS OF DATA

Introduction

In recent years, scores from official standardized tests have become the marker of successful schools to policy makers, and the public, thus to educators. Considering the unbalanced emphasis on standardized test scores over classroom assessment data, it is no surprise that teachers and principals have emerged with varying levels of assessment literacy (Popham, 2004). Teachers are asked to use different assessment and testing frameworks to please different audiences - such as parents, administrators, the community and the federal government, serve different purposes, and use different procedures to collect information (Farr, 1992). This leaves teachers, administrators and students to wonder how to prioritize their educational goals for classroom assessment practices, as well as preparation for official standardized tests.

According to Stiggins (1999), there is a need for official standardized tests because policy and programmatic decision makers at the district, state and national level need test results that are comparable across classrooms. However, educators need to rethink the relationship between testing and effective schooling. “This reevaluation must center not on how we assess student achievement but on how we use assessment in pursuit of student success” (Stiggins, p. 191). Researchers have indicated that classroom assessment practices positively impact student achievement (Rodriguez, 2004; Black & Wiliam, 1998b; Bloom, 1984).

In the seminal literature review by Black and Wiliam (1998b), classroom assessment practices were determined to be a strong indicator for improving student

achievement as measured by various factors such as teacher/researcher observations, summative assessment data such as the Third International Mathematics and Science Study (TIMSS) and student self-perception survey scores. Classroom assessment practices have been determined from a review of the literature to include instructional practices (Black, 1998), teacher adaptation of instruction (Airasian, 2000), feedback (Sadler, 1989), peer and self-assessment (Ellis, 2001) and the formative use of summative data (Fox, 2003). Few studies have taken a focused look at the specific types and levels of classroom assessment practices and student achievement as measured by official standardized tests.

While much of the recent research on assessment has focused on large-scale standardized testing and classroom assessment practices, the school principal's role in assessment has been largely ignored (Noonan & Renihan, 2006). The role that the principal plays in teachers' assessment practices, which would fall under the umbrella of responsibilities of instructional leadership, has been greatly influenced by assessment reform in the United States (Noonan & Renihan). Expectations are evident that the principal cannot neglect the role of assessment. "The only critical question remaining is how administrators will be involved" (Cizek, 1995, p.248). One definite indicator is the expectation by the Interstate School Leaders Licensure Program (ISSLC) that the "school administrator promote the success of all students by advocating, nurturing and sustaining an instructional program conducive to student learning" (Green, 2001).

Study Design

The purpose of this study was to investigate the relationships among teachers' classroom assessment practices, instructional leadership practices, and student

achievement as measured by standardized test scores. The method of analysis was quantitative, with survey data being used to determine, (a) what is descriptive of the demographics of Missouri middle level schools, (b) what is descriptive of the assessment practices of teachers in Missouri middle level schools, (c) what is descriptive of the instructional leadership practices of the school leader, and (d) if any relationships exist between the teachers' assessment practices, the principal as instructional leader and student achievement. Supplementing the quantitative data were data obtained through written responses to open-ended questions about assessment practices.

Two surveys were used in this study. The Middle Level Teacher Assessment Practices Survey (TAPS-ML), developed by the researcher and the Middle Level Center Director, was used to collect data from middle level teachers about their self-perceptions of their classroom assessment practices. The instrument was developed from the review of the literature on classroom assessment practices. The survey was organized around five scales: (a) instructional strategies, (b) teacher adaptation of instruction per assessment data, (c) teacher feedback per assessment data, (d) teacher development of peer and self-assessment strategies, and (e) the formative use of summative assessments. The instrument was piloted during the scale development phase of this study. After collection of data for this study, the instrument data were factor analyzed. Throughout this chapter, data are repeated for both the scale constructs and the factors.

Principal instructional leadership was measured using the factors of the Educational Domain portion of the Audit of Principal Effectiveness (Valentine & Bowman, 1988). The measure was completed by teachers and was based on teachers' perceptions of the principal's instructional leadership behavior. The two factors that

measured instructional leadership were: (a) instructional improvement and (b) curricular improvement. Student achievement data, as measured by the Missouri Assessment Program (MAP) tests, were obtained from the Missouri Department of Elementary and Secondary Education website for Communication Arts and Mathematics. The achievement data for this study were from the state assessment administered in the spring of 2006. The achievement data used in this study, for each participating school was computed by identifying the percent of students at each grade level who mastered the state assessment and thus creating a school average of those who reached mastery by averaging the grade level percentages. This process established a composite school achievement score for each school.

In the fall of 2006, principals from Missouri's 371 middle level schools were contacted and asked to participate in a statewide study of middle level teachers' assessment practices and principals' instructional leadership qualities. Of the 371 schools, 336 fit the criteria for this study of having either grade 7 or grade 8 or both.

Research Questions

To accomplish the purpose of this study, the following research questions were developed:

1. What are the types and levels of assessment practices used in Missouri's middle level schools according to teachers' perceptions?
2. What are the levels of instructional leadership for principals in Missouri's middle level schools?
3. Are there relationships between assessment practices and principal instructional leadership in Missouri's middle level schools?

4. Are there relationships between the types and levels of assessment practices and the levels of principal instructional leadership on student achievement in communication arts and mathematics according to teachers' perceptions?
5. Are there differences in the levels of principal instructional leadership and the types and levels of assessment practices between schools with high levels of achievement and schools with low levels of achievement in communication arts and mathematics?

Hypotheses

The following hypotheses were tested in this study:

H_{01} : There are no significant correlational relationships among the items, scales, and factors measuring assessment practices, the factors measuring principal instructional leadership and student achievement as measured by the Communication Arts and Mathematics portions of the Missouri Assessment Program.

H_{02} : There are no significant predictive linear relationships for the factors measuring principal instructional leadership, and the scales and factors measuring assessment practices.

H_{03} : There are no predictive linear relationships among the scales and factors measuring assessment practices and the factors measuring principal instructional leadership on student achievement in communication arts and mathematics.

H_{04} : There are no significant differences for the scales and factors measuring assessment practices and the factors measuring principal instructional leadership

between schools with high levels of achievement and schools with low levels of achievement in communication arts and mathematics.

Descriptive Findings

Demographic data are discussed in this section. Four categories are shared: (a) demographic data, (b) teacher demographics, (c) leader demographics, and (d) factor analysis of the items on the Middle Level Teachers' Assessment Practices (TAPS-ML).

Demographic Data

In the 2006-2007 school year, 336 middle level schools in the state of Missouri fit the criteria for this study. Of those 336 middle level schools, 187 schools provided data for this study. After removing schools with first year principals and schools where at least 40% of the faculty did not provide usable returns, 123 schools became the data set for this study. The percent of the 123 schools represented a 37% sample of the state's middle level schools. The percent of respondents and statewide percentage of schools by grade level is provided in Table 3.

Table 3

<i>Grade Level of Responding Schools and Statewide Comparison</i>				
Grade Levels	Number of Respondents	Percent of Respondents	Number Statewide	Statewide Percentage
5-6-7-8	18	14.4	48	14.32
6-7-8	64	51.2	188	56.12
7-8	27	21.6	70	20.90
7-8-9	1	.8	7	2.18
8-9	5	4	9	2.6
Other combinations of grade levels	8	6.4	13	3.88
Total	123	100.00	336	100.00

Achievement data were obtained for all Missouri middle level schools in Communication Arts and Mathematics for the 2005-2006 school year from the Missouri Department of Elementary and Secondary Education's website. The state reports the student scores in four levels: (a) below basic, (b) basic, (c) proficient and (d) advanced. Percentages from the top two levels, proficient and advanced, were averaged to create the composite school achievement for both Communication Arts and Mathematics used for this study. Communication Arts and Mathematics achievement scores are presented in Table 4 and Table 5 respectively. For communication arts, the study mean for grade 8 was 44.36 while the state mean was 42.95. In grade 7 for communication arts, the study mean was 47.35 and the state mean was 45.07. Grade 6 communication arts scores for the study averaged to be 44.17 and the state mean was 42.73. In grade 8, the study mean for Mathematics was 43.42, while the state mean was 41.88. In grade 7, the study mean for

mathematics was 41.88, and the state mean was 45.63. In grade 6, the study mean for mathematics was 46.26 and the state mean was 44.64.

Table 4

Missouri Assessment Program Student Achievement Test, 2006
Communication Arts

Measure	Mean	SD	Minimum	Maximum
<u>Study Schools</u> Communication Arts Grade 8	44.36	11.77	14.60	87.30
<u>Statewide Schools</u> Communication Arts Grade 8	42.95	13.71	4.20	87.30
<u>Study Schools</u> Communication Arts Grade 7	47.35	12.4	11.50	81.30
<u>Statewide Schools</u> Communication Arts Grade 7	45.07	14.32	2.10	81.30
<u>Study Schools</u> Communication Arts Grade 6	44.17	12.29	13.00	86.40
<u>Statewide Schools</u> Communication Arts Grade 6	42.73	14.31	1.70	86.40

Table 5

Missouri Assessment Program Student Achievement Test, 2006
Mathematics

Measure	Mean	SD	Minimum	Maximum
<u>Study Schools</u> Mathematics Grade 8	43.42	13.10	11.40	71.80
<u>Statewide Schools</u> Mathematics Grade 8	41.88	15.04	.00	73.90
<u>Study Schools</u> Mathematics Grade 7	47.62	14.54	3.90	79.70
<u>Statewide Schools</u> Mathematics Grade 7	45.63	15.64	2.80	79.70
<u>Study Schools</u> Mathematics Grade 6	46.27	14.80	7.10	85.20
<u>Statewide Schools</u> Mathematics Grade 6	44.64	16.35	1.40	85.20

Communication Arts and Mathematics scores were averaged across grades six through eight. The averages are presented in Table 6. In both subjects, the study schools scored slightly higher than the average for all of the schools in Missouri.

Table 6

Total Averages of Study and Statewide Schools

<u>Study Schools</u> Communication Arts Total	<u>Statewide Schools</u> Communication Arts Total	<u>Study Schools</u> Mathematics Total	<u>Statewide Schools</u> Mathematics Total
45.29	43.55	45.77	44.05

Overall, the schools in the study were similar to the larger population for the student achievement measures and for the demographic measures. The percent of students receiving a free or reduced priced lunch in Missouri's middle level schools was 41.80, a little less than one percentage point higher than the study average of 41.05. The state average for building enrollment was 509.28 which was slightly higher than the study average of 486.56. The state average for full time equivalent employees was 41.79, which was a little higher than the study average of 40.27. Attendance was very similar in comparison with the statewide average at 93.73 and the study average at 93.42. The comparison of survey respondents to the population of middle level schools in the state of Missouri is presented in Table 7.

Table 7

<i>Comparisons of Respondents to State of Missouri, Selected Demographic Variables</i>				
Measure	Study Average	State Average	Study SD	State SD
Percent of Students receiving free/reduced priced lunch	41.05	45.30	18.490	19.80
Building Enrollment	486.56	504.80	249.833	275.51
Faculty Full Time Equivalent	40.27	41.79	48.47	42.96
Average Daily Attendance	93.42	93.73	11.17	9.4
(SD – Standard Deviation)				

Teacher Demographics

The survey data for this study were completed by 779 teachers in the 123 middle levels schools that participated in this study. Seventy-two percent of the teacher respondents were female with 25 percent male and 487 teachers or 62 percent responded positively when asked if they had taken an assessment course as part of teacher certification, while 264 or 34 percent replied that they had not taken a course. Fifty-one percent of the teacher respondents teach core subjects with 27 teaching non-core subjects. Special education teachers followed with 14 percent and five percent taught in another area. Forty-six percent of the middle level teacher respondents taught in a team of four or more teachers and 11 percent taught on a team of two or more. Twenty-eight percent were not assigned to a team and 11.6 were assigned to another teaching configuration. Demographic information about the teacher respondents are presented in Tables 8, 9, 10, and 11.

Table 8

Teacher's Gender (N=779)

Gender	Number	Percent
Female	561	72
Male	193	25
Not Reported	25	3
Total	779	100

Table 9

<i>Assessment course taken for teacher certification (N=779)</i>		
Response	Number	Percent
Yes	487	62
No	264	34
Not Reported	28	4
Total	779	100

Table 10

<i>Teaching area (N=779)</i>		
Area	Number	Percent
Core subjects	398	51
Non-core subjects	209	27
Special Education	106	14
Other	43	5
Not Reported	23	3
Total	779	100

Table 11

Teaching assignment (N=779)

Team	Number	Percent
Team of 4+	359	46
Team of 2+	88	11
No Team	218	28
Other	91	12
Not Reported	23	3
Total	779	100

Leader Demographics

Demographic data about the study schools were reported by the principals and confirmed by the data reported by the Missouri State Department of Education website. Principals provided basic demographic information about their professional experience. The mean age of the principals in this study was 46. Total years as an educator averaged eleven years. The mean total years served as a principal at any level was seven, while the average years as a middle level principal were seven. Demographic information about the principal respondents is presented in Table 12.

Table 12

<i>Leadership Demographics (N=123)</i>					
Item	Mean	SD	Variance	Minimum	Maximum
Principal Age	46	7.963	63.407	30	66
Years in Education	11	6.036	36.434	3	28
Years as a Principal	8	5.083	25.837	0	30
Years as a Middle Level Principal	7	4.657	21.686	0	30

Factor Analysis

Through an extensive literature review, items for the MLTAP were developed. The items were clustered logically, based upon the literature, into theoretical constructs referred to as “scales” throughout this study. The teachers’ responses to the initial set of 38 Likert-type assessment items were factor analyzed. The following describes the factor analysis process.

From the literature, six scales were developed. These scales comprised 38 items. The 38 items were designed to provide forced choice response using a Likert-type scale of: (a) Never (0%); (b) Infrequently (1 to 25%); (c) Sporadically (25 to 50%); (d) Occasionally (50 to 75%); (e) Frequently (75 to 99%); and (f) Always (100%). Four additional items, numbers 39 through 42, were short answer, open-ended questions.

Table 13

<i>Assessment Items Clustered by Original Scale</i>	
Scale Description	Items
Instructional strategies	1,2,3,4,5,6,7,8,9
Teacher Adaptation of Instruction per Assessment Data	10,11,12,13,14,15
Teacher Feedback per Assessment Data	16,17,18,19,20,21,22,23
Teacher Development of Peer and Self-Assessment Strategies	24,25,26,27
The Formative Use of Summative Assessments	28,29,30,31,32,33,34,35,36,37,38
Open-ended questions	39, 40, 41, 42

Factor analysis using the extraction method of Principal Component Analysis with a rotational method of Varimax with Kaiser Normalization was conducted on respondent data. For inclusion on a factor, the items had to load at .40 or higher, and if the items cross-loaded, the difference between the loadings had to be .15 or higher to be included in the factor. The factor names, items, and item loadings defined from the data and used in this study are listed in Table 14.

Table 14

<i>Factors Identified and Loadings</i>		
Item Number	Item	Loading
<u>Factor 1</u>	<u>NAME: Preparation for Formative Assessment</u>	
1	I inform students about the expectations for their learning at the very beginning of the learning process	.543
2	I tell my students, in language that they understand, the objectives they are going to be learning	.488
3	I provide a detailed scoring guide to students before engaging them in the learning experience(s), so they understand the learning expectations for the content being studied.	.781
4	Before instruction, I provide previous work examples for students to analyze to guide their thinking about what a successful product looks like.	.530
7	I provide a detailed scoring guide as a means for students to engage in peer and self assessment.	.730
8	I carefully plan the questions that I ask in class so responses to questions give me good insight about student understanding.	.464
18	I use a detailed scoring guide to assess students' performance on a learning task.	.748
19	I work to ensure that my students clearly understand their level of performance on a classroom assessment.	.633
20	I work to ensure that my students clearly understand the desired goal of a classroom assessment and how to achieve the desired goal.	.629
<u>Factor 2</u>	<u>NAME: Application of Formative Assessments</u>	
12	When I teach a unit or concept and I have students who don't thoroughly understand the content, I design and implement different instructional strategies and experiences.	.630
13	When I teach a unit or concept and some of my students don't	.723

thoroughly understand the content, I provide individualized learning experiences so those who don't understand have more time to develop that understanding and those who do understand have time for enrichment experiences.

15	I teach, re-teach and continue to work with students until each student has truly mastered the learning concept.	.718
21	I meet with students individually at least once a week to discuss their performance.	.650
23	I explain to parents the meaning of their child's assessment scores and how those scores relate to their child's growth as a learner.	.540
24	Students in my class are comfortable talking with each other and identifying strengths as well as areas to improve.	.471
25	My students write reflections about their growth and learning over a period of time for classroom objectives.	.461
26	I use conversation or written correspondence with my students to discuss, back and forth, the student's level of performance in my class.	.547

Factor 3

NAME: Impact of Summative Assessments

28	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused me to implement changes in instructional delivery .	.560
29	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused me to implement changes in curriculum .	.703
30	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused me to implement changes in the programs my school supports.	.723
31	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests	.719

(e.g. writing assessments, SRI, DRA, math objective tests), have caused **my school** to implement changes in **instructional delivery**.

- | | | |
|----|--|------|
| 32 | Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused my school to implement changes in curriculum . | .792 |
| 33 | Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused my school to implement changes in the programs my school supports. | .781 |

<u>Factor 4</u>	<u>NAME: Collaboration Through Assessment</u>
-----------------	---

- | | | |
|----|---|------|
| 34 | I meet with my principal, colleagues, and/or team to analyze formal assessment data (e.g. writing assessments, SRI, DRA, math objective tests), to plan for further instruction | .677 |
| 35 | I meet with my principal, colleagues, and/or team to analyze informal classroom assessment data (e.g. weekly assignments, daily quizzes, and performance assessments) to plan for further instruction . | .741 |
| 36 | I meet with my principal, colleagues, and/or team to analyze and discuss formal assessment data (e.g. writing assessments, SRI, DRA, math objective tests), to look for individual students that need extra help. | .620 |
| 37 | I meet with my principal, colleagues, and/or team to analyze informal classroom assessment data (e.g. weekly assignments, daily quizzes, and performance assessments) to look for individual students that need extra help. | .685 |
-

The partial correlations, controlling for free and reduced price lunch, for the factors identified from the factor analysis are provided in Table 15. All factors were moderately or highly correlated. “Preparation for Assessment” positively correlated with “Application of Formative Assessment” ($r = .439, p = .000$), “Impact of Summative Assessment” ($r = .318, p = .000$), and “Collaboration through Assessment” ($r = .358, p$

= .000). “Application of Formative Assessment” positively correlated with “Preparation for Assessment” ($r = .439, p = .000$), “Impact of Summative Assessment” ($r = .345, p = .000$), and “Collaboration through Assessment” ($r = .305, p = .000$). “Impact of Summative Assessment” positively correlated with “Preparation for Assessment” ($r = .318, p = .000$), “Application of Formative Assessment” ($r = .345, p = .000$) and “Collaboration through Assessment” ($r = .541, p = .000$). “Collaboration through Assessment” positively correlated with “Preparation for Assessment” ($r = .358, p = .000$), “Application of Formative Assessment” ($r = .305, p = .000$) and “Impact of Summative Assessment” ($r = .541, p = .000$).

Table 15

<i>Factor Correlations with zero-order and partial controlling for Free/Reduced Lunch</i>				
	TAPS-ML Preparation for Assessment	TAPS-ML Application of Formative Assessment	TAPS-ML Impact of Summative Assessment	TAPS-ML Collaboration through Assessment
TAPS-ML Preparation for Assessment	1.00			
TAPS-ML Application of Formative Assessment	$r = .442^{**}$ $p = .000$ $pr = .439^{**}$ $p = .000$	1.00		
TAPS-ML Impact of Summative Assessment	$r = .323^{**}$ $p = .000$ $pr = .318^{**}$ $p = .000$	$r = .359^{**}$ $p = .000$ $pr = .345^{**}$ $p = .000$	1.00	
TAPS-ML Collaboration through Assessment	$r = .356^{**}$ $p = .000$ $pr = .358^{**}$ $p = .000$	$r = .299^{**}$ $p = .000$ $pr = .305^{**}$ $p = .000$	$r = .531^{**}$ $p = .000$ $pr = .541^{**}$ $p = .000$	1.00

* $p < .05$, ** $p < .01$

Hypothesis Testing

Four hypotheses were tested in this study. Hypothesis one was tested by conducting correlational relationships among variables measuring middle level teacher assessment practices, principal instructional leadership characteristics and student achievement in Missouri middle level schools using zero-order correlations and partial controlling for socioeconomic status. The second and third hypotheses were tested by conducting regression analyses for the variables measuring middle level teacher assessment practices and principal instructional leadership characteristics when controlling for socioeconomic status. The fourth hypothesis was tested by using a test of differences for the variables when sorting the schools into achievement quartiles for Communication Arts and Mathematics.

Hypothesis One

The first hypothesis tested in this study was: There are no significant correlational relationships among the items, scales, and factors measuring assessment practices, the factors measuring principal instructional leadership and student achievement as measured by the Communication Arts and Mathematics portions of the Missouri Assessment Program. Pearson-product moment (zero-order) correlations and partial correlations controlling for the percent of students receiving free or reduced priced lunch were calculated for the relationship among the scales and factors from the Middle Level Teachers Assessment Practices Survey, the two factors of the Educational Program domain portion of the Audit of Principal Effectiveness (APE) and the Communication Arts and Mathematics achievement data.

Zero-order and partial correlations for the items from the scale, Instructional Preparation Practice from the Middle Level Teachers' Assessment Practices Survey and the factors, Instructional Improvement and Curriculum Improvement, used to measure instructional leadership are presented in Table 16. Of the five survey items, no items showed significant relationships with the Instructional Improvement factor or Curriculum Improvement factor.

Table 16

Correlations: Middle Level Teachers' Assessment Practices Items for Instructional Preparation Practices Scale with Audit of Principal Effectiveness Factors

#	Item	APE Factor 8 Instructional Improvement		APE Factor 9 Curriculum Improvement	
		Zero- order	Partial	Zero- order	Partial
1	Inform students about expectation	-.101 $p = .264$	-.104 $p = .255$	-.111 $p = .221$	-.114 $p = .210$
2	Tell learning objectives before instruction	-.067 $p = .462$	-.059 $p = .517$	-.088 $p = .329$	-.082 $p = .370$
3	Provide detailed scoring guide	.115 $p = .205$.105 $p = .252$.084 $p = .352$.074 $p = .416$
4	Provide samples of successful products	.003 $p = .970$	-.018 $p = .843$.000 $p = .996$	-.017 $p = .850$
5	Spend time planning instruction	.176 $p = .051$.171 $p = .059$.131 $p = .148$.127 $p = .165$

* $p < .05$, ** $p < .01$

Zero-order and partial correlations for items from the scale, Assessment during Instruction from the Middle Level Teachers' Assessment Practices Survey, and the

factors, Instructional Improvement and Curriculum Improvement, used to measure instructional leadership are presented in Table 17. One item, “teacher does not move on until students have mastered content,” did have a significant zero-order correlation ($r = .263, p = .003$) and a significant partial correlation ($pr = .267, p = .003$) for the Instructional Improvement factor. Additionally, the same item showed a significant zero-order correlation ($r = .206, p = .022$) and a significant partial correlation ($pr = .207, p = .022$) for Curriculum Improvement.

Table 17

Correlations: Middle Level Teachers’ Assessment Practices Items for Assessment during Instruction Scale with Audit of Principal Effectiveness Factors

#		APE Factor 8 Instructional Improvement		APE Factor 9 Curriculum Improvement	
	Item	Zero- order	Partial	Zero- order	Partial
6	Think about students’ responses during discussion	.023 $p = .804$.023 $p = .800$.035 $p = .702$.039 $p = .666$
7	Provide scoring guide for self assessment model	.160 $p = .076$.153 $p = .092$.089 $p = .326$.085 $p = .355$
8	Plan questions before instruction	.026 $p = .771$.027 $p = .770$.029 $p = .745$.033 $p = .714$
9	Provide wait time for answers	.036 $p = .693$.032 $p = .723$.045 $p = .620$.048 $p = .597$
10	Do not move on until students have mastered content.	.263** $p = .003$.267** $p = .003$.206* $p = .022$.207* $p = .022$

* $p < .05$, ** $p < .01$

Zero-order and partial correlations for items from the scale, Adaptation of Instruction per Assessment Data from the Middle Level Teachers’ Assessment Practices

Survey, and the factors, Instructional Improvement and Curriculum Improvement, used to measure instructional leadership are presented in Table 18. Of the five survey items, no items showed significant relationships for the Instructional Improvement factor or Curriculum Improvement factor.

Table 18

Correlations: Middle Level Teachers' Assessment Practices Items for Adaptation of Instruction per Assessment Data Scale with Audit of Principal Effectiveness Factors

#	Item	APE Factor 8 Instructional Improvement		APE Factor 9 Curriculum Improvement	
		Zero- order	Partial	Zero- order	Partial
11	Reteach using same instructional methods	-.090 $p = .321$	-.092 $p = .316$	-.099 $p = .274$	-.101 $p = .269$
12	Reteach using different strategies	-.129 $p = .154$	-.125 $p = .170$	-.157 $p = .082$	-.152 $p = .096$
13	Reteach using individualized instruction	-.061 $p = .503$	-.077 $p = .397$	-.096 $p = .291$	-.108 $p = .238$
14	Design instruction based on assessment outcomes	-.019 $p = .832$	-.025 $p = .780$	-.036 $p = .691$	-.033 $p = .722$
15	Teach, reteach until all students master concept	.021 $p = .818$.017 $p = .856$.056 $p = .534$.052 $p = .567$

* $p < .05$, ** $p < .01$

Zero-order and partial correlations for items from the scale Feedback to Students and Parents from the Middle Level Teachers' Assessment Practices Survey, and the factors, Instructional Improvement and Curriculum Improvement, used to measure instructional leadership are presented in Table 19. One item, "meet with students

individually to discuss performance” was significant on a zero-order correlation ($r = .193, p = .032$) and a partial correlation ($pr = .190, p = .036$).

Table 19

Correlations: Middle Level Teachers’ Assessment Practices Items for Feedback to Students and Parents Scale with Audit of Principal Effectiveness Factors

#	Item	APE Factor 8 Instructional Improvement		APE Factor 9 Curriculum Improvement	
		Zero- order	Partial	Zero- order	Partial
16	Use letter grade for feedback (Reversed item)	.052 $p = .565$.057 $p = .533$.039 $p = .666$.037 $p = .687$
17	Write detailed comments on work	.106 $p = .241$.115 $p = .209$.069 $p = .447$.077 $p = .400$
18	Use scoring guide to assess	.167 $p = .063$.171 $p = .060$.122 $p = .178$.128 $p = .160$
19	Students know their level of performance on assessment	.105 $p = .247$.102 $p = .265$.100 $p = .270$.104 $p = .256$
20	Students understand goal of assessment and how to achieve it	.071 $p = .431$.065 $p = .479$.089 $p = .327$.093 $p = .309$
21	Meet with students individually to discuss performance	.193* $p = .032$.190* $p = .036$.175 $p = .052$.173 $p = .056$
22	Parents know their child’s performance	.116 $p = .200$.112 $p = .221$.103 $p = .257$.102 $p = .264$
23	Explain to parents meaning of assessment scores	.075 $p = .408$.072 $p = .430$.077 $p = .394$.077 $p = .398$

* $p < .05$, ** $p < .01$

Zero-order and partial correlations for items for the scale Development of Self and Peer Assessment from the Middle Level Teachers’ Assessment Practices Survey, and the factors, Instructional Improvement and Curriculum Improvement, used to measure

instructional leadership are presented in Table 20. Of the four survey items, three had significant relationships with the leadership factors. The item, “students talk with others about strengths and areas to improve” was significant for the Instructional Improvement factor for the zero-order correlation ($r = .362, p = .000$) and for the partial correlation ($pr = .190, p = .036$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .333, p = .000$) and a partial correlation ($pr = .377, p = .000$). The item, “students write reflections about learning” was significant for the Instructional Improvement factor for the zero-order correlation ($r = .227, p = .011$) and for the partial correlation ($pr = .223, p = .013$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .227, p = .011$) and a partial correlation ($pr = .222, p = .014$). The item, “teach self-assessment strategies” was significant for the Instructional Improvement factor for the zero-order correlation ($r = .222, p = .013$) and for the partial correlation ($pr = .230, p = .011$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .232, p = .009$) and a partial correlation ($pr = .241, p = .008$).

Table 20

Correlations: Middle Level Teachers' Assessment Practices Items for Development of Self and Peer Assessment Scales with Audit of Principal Effectiveness Factors

#	Item	APE Factor 8 Instructional Improvement		APE Factor 9 Curriculum Improvement	
		Zero- order	Partial	Zero- order	Partial
24	Students talk with others about strengths and areas to improve	.362** $p = .000$.362** $p = .000$.333** $p = .000$.377** $p = .000$
25	Students write reflections about learning	.227* $p = .011$.223* $p = .013$.227* $p = .011$.222* $p = .014$
26	Teacher and student converse, through talk or writing, about performance	.149 $p = .098$.148 $p = .103$.116 $p = .201$.111 $p = .223$
27	Teach self-assessment strategies	.222* $p = .013$.230* $p = .011$.232* $p = .009$.241* $p = .008$

* $p < .05$, ** $p < .01$

Zero-order and partial correlations for items from the scale Use of Summative Assessment Data from the Middle Level Teachers' Assessment Practices Survey, and the factors, Instructional Improvement and Curriculum Improvement, used to measure instructional leadership are presented in Table 21. Of the ten survey items, six were significant. The item, "formal data caused me to change instructional delivery students talk with others about strengths and areas to improve" was significant for the Instructional Improvement factor for the zero-order correlation ($r = .299, p = .001$) and for the partial correlation ($pr = .301, p = .001$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .293, p = .001$) and a partial correlation ($pr = .296, p = .001$). The item, "formal data caused my school to change instructional delivery" was significant for the Instructional Improvement factor for the

zero-order correlation ($r = .257, p = .004$) and for the partial correlation ($pr = .254, p = .005$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .237, p = .008$) and a partial correlation ($pr = .236, p = .009$). The item, “formal data caused my school to change curriculum” was significant for the Instructional Improvement factor for the zero-order correlation ($r = .235, p = .004$) and for the partial correlation ($pr = .232, p = .010$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .209, p = .020$) and a partial correlation ($pr = .210, p = .020$). The item, “I meet to analyze formal data to plan for instruction” was significant for the Instructional Improvement factor for the zero-order correlation ($r = .430, p = .000$) and for the partial correlation ($pr = .432, p = .000$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .416, p = .000$) and a partial correlation ($pr = .432, p = .000$). The item, “I meet to analyze classroom assessment data to plan for instruction” was significant for the Instructional Improvement factor for the zero-order correlation ($r = .267, p = .003$) and for the partial correlation ($pr = .265, p = .003$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .306, p = .001$) and a partial correlation ($pr = .303, p = .001$). The item, “I meet to analyze classroom assessment data to plan for instruction” was significant for the Instructional Improvement factor for the zero-order correlation ($r = .267, p = .003$) and for the partial correlation ($pr = .265, p = .003$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .306, p = .001$) and a partial correlation ($pr = .303, p = .001$). The item, “I meet to analyze classroom assessment data for individual student’s needs” was significant for the Instructional Improvement factor for the zero-order correlation ($r = .254, p = .004$) and for the partial correlation ($pr =$

.263, $p = .003$), as well as for the Curriculum Improvement factor at a zero-order correlation ($r = .296$, $p = .001$) and a partial correlation ($pr = .305$, $p = .001$).

Table 21

Correlations: Middle Level Teachers' Assessment Practices Items for Use of Summative Assessment Data Scales with Audit of Principal Effectiveness Factors

#	Item	APE Factor 8 Instructional Improvement		APE Factor 9 Curriculum Improvement	
		Zero- order	Partial	Zero- order	Partial
28	Formal data caused me to change instructional delivery	.299** $p = .001$.301** $p = .001$.293** $p = .001$.296** $p = .001$
29	Formal data caused me to change curriculum	.154 $p = .088$.152 $p = .094$.120 $p = .185$.125 $p = .171$
30	Formal data caused me to change programs	.154 $p = .087$.152 $p = .095$.133 $p = .140$.141 $p = .122$
31	Formal data caused my school to change instructional delivery	.257** $p = .004$.254** $p = .005$.237** $p = .008$.236** $p = .009$
32	Formal data caused my school to change curriculum	.235** $p = .004$.232** $p = .010$.209* $p = .020$.210* $p = .020$
33	Formal data caused my school to change programs	.166 $p = .066$.164 $p = .072$.152 $p = .091$.154 $p = .091$
34	I meet to analyze formal data to plan for instruction	.430** $p = .000$.432** $p = .000$.416** $p = .000$.432** $p = .000$
35	I meet to analyze classroom assessment data to plan for instruction	.267** $p = .003$.265** $p = .003$.306** $p = .001$.303** $p = .001$
36	I meet to analyze formal data to look for individual student's needs	.152 $p = .091$.151 $p = .098$.134 $p = .139$.130 $p = .155$
37	I meet to analyze classroom assessment data for individual student's needs	.254** $p = .004$.263** $p = .003$.296** $p = .001$.305** $p = .001$

* $p < .05$, ** $p < .01$

Zero-order and partial correlations for item, “overall, our faculty uses assessment to improve instruction and student achievement” from the Middle Level Teachers’ Assessment Practices Survey, and the factors, Instructional Improvement and Curriculum Improvement, used to measure instructional leadership are presented in Table 22. For the factor, Instructional Improvement, there was a zero-order correlation ($r = .353, p = .000$) and a partial correlation ($pr = .357, p = .000$). For the factor, Curriculum Improvement, there was a zero-order correlation ($r = .316, p = .000$) and a partial correlation ($pr = .323, p = .000$).

Table 22

Correlations: Middle Level Teachers’ Assessment Practices Items for Overall Collective Assessment Practices item with Audit of Principal Effectiveness Factors

#	Item	APE Factor 8 Instructional Improvement		APE Factor 9 Curriculum Improvement	
		Zero- order	Partial	Zero- order	Partial
38	Overall, our faculty uses assessment to improve instruction and student achievement.	.353** $p = .000$.357** $p = .000$.316** $p = .000$.323** $p = .000$

* $p < .05$, ** $p < .01$

Pearson-product moment correlations and partial correlations for the scales of the Middle Level Teachers Assessment Practices Survey and instructional leadership, as measured by the Educational Program domain portion of the Audit of Principal Effectiveness are presented on Table 23. One middle level teachers’ assessment practices scale was significant with Instructional Improvement for both the zero-order and partial correlations. “Scale 6 – Use of Formal Summative Assessment Data” had a positive zero-order ($r = .315, p = .000$) and partial correlation ($pr = .314, p = .000$). Two assessment practices scales were significant with Curriculum Improvement for the zero-order

correlation, “Scale 2 –Assessment during Instruction”, ($r = .178, p = .048$) and “Scale 6 – Use of Formal Summative Assessment Data” ($r = .301, p = .001$). Positive partial correlations were found in “Scale 2 –Assessment during Instruction”, ($pr = .180, p = .046$) and “Scale 6 – Use of Formal Summative Assessment Data” ($pr = .305, p = .001$).

Table 23

Correlations: Middle Level Teachers’ Assessment Practices Scales with Audit of Principal Effectiveness

	APE Factor 8 Instructional Improvement		APE 9 Curriculum Improvement	
	Zero-Order	Partial	Zero-Order	Partial
Scale 1 – Instructional Planning Practices	.099 $p = .273$.099 $p = .278$.153 $p = .090$.154 $p = .090$
Scale 2 – Assessment During Instruction	.176 $p = .051$.175 $p = .053$.178* $p = .048$.180* $p = .046$
Scale 3 – Adaptation of Instruction per Assessment	.113 $p = .212$.111 $p = .220$.056 $p = .539$.057 $p = .528$
Scale 4 – Teacher Feedback to Students and Parents	.146 $p = .104$.145 $p = .109$.161 $p = .075$.163 $p = .072$
Scale 5 – Development of Student/Peer Assessments	.105 $p = .244$.107 $p = .240$.107 $p = .237$.106 $p = .242$
Scale 6 – Use of Formal Summative Assessment Data	.315** $p = .000$.314** $p = .000$.301** $p = .001$.305** $p = .001$

* $p < .05$, ** $p < .01$

Pearson-product moment correlations and partial correlations for middle level teachers' assessment practices, as measured by the factors from the Middle Level Teachers' Assessment Practices Survey, and instructional leadership, as measured by the Educational Program domain portion of the Audit of Principal Effectiveness are presented on Table 24. One factor was significant with Instructional Improvement for both the zero-order and partial correlations. "Factor 4 – Collaboration through Assessment" had a positive zero-order ($r = .285, p = .001$) and partial correlation ($pr = .285, p = .001$). One assessment practices factor was significant with Curriculum Improvement for the zero-order correlation, "Factor 4 – Collaboration through Assessment" ($r = .268, p = .003$). Positive partial correlations were found in "Factor 4 – Collaboration through Assessment" ($pr = .268, p = .003$)

Table 24

Correlations: Middle Level Teachers' Assessment Practices Factors with Audit of Principal Effectiveness

	APE Factor 8		APE 9	
	Instructional Improvement Zero-Order	Partial	Curriculum Improvement Zero-Order	Partial
Factor 1 – Preparation for Formative Assessment	.089 $p = .326$.088 $p = .333$.057 $p = .528$.058 $p = .522$
Factor 2 – Application of Formative Assessment	-.017 $p = .849$	-.020 $p = .825$	-.057 $p = .530$	-.055 $p = .543$
Factor 3 – Impact of Summative Assessment	.145 $p = .107$.144 $p = .112$.134 $p = .138$.138 $p = .127$
Factor 4 – Collaboration through Assessment	.285** $p = .001$.285** $p = .001$.268** $p = .003$.268** $p = .003$

* $p < .05$, ** $p < .01$

Pearson-product moment correlations and partial correlations for instructional leadership, as measured by the Educational Program domain portion of the Audit of Principal Effectiveness and student achievement, as measured by the Communication Arts and Mathematics Missouri Assessment Program data, are presented on Table 25. One factor was significant at the partial correlation. “Instructional Improvement” had a positive partial correlation when controlling for free and reduced lunch ($pr = .220$, $p = .044$).

Table 25

Correlations: Educational Improvement Factors from the Audit of Principal Effectiveness with Student Achievement

	APE Factor 8		APE 9	
	Instructional Improvement Zero-Order	Partial	Curriculum Improvement Zero-Order	Partial
Communication Arts Grade 8 Top Two %	.079 $p = .394$.107 $p = .333$.102 $p = .274$.111 $p = .316$
Communication Arts Grade 7 Top Two %	.140 $p = .131$.221 $p = .054$.150 $p = .105$.197 $p = .072$
Communication Arts Grade 6 Top Two %	.079 $p = .456$.115 $p = .298$.096 $p = .365$.128 $p = .246$
Mathematics Grade 8 Top Two %	.137 $p = .140$.203 $p = .064$.163 $p = .078$.180 $p = .101$
Mathematics Grade 7 Top Two %	.148 $p = .110$.220* $p = .044$.161 $p = .082$.194 $p = .077$
Mathematics Grade 6 Top Two %	.147 $p = .161$.208 $p = .058$.137 $p = .193$.183 $p = .096$

* $p < .05$, ** $p < .01$

Pearson-product moment correlations and partial correlations for Middle Level Teachers’ Assessment Practices scales and student achievement in Communication Arts and Mathematics as measured by the Missouri Assessment Program top two percentages

are presented on Table 26. In Mathematics, one factor was significant at the partial correlation. “Teacher Development of Student and Peer Assessments” had a positive zero-order correlation ($r = .213, p = .041$) and a partial correlation when controlling for free and reduced lunch ($pr = .215, p = .050$).

Table 26

Correlations: Middle Level Teachers' Assessment Practices Scales with Student Achievement on Communication Arts and Mathematics

	Scale 1 – Instructional Preparation for Assessment		Scale 2 – Teachers Assessment during Instruction		Scale 3 – Teacher Adaptation for Instruction		Scale 4 – Feedback to Students and Parents		Scale 5 – Teacher Development of Self and Peer Assessments		Scale 6 – Use of Summative Assessment Data	
	<i>r</i>	<i>pr</i>	<i>r</i>	<i>pr</i>	<i>r</i>	<i>pr</i>	<i>r</i>	<i>pr</i>	<i>r</i>	<i>pr</i>	<i>r</i>	<i>pr</i>
CA Gr. 8 Top 2 %	.087 <i>p</i> = .353	.155 <i>p</i> = .161	.034 <i>p</i> = .713	.130 <i>p</i> = .239	-.066 <i>p</i> = .479	.011 <i>p</i> = .923	-.009 <i>p</i> = .921	.096 <i>p</i> = .386	.038 <i>p</i> = .682	.058 <i>p</i> = .603	-.108 <i>p</i> = .247	-.041 <i>p</i> = .741
CA Gr. 7 Top 2 %	.015 <i>p</i> = .870	.100 <i>p</i> = .366	-.024 <i>p</i> = .792	.103 <i>p</i> = .349	-.140 <i>p</i> = .131	-.066 <i>p</i> = .552	-.028 <i>p</i> = .760	.146 <i>p</i> = .186	.059 <i>p</i> = .523	.132 <i>p</i> = .233	.023 <i>p</i> = .808	.132 <i>p</i> = .231
CA Gr. 6 Top 2 %	.056 <i>p</i> = .595	.100 <i>p</i> = .366	.010 <i>p</i> = .922	.103 <i>p</i> = .349	-.063 <i>p</i> = .549	-.066 <i>p</i> = .552	-.039 <i>p</i> = .712	.146 <i>p</i> = .186	.092 <i>p</i> = .385	.132 <i>p</i> = .233	-.122 <i>p</i> = .247	.132 <i>p</i> = .231
MA Gr. 8 Top 2 %	.034 <i>p</i> = .717	.029 <i>p</i> = .794	.004 <i>p</i> = .969	.071 <i>p</i> = .521	-.042 <i>p</i> = .654	.065 <i>p</i> = .558	-.047 <i>p</i> = .618	-.039 <i>p</i> = .722	.027 <i>p</i> = .771	.010 <i>p</i> = .930	-.039 <i>p</i> = .675	.009 <i>p</i> = .932
MA Grade 7 Top 2 %	.024 <i>p</i> = .795	.093 <i>p</i> = .401	-.016 <i>p</i> = .863	.047 <i>p</i> = .671	-.061 <i>p</i> = .510	-.024 <i>p</i> = .829	-.023 <i>p</i> = .805	.104 <i>p</i> = .346	.123 <i>p</i> = .184	.157 <i>p</i> = .153	.005 <i>p</i> = .960	.196 <i>p</i> = .075
MA Grade 6 Top 2 %	.062 <i>p</i> = .554	.142 <i>p</i> = .197	-.023 <i>p</i> = .829	.078 <i>p</i> = .482	-.063 <i>p</i> = .553	-.060 <i>p</i> = .585	.070 <i>p</i> = .507	.186 <i>p</i> = .091	.213* <i>p</i> = .041	.215* <i>p</i> = .050	-.043 <i>p</i> = .684	.087 <i>p</i> = .433

p* < .05, *p* < .01

Teachers' assessment practices factors and student achievement, as measured by the Communication Arts and Mathematics Missouri Assessment Program top two

percentages, produced Pearson-product moment correlations and partial correlations which are presented on Table 27. There were no significant relationships for either zero-order or partial correlations.

Table 27

Correlations: Middle Level Teachers' Assessment Practices Factors with Student Achievement on Communication Arts

	F1: Preparation for Assessment		F2: Application of Formative Assessment		F3: Impact of Summative Assessments		F4: Collaboration through Assessment	
	<i>r</i>	<i>pr</i>	<i>r</i>	<i>pr</i>	<i>r</i>	<i>pr</i>	<i>r</i>	<i>pr</i>
CA Gr. 8	-.015	.039	-.039	-.137	-.137	.026	.027	.031
Top Two %	p = .870	p = .723	p = .677	p = .301	p = .141	p = .813	p = .769	p = .780
CA Gr. 7	-.131	.000	-.091	.111	-.036	.137	.117	.172
Top Two %	p = .158	p = .998	p = .328	p = .316	p = .702	p = .214	p = .207	p = .118
CA Gr. 6	-.077	.037	-.039	.067	-.130	-.015	-.034	.059
Top Two %	p = .463	p = .738	p = .711	p = .546	p = .215	p = .896	p = .750	p = .596
MA Gr. 8	-.006	-.002	-.080	-.012	-.069	.048	.036	.006
Top Two %	p = .949	p = .989	p = .390	p = .911	p = .458	p = .667	p = .701	p = .960
MA Grade 7	-.088	-.028	-.071	.054	-.055	.167	-.088	.052
Top Two %	p = .345	p = .800	p = .328	p = .629	p = .702	p = .129	p = .215	p = .641
MA Grade 6	-.082	-.001	.026	.145	-.088	.052	-.025	.069
Top Two %	p = .463	p = .738	p = .711	p = .187	p = .215	p = .641	p = .750	p = .534

*p < .05, **p < .01

Hypothesis Two

The second hypothesis tested in this study was: There are no significant predictive linear relationships between the factors measuring principal instructional leadership and the scales and factors measuring assessment practices.

Multiple linear regression analysis was the primary analytical technique used to explain the variance in the linear models. Stepwise regression was the regression method conducted to determine the predictor (independent) variables with the best estimate, or predictive power for the selected criterion (dependent) variables. The selection of independent variables, “Instructional Improvement” and “Curricular Improvement”, was based on the findings from the partial correlations of the factors and scales of the Middle Level Teachers’ Assessment Practices Survey. Five factors and scales, “Factor 4 – Collaboration through Assessment”, “Scale 2 – Assessment during Instruction”, “Scale 6 – Use of Summative Data”, “Overall Assessment”, and “Teachers’ Assessment Practices average”, were deemed appropriate for the regressions based on the correlations. Separate stepwise regressions were conducted for each criterion variable that at some point showed influence in the model. For all analyses, the percent of students receiving free or reduced price lunch were entered into the stepwise regression. If the percent of students receiving free or reduced price lunch was significant, then that will be shown, if it is absent that means it was not significant in its ability to account for explained variance.

In total, five stepwise linear regressions were conducted. The statistics presented from the stepwise regressions include the model summary data, specifically, the number of models produced, R , R^2 , R^2 Change, Adjusted R^2 , and Standard Error of the Estimate. The statistics presented for the individual models include sum of squares (SS) (regression,

residual, and total), degrees of freedom, mean square (*MS*), *F* value and significance of *F*. The variable statistics, shown via the coefficients from the models, include the beta value, the standard error of the beta, the standardized beta coefficient (β), the *t* value, and significance of each *t* value.

The results of the multiple regressions for “Collaboration through Assessment” are presented in Table 28. The factor “instructional improvement” was the only factor included in model. The model accounted for 7% of the variance in “Instructional Improvement” (Adjusted $r^2 = .074$, $F = 10.761$, $p = .001$). “Instructional Improvement” was the influential variable in the first model with a standardized beta coefficient (β) of .285 ($p = .001$). Two variables, “percent of students receiving free or reduced price lunch” ($p = .751$) and “Curriculum Improvement” ($p = .872$), did not achieve significance in the model.

Table 28

Linear Regression: Instructional Leadership Factors with Collaboration through Assessment (Factor 4 – TAPS-ML)

Model Summary					
Model	<i>R</i>	R^2	R^2 Change	Adjusted R^2	Std. Error of the Est.
1	.285	.081	.081	.074	.76021
ANOVA					
Model	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	Sig. <i>F</i>
Regression	6.219	1	6.219	10.761	.001
Residual	70.505	122	.578		
Total	76.724	123			
Coefficients					
Variables	<i>B</i>	Std. Error	β	<i>t</i>	Sig. <i>t</i>
(Constant)	2.207	.574		3.844	.000
APE Factor 8 Instructional Improvement	.379	.115	.285	3.280	.001

a. Predictors: (Constant), Instructional Improvement

The results of the multiple regressions for “Assessment during Instruction” are presented in Table 29. The factor “Curriculum Improvement” was the only factor included in model. The model accounted for 2% of the variance in “Assessment during Instruction” (Adjusted $R^2 = .024$, $F = 3.986$, $p = .048$). “Curriculum Improvement” was the influential variable in the first model with a standardized beta coefficient (β) of .178 ($p = .048$). Two variables, “percent of students receiving free or reduced price lunch” ($p = .286$) and “Instructional Improvement” ($p = .740$), did not achieve significance in the model.

Table 29

Linear Regression: Instructional Leadership Factors with Assessment during Instruction (Scale 2 – TAPS-ML)

Model Summary					
Model	R	R^2	R^2 Change	Adjusted R^2	Std. Error of the Est.
1	.178	.032	.032	.024	.35658
ANOVA					
Model	SS	df	MS	F	Sig. F
Regression	.507	1	.507	3.986	.048
Residual	15.512	122	.127		
Total	16.019	123			
Coefficients					
Variables	B	Std. Error	β	t	Sig. t
(Constant)	4.219	.264		15.983	.000
APE Factor 9 Curricular Improvement	.106	.053	.178	1.996	.048

a. Predictors: (Constant), Curriculum Improvement

The results of the multiple regressions for “Use of Summative Assessment Data” are presented in Table 30. The factor “Instructional Improvement” was the only factor included in model. The model accounted for 9% of the variance in “Use of Summative

Data” (Adjusted $R^2 = .092$, $F = 13.394$, $p = .000$). “Instructional Improvement” was the influential variable in the first model with a standardized beta coefficient (β) of .315 ($p = .000$). Two variables, “percent of students receiving free or reduced price lunch” ($p = .260$) and “Curriculum Improvement” ($p = .745$), did not achieve significance in the model.

Table 30

Linear Regression: Instructional Leadership Factors with Use of Summative Assessment Data (Scale 6 – TAPS-ML)

Model Summary					
Model	R	R^2	R^2 Change	Adjusted R^2	Std. Error of the Est.
1	.315	.099	.099	.092	.55756
ANOVA					
Model	SS	df	MS	F	Sig. F
Regression	4.164	1	4.164	13.394	.000
Residual	37.926	122	.311		
Total	42.090	123			
Coefficients					
Variables	B	Std. Error	β	t	Sig. t
(Constant)	2.619	.421		6.219	.000
APE Factor 8 Instructional Improvement	.310	.085	.315	3.660	.000

a. Predictors: (Constant), Instructional Improvement

The results of the multiple regressions for “Overall Collective Assessment Practices” are presented in Table 31. The factor “Instructional Improvement” was the only factor included in model. The model accounted for 12% of the variance in “Overall Collective Assessment Practices” (Adjusted $R^2 = .117$, $F = 17.272$, $p = .000$). “Instructional Improvement” was the influential variable in the first model with a

standardized beta coefficient (β) of .352 ($p = .000$). Two variables, “percent of students receiving free or reduced price lunch” ($p = .132$) and “Curriculum Improvement” ($p = .780$), did not achieve significance in the model.

Table 31

Linear Regression: Educational Improvement Factors with Overall Collective Assessment Practices

Model Summary					
Model	<i>R</i>	<i>R</i> ²	<i>R</i> ² Change	Adjusted <i>R</i> ²	Std. Error of the Est.
1	.352	.124	.124	.117	.50727
ANOVA					
Model	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	Sig. <i>F</i>
Regression	4.444	1	4.444	17.272	.000
Residual	31.394	122	.257		
Total	35.838	123			
Coefficients					
Variables	<i>B</i>	Std. Error	β	<i>t</i>	Sig. <i>t</i>
(Constant)	3.201	.384		8.367	.000
APE Factor 8 Instructional Improvement	.320	.077	.352	4.156	.000

a. Predictors: (Constant), Instructional Improvement

Hypothesis Three

The third hypothesis in this study was: There are no significant predictive relationships for the scales and factors measuring assessment practices and the factors measuring principal instructional leadership and student achievement in mathematics and communication arts.

Multiple linear regression analysis was the primary analytical technique used to explain the variance in the linear models. Stepwise regression was the regression method conducted to determine the predictor (independent) variables with the best estimate, or

predictive power for the selected criterion (dependent) variables. Separate stepwise regressions were conducted for each criterion variable that at some point showed influence in the model. The linear regressions were first conducted for Communication Arts and Mathematics. Additional stepwise regressions were then conducted on the variables that were identified as significant and predictive, until all variables had been exhausted. For all analyses, the percent of students receiving free or reduced price lunch were entered into the stepwise regression. If the percent of students receiving free or reduced price lunch was significant, then that will be shown; if it is absent that means it was not significant in its ability to account for explained variance.

In total, two stepwise linear regressions were conducted. The statistics presented from the stepwise regressions include the model summary data, specifically, the number of models produced, R , R^2 , R^2 Change, Adjusted R^2 , and Standard Error of the Estimate. The statistics presented for the individual models include sum of squares (SS) (regression, residual, and total), degrees of freedom, mean square (MS), F value and significance of F . The variable statistics, shown via the coefficients from the models, include the beta value, the standard error of the beta, the standardized beta coefficient (β), the t value, and significance of each t value.

The results of the regression equations estimated for Communication Arts are presented in Table 32. The independent factors entered were “percent of students receiving free or reduced price lunch”, the four factors from the Middle Level Teachers’ Assessment Practices Survey: (a) Preparation for Assessment, (b) Application of Formative Assessment, (c) Impact of Summative Assessment and (d) Collaboration through Assessment and the two factors from the Educational Improvement section of the

Audit of Principal Effectiveness survey: (a) Instructional Improvement and (b) Curriculum Improvement. The first model accounted for 49% of the variance in Communication Arts scores (Adjusted $R^2 = .491$, $F = 119.502$, $p = .000$). The second model accounted for around 50% of the variation in Communication Arts scores with an R^2 change of .017 (Adjusted $R^2 = .504$, $F = 63.460$, $p = .000$). The “percent of students receiving free or reduced price lunch” was the influential variable in the first model. Examination of the standardized beta coefficients (β) explains the influence of two of the three independent variables entered into the second model. The variable, “percent of students receiving free or reduced price lunch” ($\beta = -.703$, $p = .000$) from the first model entered into the second explaining the majority of the variance (R^2 change = .017), while “Instructional Improvement” ($\beta = .131$, $p = .042$) also significantly contributed to the overall Adjusted R^2 . The variables, “Curriculum Improvement” ($p = .691$), “Preparation for Formative Assessment” ($p = .507$), “Application of Formative Assessment” ($p = .593$), “Impact of Summative Assessment” ($p = .877$), and “Collaboration through Assessment” ($p = .737$), did not explain variance in Communication Arts scores.

Table 32

Linear Regression: Instructional Leadership Factors and Middle Level Teachers Assessment Practices with Student Achievement in Communication Arts

Model Summary					
Model	<i>R</i>	<i>R</i> ²	<i>R</i> ² Change	Adjusted <i>R</i> ²	Std. Error of the Est.
1	.703 ^a	.495	.495	.491	8.06067
2	.715 ^b	.512	.017	.504	7.95565

a. Predictors: (Constant), Free/Reduced Lunch

b. Predictors: (Constant), Free/Reduced Lunch, Instructional Improvement

ANOVA					
Model	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	Sig. <i>F</i>
Regression	7764.577	1	7764.577	119.502	.000
Residual	7926.880	122	64.974		
Total	15691.457	123			
Regression	8033.083	2	4016.542	63.460	.000
Residual	7658.374	121	63.292		
Total	15691.457	123			

Coefficients					
Variables	<i>B</i>	Std. Error	β	<i>t</i>	Sig. <i>t</i>
(Constant)	63.534	1.769		35.921	.000
Free/Reduced Lunch	-.430	.039	-.703	-10.932	.010
(Constant)	51.306	6.188		8.291	.000
Free/Reduced Lunch	-.431	.039	-.706	-11.144	.000
APE Factor 8 Instructional Improvement	2.189	1.208	.131	2.060	.042

The results of the regression equations estimated for Mathematics are presented in Table 33. The independent factors entered were “percent of students receiving free or reduced price lunch”, the four factors from the Middle Level Teachers’ Assessment Practices Survey: (a) Preparation for Assessment, (b) Application of Formative Assessment, (c) Impact of Summative Assessment and (d) Collaboration through Assessment and the two factors from the Educational Improvement section of the Audit

of Principal Effectiveness survey: (a) Instructional Improvement and (b) Curriculum Improvement. The first model accounted for 50% of the variance in Mathematics scores (Adjusted $R^2 = .508$, $F = 127.841$, $p = .000$). The second model accounted for around 54% of the variation in Mathematics scores (Adjusted $R^2 = .535$, $F = 71.825$, $p = .000$). The “percent of students receiving free or reduced price lunch” was the influential variable in the first model. Examination of the standardized beta coefficients (β) explains the influence of two of the three independent variables entered into the second model. The variable, “percent of students receiving free or reduced price lunch” ($\beta = .715$, $p = .000$) from the first model entered into the second explaining the majority of the variance (Adjusted $r^2 = .535$), while “Curriculum Improvement” ($\beta = .176$, $p = .000$) also significantly contributed to the overall Adjusted R^2 . The variables, “Instructional Improvement” ($p = .751$), “Preparation for Formative Assessment” ($p = .920$), “Application of Formative Assessment” ($p = .231$), “Impact of Summative Assessment” ($p = .738$), and “Collaboration through Assessment” ($p = .925$), did not explain variance in Mathematics scores.

Table 33

Linear Regression: Instructional Leadership Factors and Middle Level Teachers Assessment Practices with Student Achievement in Mathematics

Model Summary					
Model	<i>R</i>	<i>R</i> ²	Adjusted <i>r</i> ²	Std. Error of the Est.	
1	.715 ^a	.512	.508	8.99940	
2	.737 ^b	.543	.535	8.74398	
a. Predictors: (Constant), Free/Reduced Lunch					
b. Predictors: (Constant), Free/Reduced Lunch, Curriculum Improvement					

ANOVA					
Model	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	Sig. <i>F</i>
Regression	10353.720	1	10353.720	127.841	.000
Residual	9880.673	122	80.989		
Total	20234.392	123			
Regression	10983.072	2	5491.536	71.825	.000
Residual	9251.321	121	76.457		
Total	20234.392	123			

Coefficients					
Variables	<i>B</i>	Std. Error	β	<i>t</i>	Sig. <i>t</i>
(Constant)	3.516	.284		12.367	.000
APE Factor 9 Curricular Improvement	.149	.057	.229	2.601	.010

Hypothesis Four

The fourth hypothesis tested in this study was: There are no significant differences for the scales and factors measuring assessment practices and the factors measuring principal instructional leadership between schools with high levels of achievement and schools with low levels of achievement in mathematics and communication arts.

A general linear model was developed to test for significant differences in the factors and scales of the Middle Level Teachers' Assessment Practices Survey and Educational Improvement factors from the Audit of Principal Effectiveness Survey when

the schools were grouped into quartiles according to student achievement as measured by the Communication Arts and Mathematics scores from the MAP. Analysis of variance determined the F value along with an F significance level. The means of factors and scales from the TAPS-ML and the factors from the APE from each quartile of student achievement are presented. The results of the Tukey post hoc test are presented followed by the mean differences and significance level.

Table 34 contains the results of the ANOVA with the factors from the TAPS-ML as the dependent variable with schools divided into quartiles as determined by the scores from the communication arts portion of the MAP. The post hoc Tukey HSD data are also presented. “Factor 2 – Application of Formative Assessment” showed significance at an alpha level less than .05 ($F = 3.163, p = .027$). Mean differences between quartile one ($M = 4.0606$) and quartile two ($M = 3.7826$) were significant ($p = .050$). Mean differences between quartile one ($M = 4.0606$) and quartile three ($M = 3.716$) were also significant ($p = .048$).

Table 34

*Test of Differences: Teachers' Assessment Practices Factors and Communication Arts
(Quartiles 4 High – 1 Low)*

Factor Items	<i>F</i>	<i>F Sig.</i>	MAP CA Quartile	Mean	Tukey HSD		Mean Differences	Sig.
TAPS-ML Factor 1 – Preparation for Assessment	2.059	.109	1	4.7165	1	2	.30355	.084
						3	.22613	.287
						4	.17484	.516
			2	4.4129	2	1	-.30355	.084
						3	-.07742	.929
						4	-.12871	.742
			3	4.4903	3	1	-.22613	.287
						2	.07742	.929
						4	-.05129	.978
			4	4.5416	4	1	-.17484	.516
						2	.12871	.742
						3	.05129	.978
TAPS-ML Factor 2 – Application of Formative Assessment	3.163*	.027	1	4.0606	1	2	.27806*	.050
						3	.27903*	.048
						4	.13194	.604
			2	3.7826	2	1	-.27806*	.050
						3	.00097	1.00
						4	-.14613	.520
			3	3.7816	3	1	-.27903*	.048
						2	-.00097	1.00
						4	-.14710	.514
			4	3.9287	4	1	-.13194	.604
						2	.14613	.520
						3	.14710	.514
TAPS-ML Factor 3 – Impact of Summative Assessment	.746	.527	1	4.3735	1	2	.10677	.910
						3	.23968	.446
						4	.10387	.917
			2	4.2668	2	1	-.10677	.910

						3	.13290	.842
						4	-.00290	1.00
			3	4.1339	3	1	-.23968	.446
						2	-.13290	.842
						4	-.13581	.833
			4	4.2697	4	1	-.10387	.917
						2	.00290	1.00
						3	.13581	.833
TAPS-ML	.923	.432	1	4.0084	1	2	-.16323	.848
Factor 4 –						3	.08774	.972
Collaboration						4	-.20161	.747
through								
Assessment			2	4.1716	2	1	.16323	.848
						3	.25097	.597
						4	-.03839	.998
			3	3.9206	3	1	-.08774	.972
						2	-.25097	.597
						4	-.28935	.476
			4	4.2100	4	1	.20161	.747
						2	.03839	.998
						3	.28935	.476

* $p < .05$, ** $p < .01$

Table 35 contains the results of the ANOVA with the factors from the TAPS-ML as the dependent variable with schools divided into quartiles as determined by the scores from the mathematics portion of the MAP. No significant differences were found for the assessment factor and mathematics achievement by quartile.

Table 35

*Test of Differences: Teachers' Assessment Practices Factors and Mathematics
(Quartile 4 High – 1 Low)*

Factor Items	<i>F</i>	<i>F Sig.</i>	MAP MA Quartile	Mean	Tukey HSD	Mean Differences	Sig.
TAPS-ML Factor 1 – Preparation for Assessment	.938	.425	1	4.5700	1	2 .03032	.995
						3 .14903	.654
						4 -.06065	.965
			2	4.5397	2	1 -.03032	.995
						3 .11871	.793
						4 -.09097	.894
			3	4.4210	3	1 -.14903	.654
						2 -.11871	.793
						4 -.20968	.366
			4	4.6306	4	1 .06065	.965
						2 .09097	.894
						3 .20968	.366
TAPS-ML Factor 2 – Application of Formative Assessment	.131	.942	1	3.8919	1	2 -.01387	.999
						3 .04452	.978
						4 -.01645	.999
			2	3.9058	2	1 .01387	.999
						3 .05839	.952
						4 -.00258	1.00
			3	3.8474	3	1 -.04452	.978
						2 -.05839	.952
						4 -.06097	.946
			4	3.9084	4	1 .01645	.999
						2 .00258	1.00
						3 .06097	.946
TAPS-ML Factor 3 – Impact of Summative Assessment	.241	.868	1	4.2248	1	2 -.11516	.892
						3 -.03677	.996
						4 .00742	1.00
			2	4.3400	2	1 .11516	.892

						3	.07839	.962
						4	.12258	.873
		3		4.2616	3	1	.03677	.996
						2	-.07839	.962
						4	.04419	.993
		4		4.2174	4	1	-.00742	1.00
						2	-.12258	.873
						3	-.04419	.993
TAPS-ML	.466	.707	1	4.0558	1	2	.09161	.969
Factor 4 –						3	-.03484	.998
Collaboration						4	-.14419	.891
through								
Assessment			2	3.9642	2	1	-.09161	.969
						3	-.12645	.923
						4	-.23581	.648
		3		4.0906	3	1	.03484	.998
						2	.12645	.923
						4	-.10935	.949
		4		4.2000	4	1	.14419	.891
						2	.23581	.648
						3	.10935	.949

* $p < .05$, ** $p < .01$

Table 36 contains the results of the ANOVA with the scales from the TAPS-ML as the dependent variable with schools divided into quartiles as determined by the scores from the communication arts portion of the MAP. The analyses for two scales produced significance.

“Scale 4 – Teachers Feedback to Students and Parents” showed significance at an alpha level less than .01 ($F = 7.412$, $p = .000$). Mean differences between quartile one ($M = 4.1019$) and quartile two ($M = 3.7816$) were significant ($p = .002$). Mean differences between quartile one ($M = 4.1019$) and quartile three ($M = 3.8310$) were also significant

($p = .012$). Quartile four ($M = 4.0845$) was different from quartile two ($M = 3.7816$) ($p = .004$) and from quartile three ($M = 3.8310$) ($p = .021$).

“Scale 5 – Teachers Develops Self and Peer Assessments” showed significance at an alpha level less than .01 ($F = 4.914$, $p = .003$). Mean differences between quartile one ($M = 3.9245$) and quartile three ($M = 3.6294$) were significant ($p = .048$). Mean differences between quartile three ($M = 3.6294$) and quartile four ($M = 4.0113$) were also significant ($p = .005$).

Table 36

Test of Differences: Teachers' Assessment Practices Scales and Communication Arts (Quartile 4 High – 1 Low)

Scale Items	<i>F</i>	<i>F Sig.</i>	MAP CA Quartile	Mean	Tukey HSD	Mean Differences	Sig.	
TAPS-ML Scale 1 – Instructional Preparation for Assessment	2.321	.079	1	4.6987	1	2	.13935	.485
						3	.16258	.347
						4	-.05323	.948
			2	4.5594	2	1	-.13935	.485
						3	.02323	.995
						4	-.19258	.205
			3	4.5361	3	1	-.16258	.347
						2	-.02323	.995
						4	-.21581	.127
			4	4.7519	4	1	.05323	.948
						2	.19258	.205
						3	.21581	.127
TAPS-ML Scale 2 – Teacher Assessment during Instruction	1.523	.212	1	4.8265	1	2	.17484	.225
						3	.12323	.531
						4	.03806	.975
			2	4.6516	2	1	-.17484	.225
						3	-.05161	.942
						4	-.13677	.440

			3	4.7032	3	1	-.12323	.531
						2	.05161	.942
						4	-.08516	.786
			4	4.7884	4	1	-.03806	.975
						2	.13677	.440
						3	.08516	.786
TAPS-ML Scale	1.037	.379	1	3.9255	1	2	.06323	.876
3 – Teacher						3	.14806	.300
Adaptation of						4	.07548	.807
Instruction per								
Assessment			2	3.8623	2	1	-.06323	.876
						3	.08484	.746
						4	.01226	.999
			3	3.7774	3	1	-.14806	.300
						2	-.08484	.746
						4	-.07258	.825
			4	3.8500	4	1	-.07548	.807
						2	-.01226	.999
						3	.07258	.825
TAPS-ML Scale	7.412**	.000	1	4.1019	1	2	.32032*	.002
4 – Teachers						3	.27097*	.012
Feedback to						4	.01742	.997
Students and								
Parents			2	3.7816	2	1	-.32032*	.002
						3	-.04935	.941
						4	-.30290*	.004
			3	3.8310	3	1	-.27097*	.012
						2	.04935	.941
						4	-.25355*	.021
			4	4.0845	4	1	-.01742	.997
						2	.30290*	.004
						3	.25355*	.021
TAPS-ML Scale	4.914**	.003	1	3.9245	1	2	.20387	.275
5 – Teacher						3	.29516*	.048
Develops Self						4	-.08677	.868
and Peer								

Assessment				2	3.7206	2	1	-.20387	.275						
							3	.09129	.850						
							4	-.29065	.054						
							3	3.6294	3	1	-.29516*	.048			
										2	-.09129	.850			
										4	-.38194*	.005			
							4	4.0113	4	1	.08677	.868			
										2	.29065	.054			
										3	.38194*	.005			
							TAPS-ML Scale 6 – Teacher Use of Summative Data	1.267	.289	1	4.1755	1	2	-.08355	.942
													3	.19355	.560
													4	-.00484	1.00
2	4.2590	2	1	.09355	.942										
			3	.27710	.246										
			4	.07871	.951										
3	3.9819	3	1	-.19355	.560										
			2	-.27710	.246										
			4	-.19839	.540										
4	4.1803	4	1	.00484	1.00										
			2	-.07871	.951										
			3	.19839	.540										

* $p<.05$, ** $p<.01$

* $p < .05$, ** $p < .01$

Table 37 contains the results of the ANOVA with the scales from the TAPS-ML as the dependent variable with schools divided into quartiles as determined by the scores from the mathematics portion of the MAP. Two scales had significant differences.

“Scale 4 – Teachers Feedback to Students and Parents” had an alpha level of less than .05 ($F = 2.840$, $p = .041$). Mean differences between quartile three ($M = 3.8258$) and quartile four ($M = 4.0774$) were significant ($p = .034$).

“Scale 5 – Teacher Develops Self and Peer Assessments” showed significance at an alpha level less than .05 ($F = 3.748, p = .013$). Mean differences between quartile three ($M = 3.6945$) and quartile four ($M = 4.0490$) were significant ($p = .013$).

Table 37

Test of Differences: Teachers' Assessment Practices Scales and Mathematics (Quartile 4 High – 1 Low)

Scale Items	<i>F</i>	<i>F Sig.</i>	MAP MA Quartile	Mean	Tukey HSD		Mean Differences	Sig.
TAPS-ML Scale 1 – Instructional Preparation for Assessment	2.169	.095	1	4.6755	1	2	.08419	.825
						3	.15323	.402
						4	-.08161	.838
			2	4.5913	2	1	-.08419	.825
						3	.06903	.895
						4	-.16581	.331
			3	4.5223	3	1	-.15323	.402
						2	-.06903	.895
						4	-.23484	.083
			4	4.7571	4	1	.08161	.838
						2	.16581	.331
						3	.23484	.083
TAPS-ML Scale 2 – Teacher Assessment during Instruction	.486	.693	1	4.7471	1	2	.00968	1.00
						3	.06000	.915
						4	-.05097	.946
			2	4.7374	2	1	-.00968	1.00
						3	.05032	.948
						4	-.06065	.913
			3	4.6871	3	1	-.06000	.915
						2	-.05032	.948
						4	-.11097	.626
			4	4.7981	4	1	.05097	.946
						2	.06065	.913
						3	.11097	.626

TAPS-ML Scale 3 – Teacher Adaptation of Instruction per Assessment	.697	.556	1	3.8700	1	2	-.02355	.992
						3	.08935	.717
						4	-.00097	1.00
			2	3.8935	2	1	.02355	.992
						3	.11290	.543
						4	.02258	.993
			3	3.7806	3	1	-.08935	.717
						2	-.11290	.543
						4	-.09032	.710
			4	3.8710	4	1	.00097	1.00
						2	-.02258	.993
						3	.09032	.710
TAPS-ML Scale 4 – Teachers Feedback to Students and Parents	2.840*	.041	1	3.9042	1	2	-.08742	.774
						3	.07839	.826
						4	-.17323	.235
			2	3.9916	2	1	.08742	.774
						3	.16581	.271
						4	-.08581	.783
			3	3.8258	3	1	-.07839	.826
						2	-.16581	.271
						4	-.25161*	.034
			4	4.0774	4	1	.17323	.235
						2	.08581	.783
						3	.25161*	.034
TAPS-ML Scale 5 – Teacher Develops Self and Peer Assessment	3.748*	.013	1	3.7858	1	2	.02935	.994
						3	.09129	.855
						4	-.26323	.103
			2	3.7565	2	1	-.02935	.994
						3	.06194	.949
						4	-.29258	.056
			3	3.6945	3	1	-.09129	.855
						2	-.06194	.949
						4	-.35452*	.013

TAPS-ML Scale 6 – Teacher Use of Summative Data	.179	.910	4	4.0490	4	1	.26323	.103
						2	.29258	.056
						3	.35452*	.013
			1	4.1374	1	2	.01097	1.00
						3	.02000	.999
						4	-.07806	.954
			2	4.1265	2	1	-.01097	1.00
						3	.00903	1.00
						4	-.08903	.934
			3	3.1174	3	1	-.0200	.999
						2	-.00903	1.00
						4	-.09806	.914
			4	4.2155	4	1	.07806	.954
						2	.08903	.934
						3	.09806	.914

* $p < .05$, ** $p < .01$

Table 38 contains the results of the ANOVA with the factors from the Educational Improvement portion of the Audit of Principal Effectiveness Scale as the dependent variable with schools divided into quartiles as determined by the scores from the communication arts portion of the MAP. “Factor 1 –Instructional Improvement” ($F = 2.081, p = .106$) and “Factor 2 – Curriculum Improvement” ($F = 1.881, p = .136$) were not significant for any quartiles.

Table 38

Test of Differences: Educational Improvement Factors from Audit of Principal Effectiveness Survey and Communication Arts
(Quartile 4 High – 1 Low)

Items	<i>F</i>	<i>F Sig.</i>	MAP CA Quartile	Mean	Tukey HSD		Mean Differences	Sig.
APE Factor 1 – Instructional Improvement	2.081	.106	1	4.8777	1	2	.10581	.892
						3	-.24871	.343
						4	-.10871	.884
			2	4.7719	2	1	-.10581	.891
						3	-.35452	.086
						4	-.21452	.476
			3	5.1265	3	1	.24871	.343
						2	.35452	.086
						4	.14000	.783
			4	4.9865	4	1	.10871	.881
						2	.21452	.476
						3	-.14000	.783
APE Factor 2 – Curriculum Improvement	1.881	.136	1	4.8748	1	2	.09613	.922
						3	-.21097	.513
						4	-.18613	.616
			2	4.7787	2	1	-.09613	.922
						3	-.30710	.189
						4	-.28226	.256
			3	5.0858	3	1	.21097	.513
						2	.30710	.189
						4	.02484	.998
			4	5.0610	4	1	.18613	.616
						2	.28226	.256
						3	-.02484	.998

* $p < .05$, ** $p < .01$

Table 39 contains the results of the post hoc ANOVA with the factors from the Educational Improvement portion of the Audit of Principal Effectiveness Scale as the dependent variable with schools divided into quartiles as determined by the scores from the mathematics portion of the MAP. “Factor 1 –Instructional Improvement” ($F = 1.993$, $p = .119$) and “Factor 2 – Curriculum Improvement” ($F = 2.123$, $p = .101$) were not significant for any quartiles.

Table 39

Test of Differences: Educational Improvement Factors from Audit of Principal Effectiveness Survey and Mathematics (Quartile 4 High – 1 Low)

Items	<i>F</i>	<i>F Sig.</i>	MAP MA Quartile	Mean	Tukey HSD		Mean Differences	Sig.
APE Factor 1 – Instructional Improvement	1.993	.119	1	4.7242	1	2	-.31226	.160
						3	-.23839	.382
						4	-.31516	.154
			2	5.0365	2	1	.31226	.160
						3	.07387	.960
						4	-.00290	1.00
			3	4.9626	3	1	.23839	.382
						2	-.07387	.960
						4	-.07677	.955
			4	5.0394	4	1	.31516	.154
						2	.00290	1.00
						3	.07677	.955
APE Factor 2 – Curriculum Improvement	2.123	.101	1	4.7465	1	2	-.22290	.462
						3	-.20935	.517
						4	-.38226	.063
			2	4.9694	2	1	.22290	.462
						3	.01355	1.00
						4	-.15935	.722
			3	4.9558	3	1	.20935	.517
						2	-.01355	1.00
						4	-.17290	.668
			4	5.1287	4	1	.38226	.063
						2	.15935	.722
						3	.17290	.668

* $p < .05$, ** $p < .01$

Table 40 provides the quartile mean for students receiving free and reduced price lunch. Free and reduced lunch is used to measure the socioeconomic status for a school.

Quartile one had a mean of 57 percent of students receiving free and reduced price lunch. Quartile two's average percentage was 47 percent, followed by the third quartile with 36 percent. The highest achieving of the MAP scores, quartile four, had an average of 24 percent receiving free and reduced price lunch.

Table 40

Average percentage of students receiving free or reduced price lunch by achievement quartile

Quartile 1	Quartile 2	Quartile 3	Quartile 4
57	47	37	24

Supplemental Data

Four open-ended questions were asked at the end of the Middle Level Teachers' Assessment Practices Survey (TAPS-ML). The four questions were: (a) What assessment strategies do you voluntarily use because you believe they are truly beneficial in raising student achievement on the MAP test? (b) What assessment strategies are mandated by your district or principal with the intention of raising student achievement on the MAP test? (c) Which forms of assessment do you believe are the most beneficial in raising student achievement on the MAP test? Why? and (d) Which forms of assessment do you believe are the least beneficial in raising student achievement on the MAP test? Why?

Responses from each question were copied and divided into quartiles based on mean student achievement scores in Communication Arts as measured by the MAP and then divided again into quartiles based on mean student achievement scores in Mathematics. Schools that were placed in the same quartile for both Communication Arts

and Mathematics were studied. The responses were analyzed through the lens of finding similarities and differences among the responses by quartile.

The first question, “What assessment strategies do you voluntarily use because you believe they are truly beneficial in raising student achievement on the MAP test?” revealed a common answer throughout all of the quartiles. The most consistent reply described the use of constructed response questions. An example reply was, “I use constructed response questions which are open ended questions that ask about real world situations” (Teacher from School #257).

Teachers’ responses from school placed in the highest achievement quartile in both Communication Arts and Mathematics, quartile four, revealed a common response to the first question not present in any of the other quartiles. Quartile four teachers placed an emphasis on writing. As one teacher attested, “We use writing in one capacity or another everyday.” Quartile two and three were similar in that individual teachers listed many different types of assessments considered beneficial in improving student achievement. Unit tests based on the Grade Level Expectations (GLE), performance events, weekly quizzes, and objective tests were examples of items included on the lists generated by the teachers in these quartiles. Quartile one, the lowest quartile, placed a majority of interest on administering “MAP-like” tests and the released items from the Missouri Department of Elementary and Secondary Education. Another common response from this group that didn’t appear in any other group was the use of self-assessments and self-reflection activities.

The second question was “What assessment strategies are mandated by your district or principal with the intention of raising student achievement on the MAP test?”

Common throughout all of the quartiles was the district implementation of common assessments or benchmark tests. Other similar examples included, “a summative test for all kids in all subjects at the end of the first semester” (Teacher from School #240), “tests that are based on the Grade Level Expectations” (Teacher from School #14) and “chapter assessments related to the GLEs” (Teacher from School #222).

The third question was “Based on your responses from the previous two questions, which forms of assessment are the most beneficial in raising student achievement on the MAP test? Why?” Constructed response assessments were a common selection throughout all four quartiles, however, the reasoning for the administration of these assessments varied among the quartiles. Teachers’ answers from the top two quartiles to why constructed response assessments were beneficial centered on the thinking process associated with constructed response assessments. Teachers expressed that students need to “form their own thoughts” (Teacher from School #56) and gain “exposure to the process side of answering questions” (Teacher from School #154) as a way to “internalize information” (Teacher from School #287). Conversely, teachers’ responses from schools in the bottom two quartiles focused their reasoning on the similarity to the MAP and how the replication will help students score well on the test. Typical answers indicated that constructed response assessments “allow the students to sample the MAP test before it arrives” (Teacher from School #134), constructed response assessments “looked like the MAP test” (Teacher from School #165) and allow students to become “comfortable with the MAP format” (Teacher from School #58) “so you get a high score” (Teacher from School #38).

The final question was “Which forms of assessment do you believe are the least beneficial in raising student achievement on the MAP test? Why?” Common through all of the quartiles was a belief that rote memorization assessments were not beneficial to increasing student achievement. Examples cited were multiple choice tests, true/false tests, fill in the blank questions, matching and yes/no questions with no probing question for students to explain their thinking. The top two quartiles focused their reasoning on eliminating teaching practices that were not sustainable and beneficial to the student for lasting learning. Suggestions included, “not teaching to the MAP test, but incorporate testing skills through daily teaching” (Teacher from School #326). Another suggestion indicated that teachers need to match assessments with what was taught during instructional time because “kids will be more likely to retain the information” (Teacher from School #222).

Teachers from the lower two quartiles felt that assessments such as oral responses, performance events and PowerPoint instruction were not beneficial because they were not similar to questions on the MAP test. Another criticism listed was that “they are least beneficial because they count for little points” (Teacher from School #40) and “they are hard to grade” (Teacher from School # 16). Other comments common to the lower two quartiles focused on the use of providing proper feedback to students. One teacher responded, “Students need more. A percentage doesn’t tell much” (Teacher from School #38). Another response common to the lower two quartiles was a feeling that all types of assessment help in increasing student achievement especially for students from diverse backgrounds.

Summary of Results

The results from each of the hypotheses are discussed in the following section. Two models were developed illustrating the findings.

Hypothesis Testing

Hypothesis one, there are no significant correlational relationships among the scales and factors measuring assessment practices, the factors measuring principal instructional leadership and student achievement as measured by the Mathematics and Communication Arts portions of the Missouri Assessment Program, was rejected. A summary of the findings from the partial correlations are presented in Table 41. The positive significant relationships are indicated with a “+,” and the negative significant relationships with a “-.” The last row of the table provides the total correlations.

Factors one through four correlated with the other. APE Factor 8 – Instructional Improvement correlated with TAPS-ML Factor 4 – Collaboration through Assessment, TAPS-ML Scale 6 – Teacher Use of Summative Assessment Data, Overall Teachers Assessment Practices, and Mathematics MAP scores. Factor 9 – Curricular Improvement correlated with TAPS-ML Factor 4 – Collaboration through Assessment, Scale 2 – Teacher Assessment during Instruction, Scale 6 – Teachers Use of Summative Data and Overall Teachers Assessment Practices. Scale 5 – Teacher Develops Self and Peer Assessments correlated with Mathematics MAP.

Table 41

Summary of Partial Correlation Findings

	F1	F2	F3	F4	S1	S2	S3	S4	S5	S6	Overall	Inst.	Cur.	Math	CA
F1		+	+	+											
F2	+		+	+											
F3	+	+		+											
F4	+	+	+									+	+		
S1															
S2													+		
S3															
S4															
S5														+	
S6												+	+		
Overall												+	+		
Inst.				+						+	+			+	
Cur.				+		+				+	+				
Math									+			+			
CA															
Total	3	3	3	5	0	1	0	0	1	2	2	4	4	2	0

F1 – Preparation for Assessment, F2 – Application of Formative Assessment, F3 – Impact of Summative Assessment, F4 – Collaboration through Assessment, S1 – Instructional Preparation for Assessment, S2 – Teacher Assessment during Instruction, S3 – Teacher Adaptation of Instruction, S4 – Teacher Feedback to Students and Parents, S5 – Teacher Develops Self and Peer Assessment, S6 – Teacher Use of Summative Data, Overall – Overall Teachers Assessment Practices, Ave. – Middle Level Teachers Assessment Practices Average, Inst. – APE Factor 8 – Instructional Improvement, Cur. – APE Factor 9 – Curriculum Improvement, Math – Mathematics MAP, CA – Communication Arts MAP

Hypothesis two, there are no significant predictive linear relationships for the factors measuring principal instructional leadership and the scales and factors measuring assessment practices, and Hypothesis three, there are no significant predictive linear relationships for the scales and factors measuring assessment practices and the factors measuring principal instructional leadership and student achievement in communication arts and mathematics, were rejected. A summary to the findings from the multiple regressions are presented in Table 42. The positive significant relationships are indicated with a “+”, and the negative significant relationship with a “-“. Only the variables that significantly explained a portion of the variance are represented. Students who receive free and reduced price lunches explained most of the variance in both mathematics and communication arts. APE Factor 8 – Instructional Improvement had positive influences on TAPS-ML Factor 4 – Collaboration through Assessment, TAPS-ML Scale 6 – Teachers Use of Summative Assessments, TAPS-ML Overall Teachers Assessment Practices and Communication Arts. APE Factor 9 – Curriculum Improvement had a positive relationship with TAPS-ML Scale 2 – Teacher Assessment during Instruction and Mathematics.

Table 42

Summary of the Influence of Independent on Dependent Variables from the Multiple Regression Findings

	F/R	F4	S2	S6	Overall	Inst.	Cur.	Math	CA
F/R								+	+
F4						+			
S2							+		
S6						+			
Overall						+			
Inst.		+		+	+				+
Cur.			+					+	
Math	+						+		
CA	+					+			

Hypothesis four, there are no significant differences for the scales and factors measuring assessment practices and the factors measuring principal instructional leadership among schools with high levels of achievement and schools with low levels of achievement in mathematics and communication arts, was rejected. Differences occurred in TAPS-ML Factor 2 –Application of Formative Assessment, TAPS-ML Scale 4 – Teacher Feedback to Students and Parents and TAPS-ML Scale 5 – Teacher Develops Self and Peer Assessment with both Mathematics and Communication Arts. There were no significant differences for Instructional Improvement or Curriculum Improvement.

Tables 43 depicts the significant differences in the factors and scales of the Middle Level Teachers' Assessment Practices Survey and Educational Improvement

factors from the Audit of Principal Effectiveness Survey when the schools were grouped into quartiles according to student achievement as measured by the Communication Arts and Mathematics scores from the MAP with quartile one including those schools with the mean average in the lower 25% of the total sample and quartile four including those schools in the upper 25% of the total sample.

Table 42

Differences among Student Achievement Quartiles for the TAPS-ML Factors

	F1	F2	F3	F4	S1	S2	S3	S4	S5	S6	APE 8	APE 9
CA1		>CA2 >CA3						>CA2 >CA3	>CA2 >CA3 >CA4			
CA2		<CA1						<CA4				
CA3		<CA1						<CA4	<CA4			
CA4												
MA1												
MA2												
MA3								<MA4	<MA4			
MA4												

Given the findings in tables 40-41, Figure 4 was developed to depict the exploratory findings of this study.

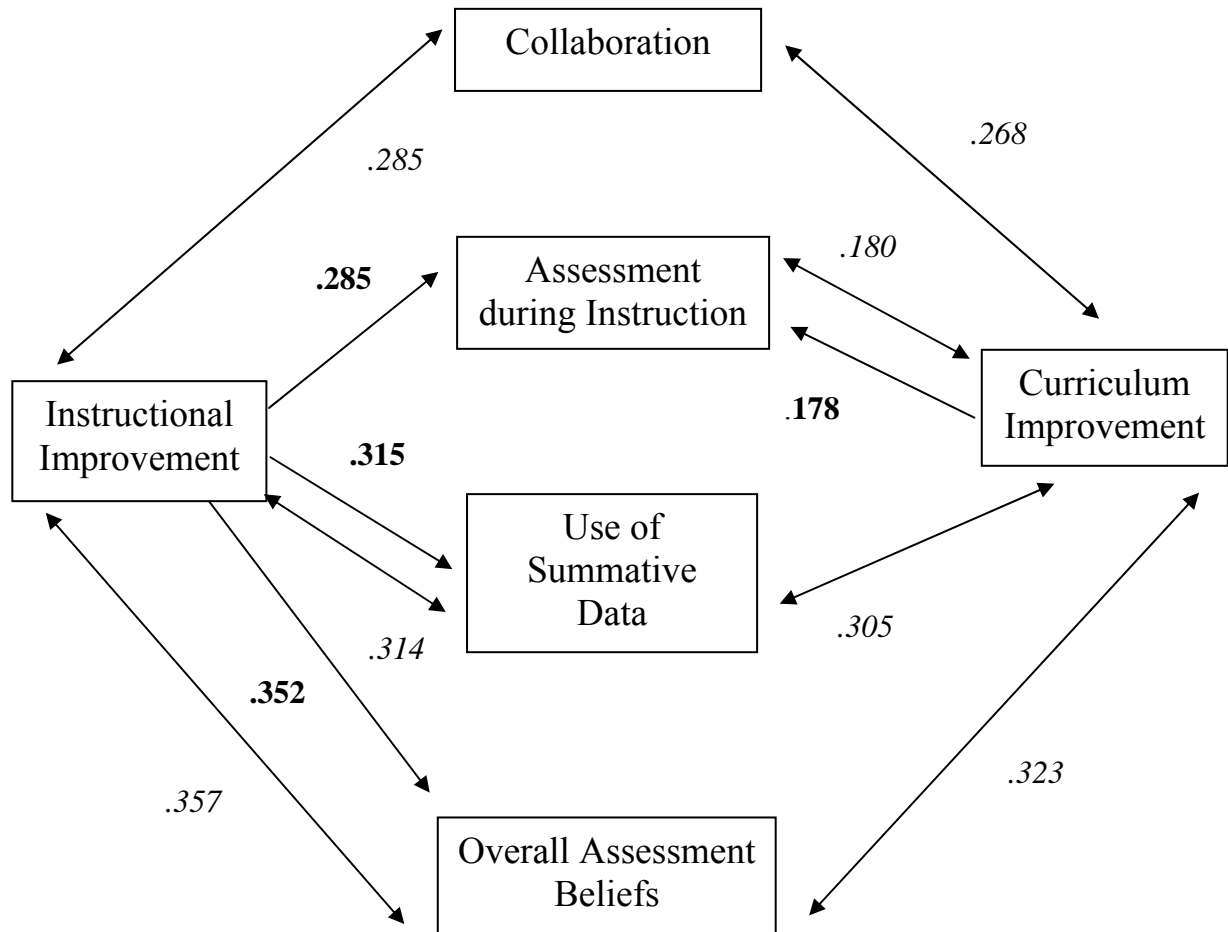


FIGURE 4. *Correlations and Regressions between Instructional Leadership factors, Instructional Improvement and Curriculum, and Middle Level Assessment Practices scales and factors, Collaboration through Assessment, Assessment during Instruction, Use of Summative Data, Overall Collective Assessment Practices*

- ↔ = Mutually Influential
(Partial Correlation in italics)
- = Directly Influenced
(Standardized Beta Coefficient, β , in bold)

Partial correlations for the TAPS-ML variables, Collaboration through Assessment, Use of Summative Assessment Data, Assessment during Instruction and Overall Collective Assessment Practices are presented in Table 44. All variables showed moderate to strong correlations with the other variables.

Table 44

Correlations for Collaboration through Assessment, Use of Summative Assessment Data, Assessment during Instruction and Overall Collective Assessment Practices controlling for Free/Reduced Lunch

	Collaboration through Assessment	Use of Summative Assessment Data	Teacher Assessment during Instruction	Overall Teacher Assessment
Collaboration through Assessment	1.00			
Use of Summative Assessment Data	.733** $p = .000$	1.00		
Teacher Assessment during Instruction	.252** $p = .005$.337** $p = .000$	1.00	
Overall Teacher Assessment	.457** $p = .000$.516** $p = .000$.298** $p = .001$	1.00

* $p < .05$, ** $p < .01$

The results of the multiple regressions for Overall Collective Assessment Practices are presented in Table 45. The factors measuring Assessment during Instruction

and Use of Summative Data were included in model. The model accounted for 28% of the variance in Overall Collective Assessment Practices (Adjusted $R^2 = .279$, $F = 24.829$, $p = .000$). Use of Summative Assessment Data was the influential variable in the model with a standardized beta coefficient (β) of .473 ($p = .000$).

Table 45

Linear Regression: Middle Level Teachers Assessment Practices - Collaboration through Assessment, Assessment during Instruction and Use of Summative Assessment with Overall Collective Assessment Practices

Model Summary					
Model	R	R^2	R^2 Change	Adjusted R^2	
1	.549	.301	.017	.284	
2	.539	.291	.012	.279	

ANOVA					
Model	SS	df	MS	F	Sig. F
Regression	10.428	2	5.214	24.829	.000
Residual	25.410	121	.210		
Total	35.838	123			

Coefficients					
Variables	B	Std. Error	β	t	Sig. t
(Constant)	1.963	.556		3.530	.001
Assessment during Instruction	.215	.122	.144	1.764	.048
Use of Summative Assessment	.436	.075	.473	5.800	.000

a. Predictors: (Constant), Assessment during Instruction, Use of Summative Data

A final figure was developed based on previous findings on the instructional leadership factors and the correlations and regressions in Tables 43 and 44. Further discussion for this model will occur in Chapter Five.

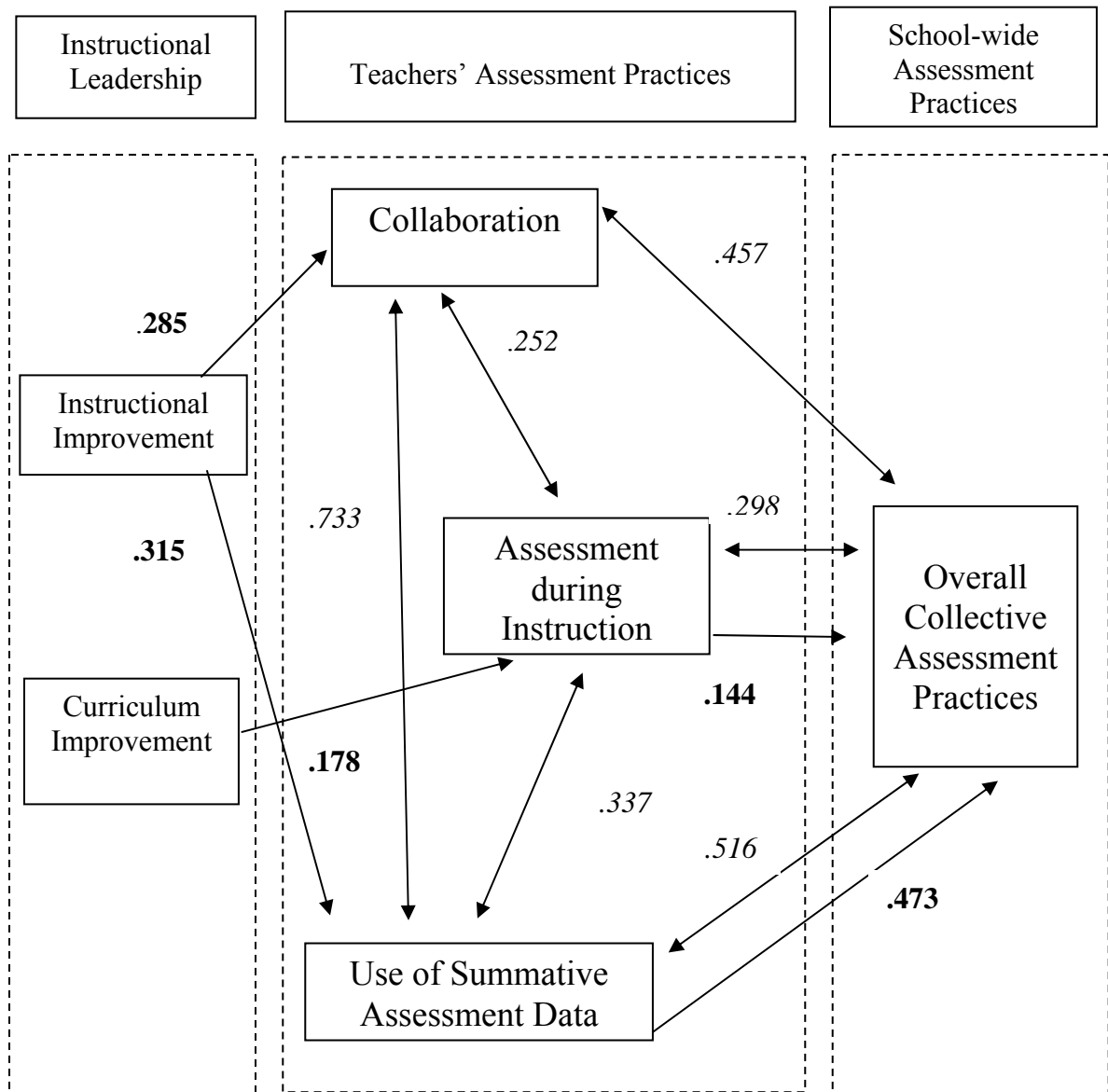


FIGURE 5. *Correlations and Regressions for Instructional Leadership factors and Assessment Practices on Overall Collective Assessment Practices*

- \longleftrightarrow = Mutually Influential
(Partial Correlation in italics)
- \longrightarrow = Directly Influenced
(Standardized Beta Coefficient, β , in bold)
- = Broad Components of the model

CHAPTER 5

DISCUSSION OF FINDINGS

Introduction

In recent years, assessments and accountability have become the focus for most educational improvement initiatives. The No Child Left Behind accountability systems have demanded that educators concentrate on student achievement as measured by annual standardized tests. External pressures to increase student achievement have caused principals and teachers to work to increase scores on the annual tests. The mistake that is being made throughout schools is that the annual standardized test is not the only assessment that is taken throughout the year and may not be the best source of information about student achievement for teachers (Stiggins, 2004). The classroom assessments that teachers administer everyday in classrooms provide ample information about student learning too often overlooked by educators (Stiggins).

Overview of Study

The purpose of this study was to analyze the relationships among teachers' assessment practices, instructional leadership and student achievement in the middle level schools in Missouri. The method of analysis was primarily quantitative, with survey data being used to determine, (a) what is descriptive of the demographics of Missouri middle level schools, (b) what is descriptive of the assessment practices of teachers in Missouri middle level schools, (c) what is descriptive of the instructional leadership practices of the school leader, and (d) if any relationships exist between the teachers' assessment practices, the principal as instructional leader and student achievement. Supplementing

the quantitative data were data obtained through written responses to open-ended questions.

Two surveys were used in this study. The Middle Level Teacher Assessment Practices Survey (TAPS-ML), developed by the researcher and the Middle Level Center Director, was used to collect data from middle level teachers about their self-perceptions of their classroom assessment practices. The instrument was developed from the review of the literature on classroom assessment practices. The survey was organized around five scales: (a) instructional strategies, (b) teacher adaptation of instruction per assessment data, (c) teacher feedback per assessment data, (d) teacher development of peer and self-assessment strategies, and (e) the formative use of summative assessments. The four factors that were analyzed and developed were: (a) preparation for assessment, (b) application of formative assessment, (c) impact of summative assessment, and (d) collaboration through assessment.

Principal instructional leadership was measured using the factors of the Educational Domain portion of the Audit of Principal Effectiveness (Valentine & Bowman, 1988). The measure was completed by teachers and was based on teachers' perceptions of the principal's instructional leadership behavior. The two factors that measured instructional leadership were: (a) Instructional Improvement and (b) Curricular Improvement. Student achievement data, as measured by the Missouri Assessment Program (MAP) tests, were obtained from the Missouri Department of Elementary and Secondary Education website for communication arts and mathematics. The achievement data for this study were from the state assessment administered in the spring of 2006.

In the fall of 2006, principals from Missouri's 371 middle level schools were contacted and asked to participate in a statewide study of middle level teachers' assessment practices and principals' instructional leadership qualities. Of the 371 schools, 336 fit the criteria for this study of having two or more grades and at least one of those being grade seven or eight. All of the 336 school principals were sent reminders via email in the winter of 2007.

Research Questions

To accomplish the purpose of this study, the following research questions were developed:

1. What are the types and levels of assessment practices used in Missouri's middle level schools according to teachers' perceptions?
2. What are the levels of instructional leadership for principals in Missouri's middle level schools?
3. Are there relationships between assessment practices and principal instructional leadership in Missouri's middle level schools?
4. Are there relationships between the types and levels of assessment practices and the levels of principal instructional leadership on student achievement in communication arts and mathematics according to teachers' perceptions?
5. Are there differences in the levels of principal instructional leadership and the types and levels of assessment practices between schools with high levels of achievement and schools with low levels of achievement in communication arts and mathematics?

Hypotheses

The following hypotheses were tested in this study:

H_{01} : There are no significant correlational relationships among the items, scales, and factors measuring assessment practices, the factors measuring principal instructional leadership and student achievement as measured by the Communication Arts and Mathematics portions of the Missouri Assessment Program.

H_{02} : There are no significant predictive linear relationships for the factors measuring principal instructional leadership, and the scales and factors measuring assessment practices.

H_{03} : There are no predictive linear relationships among the scales and factors measuring assessment practices and the factors measuring principal instructional leadership on student achievement in communication arts and mathematics.

H_{04} : There are no significant differences for the scales and factors measuring assessment practices and the factors measuring principal instructional leadership between schools with high levels of achievement and schools with low levels of achievement in communication arts and mathematics.

Summary of Findings

The descriptive results and hypothesis testing are presented in this section. A discussion of the findings follows.

Descriptive Results

The middle schools in this study were representative of the state of Missouri. The majority of the schools in the study had the grade configurations of 6-7-8 (51.2%), 7-8

(21.6%), and 5-6-7-8 (14.4%). These percentages were similar to the grade patterns of the all middle level schools in Missouri. The average population for schools in this study was slightly less than the average population for Missouri's middle level schools. The participating schools had a slightly lower mean percent of students receiving free and reduced priced lunch than the state average. The mean full time equivalent of staffing units for teachers in the study schools was slightly less for the state. The participating schools communication arts and mathematics state achievement results were slightly higher than state averages. In all, the 123 schools in this study were very similar to the 336 middle level schools in Missouri.

Teachers in 123 middle level schools participated in this study. Seventy-two percent of the teacher respondents were female and 25 percent were male. Sixty-two percent of the teachers in this study took some form of an assessment course as part of their teacher certification while 34 had not. Fifty-one percent of the teacher respondents teach core subjects, 27 percent teach non-core subjects, 14 percent of the teachers teach special education courses while five percent teach in other areas.

Middle level principals in the 123 schools had a mean age of 46 years old. The average years spent in education among the principals was eleven years with seven average years spent as a middle level principal.

Hypotheses Testing

Hypothesis one, there are no significant correlational relationships among the items, scales and factors measuring assessment practices, the factors measuring principal instructional leadership and student achievement as measured by the Mathematics and Communication Arts portions of the Missouri Assessment Program, was rejected. The

Audit of Principal Effectiveness Factor 8, Instructional Improvement, showed positive zero-order and partial correlations with TAPS-ML Factor 4 – Collaboration through Assessment, TAPS-ML Scale 6 – Teacher Use of Summative Assessment Data, Overall Teachers Assessment Practices, and Mathematics MAP scores. The Audit of Principal Effectiveness Factor 9, Curriculum Improvement, revealed positive zero-order and partial correlations with TAPS-ML Factor 4 – Collaboration through Assessment, TAPS-ML Scale 2 – Teacher Assessment during Instruction, TAPS-ML Scale 6 – Teachers Use of Summative Data and Overall Teachers Assessment Practices. TAPS-ML Scale 5 – Teacher Develops Self and Peer Assessments correlated with the Mathematics MAP.

The second hypothesis, there was no significant predictive linear relationships between the factors measuring principal instructional leadership and the scales and factors measuring assessment practices, was rejected. All of the variables selected for the model based on correlations explained some variance. The variable, Instructional Improvement, helped to explain part of the variance in Collaboration through Assessment, Use of Summative Assessment Data, and teachers' Overall Assessment Beliefs. The variable, Curriculum Improvement, helped to predict the factor, Assessment during Instruction. It is also important to note that the variable, free and reduced lunch, was not significant in the model, implying that these findings are common across all schools regardless of socioeconomic status.

The third hypothesis, there are no significant relationships among the scales and factors measuring assessment practices and the factors measuring principal instructional leadership and student achievement in mathematics and communication arts, was rejected. While students receiving free and reduced price lunches explained a large

portion of the variance, principals' instructional improvement was a predictor of student achievement in communication arts and principals' curriculum improvement was a predictor of student achievement in mathematics.

The fourth hypothesis, there are no significant differences for the scales and factors measuring assessment practices and the factors measuring principal instructional leadership between schools with high levels of achievement and schools with low levels of achievement in mathematics and communication arts, was rejected. Differences occurred in the Communication Arts quartiles for the variables, Application of Formative Assessment, Teacher Feedback to Students and Parents and Teachers Use of Summative Data. In the Mathematic quartiles, there were differences in the variable, Teachers Feedback to Students and Parents.

Discussion of Findings

The relationships among teachers' assessment practices, instructional leadership and student achievement are discussed in this section. The discussion is organized into sections corresponding to three major findings in this study: a) the relationship of instructional leadership and teachers' assessment practices, b) the relationship of teachers' assessment practices and student achievement and c) the relationship of instructional leadership and student achievement.

Relationship of instructional leadership and teachers' assessment practices

The findings in this study suggest that the principal has an impact on teachers' assessment practices, as well as teachers' beliefs about the collective assessment practices of their school. An examination of the correlations and regressions between instructional leadership and teachers' assessment practices revealed the two instructional leadership

variables, Instructional Improvement and Curriculum Improvement, helped explain four of the assessment practices variables, Collaboration through Assessment, Assessment during Instruction, Use of Summative Data, and Overall Assessment Beliefs.

Strong instructional leaders “affect instructional skills positively through effective clinical supervision, knowledge of effective schooling, and commitment to quality instruction” (Valentine & Bowman, 1984, p. 25). The part of leadership that is described as Instructional Improvement demands a commitment from the principal. This is shown through: (a) a general knowledge of content, (b) a familiarity of teaching strategies and what current research suggests as best instructional practices, and (c) the ability to observe instruction and provide strategies for improvement, as well as recognize strong instruction.

Furthermore, the aspect of leadership known as Curriculum Improvement implies “the principal promotes an articulated, outcome-based curriculum through diagnosis of student needs and systematic program review and change” (Valentine & Bowman, p. 25). This is done through: a) the development of educational goals and objectives that reflect societal needs and trends, b) promoting the diagnosis of individual and group learning needs of students and the application of appropriate instruction to meet those needs, c) administration of a school-wide curricular program based upon identification of content goals and objective and the monitoring of student achievement toward those goals and objectives, d) has a systematic process for program review and change, and e) encourages the articulation of the curriculum program.

The purpose of this section is to illustrate the findings from this study which imply that leadership through instructional improvement and curriculum improvement

has an impact on teachers' assessment practices. The particular assessment practices found to be related to instructional leadership in this study were collaboration through assessment, assessment during instruction, use of summative assessment and overall beliefs about assessment practices.

Collaboration through Assessment.

Research has shown that collaboration among teachers has positive outcomes for students in areas such as engagement, attendance and achievement (Arhar, 1997). Building on that idea, the findings from this study suggest that instructional leadership directly impacts teacher collaboration. The factors, Instructional Improvement and Curriculum Improvement, were both strongly related with the factor, Collaboration through Assessment. In other words, the stronger the instructional leadership in a school, the stronger the collaboration based on assessment data. Additionally, Instructional Improvement was shown to be a predictor for Collaboration through Assessment (Adjusted $r^2 = .074$, $F = 10.761$, $p = .001$).

Collaboration through assessment is more than the process of teachers coming together in some form of a meeting. It is an established process, often led or initiated by the principal, in which groups of teachers come together to discuss student achievement and plan instruction based on assessment data. Informal assessment data, such as classroom assignments or observation records, and formal assessment data, such as writing assessments or objective tests, are used as the basis for the collaboration. This study suggests that the greater the capacity a principal has as instructional leader is related to the level that teachers use collaboration as a form of classroom assessment improvement. The findings in this study are similar to what is found in the literature.

Teacher collaboration without strong leadership may increase the culture among collaborative groups but it takes a leader to implement organizational structures and ongoing professional learning opportunities for collaborative groups to focus on instructional improvement and develop cultures of instructional exploration and improvement (Supovitz, 2002). This supports Marks and Printy's (2003) notion of "shared instructional leadership" (p. 371). Shared instructional leadership focuses on the "active collaboration between principals and teachers around instructional issues meant to enhance the quality of teaching and student performance" (p. 370). In this study, the instructional issue was assessment data, both formal and informal, and teachers indicated that leadership had a strong relationship with the level to which collaboration through assessment occurred.

Assessment during Instruction.

Assessment during instruction practices include: a) thinking about student responses and reacting accordingly, b) providing scoring guides prior to learning as a model of success, c) properly planning questions to provoke student engagement, and d) not advancing to new material until students have mastered the current content. The items illustrate the part of teaching that requires thorough planning which enhances teachers' ability to be very cognizant while teaching, in other words, "think on their feet". This invokes Shepard's (2000; 2005) vision of curriculum in which it is believed that learning and development are culturally embedded, socially supported processes, so socialization between teachers and students is an important part of academic practices and curriculum. In other words, the interactions that occur while instruction occurs are a part of the curriculum. Teachers develop, plan, and take into consideration students' assessment

results, which in this case may simply be an answer to a question or a response during a discussion. The assessment results are used to make decisions about how to proceed with instruction.

In this study, the instructional leadership factor, Curriculum Improvement, was related to the classroom assessment practices that teachers employ during instruction (Adjusted $R^2 = .024$, $F = 3.986$, $p = .048$). Strong leaders provide support through obtaining enough knowledge about assessments to lead teachers and teams of teachers to develop assessment strategies, such as scoring guides and classroom instructional methods (Bottoms, n.d.). Strong leaders also “promote the diagnosis of individual and group learning needs of student and application of appropriate instruction to meet those needs” which is an indicator of Curriculum Improvement (Valentine & Bowman, 1984).

The process of assessment during instruction is, more than likely, the part of instruction that occurs when it is just the teacher and the students in the room so teachers need to be prepared, through knowledge about informal classroom assessments, such as questioning or discussions, and equipped for how to respond to those assessments. The findings from this study indicate in schools where leaders are strong at curriculum improvement, the teachers are more likely to exhibit this skill.

Use of Summative Data.

Use of summative data is an assessment practice that occurs after final summative assessments, such as benchmark assessments or district-mandated tests, have been administered. The assessment data are analyzed at the building level, the teacher and classroom level, and finally, the individual student level. Strengths and weaknesses are established forming the basis for change at the classroom and school-wide levels.

Changes may occur in instructional practices, curriculum or programs that the school supports.

In this study, instructional improvement proved to be a predictor of principals' influence on the changes that are made in instructional delivery, curriculum and/or programs based on formal assessment data. Instructional leaders maintain an awareness and knowledge of recent research about the learning process and lead teachers through the analysis of summative data thus establishing common goals and a united focus for instructional improvement.

Relevant literature supports the relationship. Black and Wiliam (1996) emphasized that the most important aspect of assessment is “not the assessments themselves, but the functions they serve” (p.1). The principal is responsible for providing a “vision” about assessment and ensuring that something is actually done with the information that is amassed from the assessments that are administered in their buildings (Cizek, 1995). From a school-wide perspective, this would emphasize continuous improvement with the objectives of aligning school expectations, providing regular feedback on student learning, and promoting thinking about classroom strategies for enhancing learning (Cizek; Noonan & Renihan, 2006).

Considering leadership is closely tied to teachers' use of summative data, it is important to recognize the relevant literature on instructional leadership and intellectual stimulation. The principal has the managerial task of physically bringing teachers together to study summative data, but it is also a responsibility of the leadership to “ensure that faculty and staff are aware of the most current theories and practices regarding effective schooling and make discussions of those theories and practices a

regular aspect of the school's culture" (Marzano, et. al, 2005). In addition, strong leaders help teachers use student assessment data to measure their own effectiveness as an educator as well as for student learning (Popham, 2004 Principals can accomplish this through engaging teachers in reflective dialogue about classroom assessment practices, gauging levels of assessment literacy, initiating school and classroom-level action plans based upon student data, and promoting wide-range discussions of assessment practices and beliefs (Noonan & Renihan, 2006).

Overall Collective Assessment Practices.

Respondents were asked about the collective ability of the faculty of their school to use assessment data to improve instruction and student achievement. This required teachers to think beyond their own abilities as classroom assessment practitioners and think about the process of classroom assessment from a building-wide perspective.

Overall Collective Assessment Practices were strongly related to the Instructional Leadership factors of Instructional Improvement ($r = .353, p = .000; pr = .357, p = .000$) and Curriculum Improvement ($r = .316, p = .000; pr = .323, r = .000$). Instructional Leadership proved to be a predictor for this variable (Adjusted $R^2 = .117, F = 17.272, p = .000$). This suggests that instructional leadership makes an impact on the overall collective assessment practices in a school.

Pederson and Liu's (2003) study on teachers' perceptions about building-wide assessment practices revealed that teachers will use assessment practices for others' intentions (i.e. the school or district) as long as the assessment is also designed with the individual teacher's intentions for assessment in mind. The importance of collaboration was emphasized as key to shaping individual teacher's intentions and beliefs about

assessment as well as creating a common language that promotes the assessment “vision” for the collaborative group.

Pederson and Liu’s (2003) study link the findings from this study about collaboration through assessment and the use of summative assessments. This study implies that leadership is instrumental in developing the focus and “vision” for assessment practices necessary to shape teachers’ beliefs about the Overall Collective Assessment Practices in their school.

Figure 6 illustrates the relationship that the instructional leadership factors, Instructional Improvement and Curriculum Improvement, have on Teachers’ Assessment Practice variables, Collaboration through Assessment, Assessment during Instruction, Use of Summative Data, and Overall Assessment Beliefs.

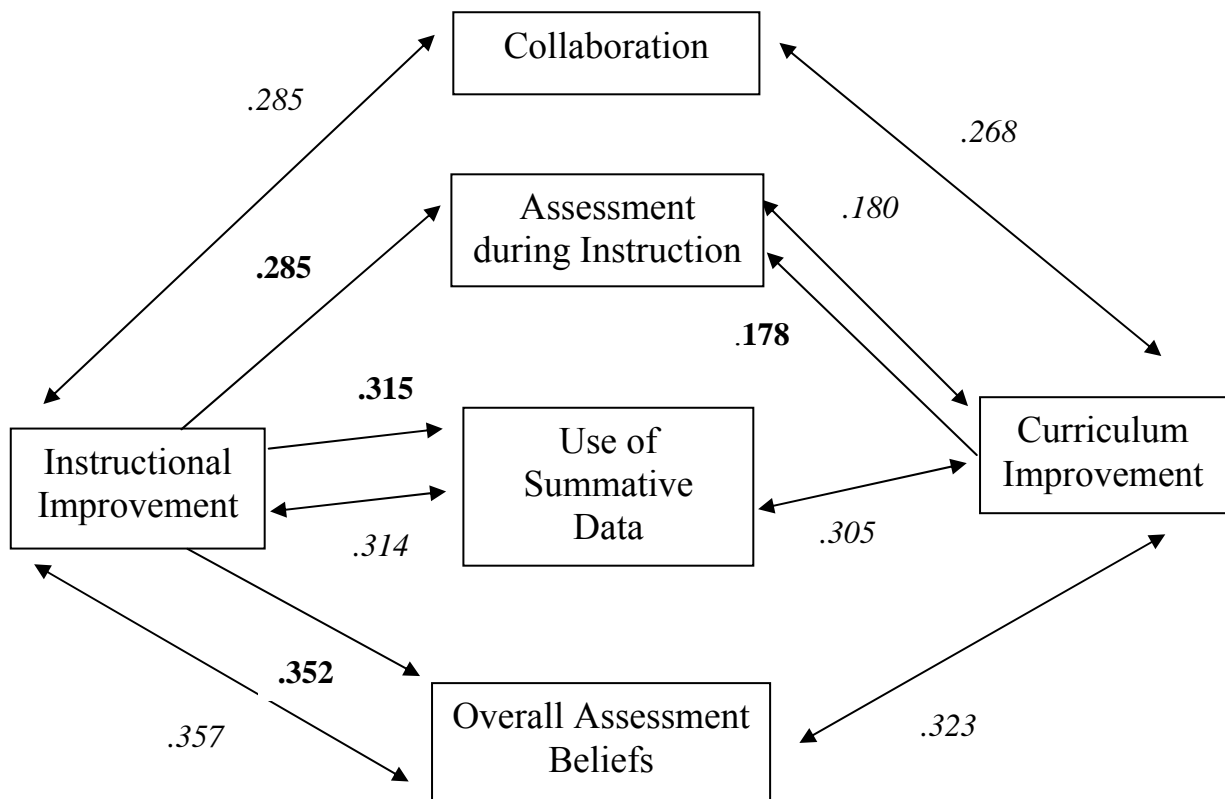


FIGURE 6. *Correlations and Regressions between Instructional Leadership factors of Instructional Improvement and Curriculum, and Middle Level Assessment Practices scales and factors of Collaboration through Assessment, Assessment during Instruction, Use of Summative Data, Overall Assessment Beliefs.*

↔ = Mutually Influential (Partial Correlation in italics)

→ = Directly Influenced (Standardized Beta Coefficient, β , in bold)

The discussion on instructional leadership and its influence on teachers' assessment practices prompted further investigation into the relations among the assessment variables. The findings from this study suggest that instructional leadership, both instructional improvement and curriculum improvement, influence three aspects of teachers' assessment practices – collaboration, assessment during instruction and use of

summative data. These variables, in turn, influence teachers' perceptions about the collective ability of their school to use assessment data to improve instruction and student achievement (See Figure 7).

When considering the time and place in which these practices occur, a model for assessment practices was developed. Collaboration could be seen as an activity that occurs before the assessment process takes place, through planning meetings and analysis of prior assessments. Assessment during instruction occurs while the assessment is in progress. Teachers actively interact with students to glean information through discussions and informal classroom assessment data. The use of summative assessment data occurs after assessments have been administered when time is taken to thoroughly think about, analyze and make changes, if necessary, based on the assessment results. An interpretation of the findings of this study would suggest that when teachers receive support on assessment strategies through instructional leadership, then the capacity to impact an "assessment cycle", or the practices needed before, during and after assessments, are enhanced. The level to which the "assessment cycle" is established would thus have an influence on teachers' overall beliefs about the assessment practices in their schools.

Figure 7 depicts the influences the Instructional Leadership factors had on teachers' assessment practices which impact those assessment practices then, in turn, teachers' overall collective assessment practices.

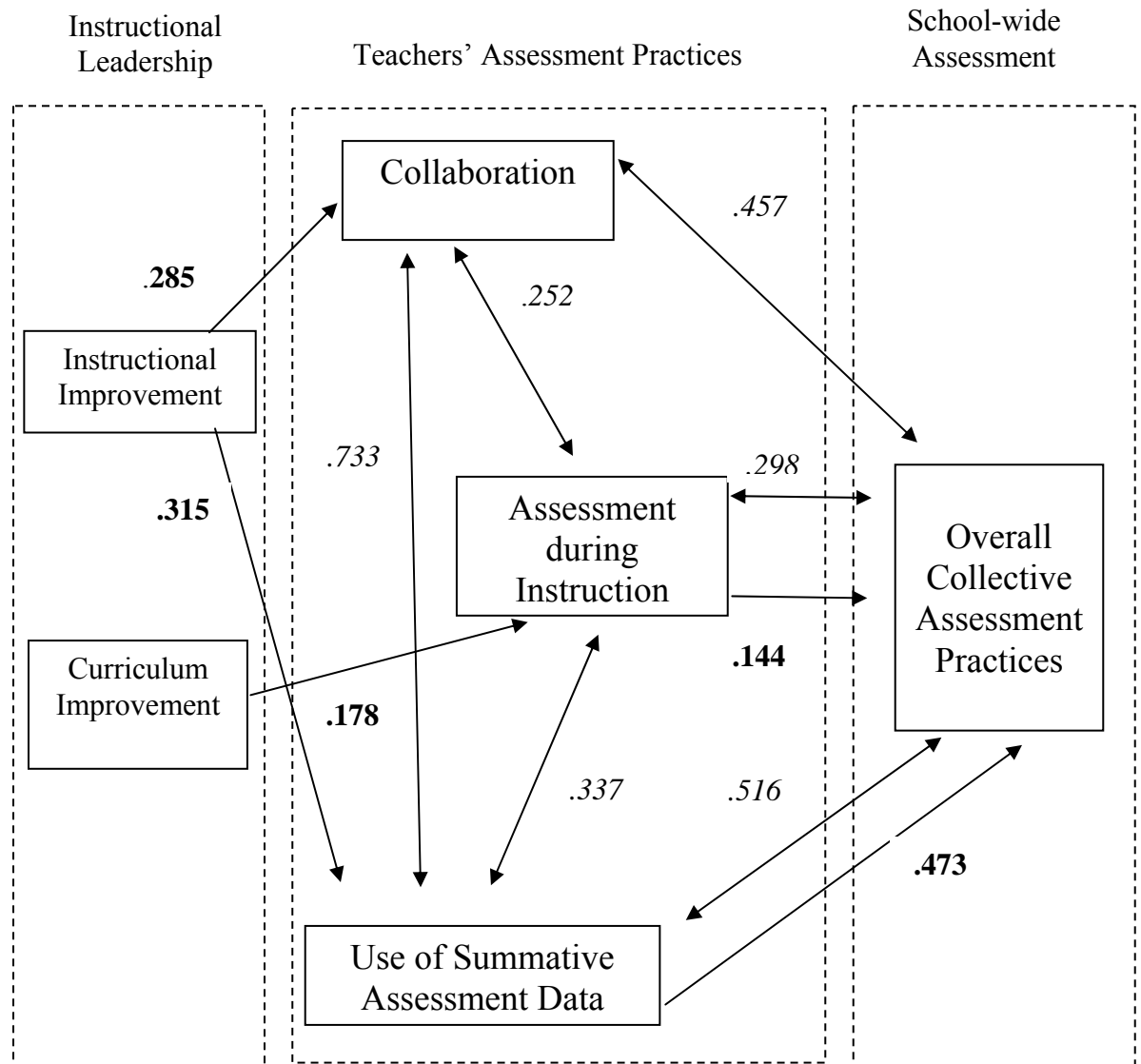
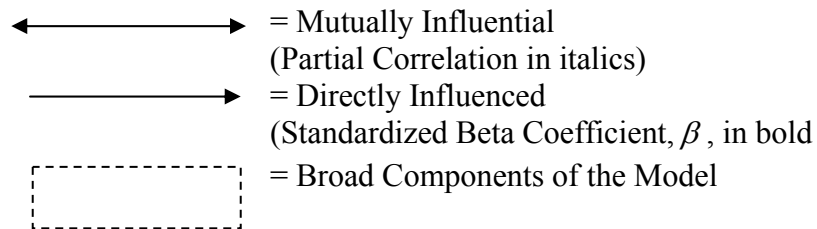


FIGURE 7. *Correlations and Regressions for Instructional Leadership factors and Assessment Practices on Overall Collective Assessment Practices*



Relationship of teachers' assessment practices and student achievement

The effective use of assessment data provides an avenue for understanding not only individual student achievement, but, in a larger perspective allows staff to reflect on effectively understanding design, adaptation of instruction, and use of summative data to modify programs and practices. For example, teachers need to thoroughly plan instruction through the use of previous assessments as a guide (Black & Wiliam, 1998a). They need to provide ample information to students throughout the entire assessment process through proper feedback (Sadler, 1989). Teachers need to remediate, adjust instruction and reteach (Chappuis, et al., 2005). Final summative data needs to be analyzed to determine students' performance level and as measure for the teachers to determine their own effectiveness as an educator (Popham, 2004). One of the driving premises for this study was that if teachers are properly using classroom assessments to improve student achievement as measured by informal assessments, then classroom assessment practices should impact student achievement as measured by annual standardized achievement tests.

The purpose of this section is to interpret the findings comparing teachers' assessment practices and student achievement as measured by the Communication Arts and Mathematics sections of the Missouri Assessment Program. The achievement data used in this study for each participating school was computed by identifying the percent of students at each grade level who mastered the state assessment, thus creating the school composite average of those who reached mastery.

This study used two approaches to examine possible relationships between middle level teachers' assessment practices and student achievement. First, partial correlations

and linear regressions between each of the scales and factors of the Middle Level Teachers Assessment Practices Survey and the Communication Arts and Mathematics portions of the MAP were analyzed. An examination of the correlational relationships between teachers' assessment practices and student achievement produced a single correlation with Scale 5 – Teacher Development of Self and Peer Assessments and student achievement in mathematics ($r = .213, p = .041$; $pr = .215, p = .050$). There were no predictive linear relationships between assessment practices and student achievement.

A possible explanation for the lack of significant relationships for teachers' assessment practices and student achievement may be that teachers in Missouri have not had enough training and exposure to classroom assessment strategies to influence practices and thus create a true difference among assessment practices. This explanation is supported by the literature on classroom assessment practices in which teachers' assessment practices have been deemed as extremely weak. Key weaknesses were: (a) the main type of assessments were superficial and rote memorization, (b) teachers do not generally prepare questions asked in class, (c) feedback remains mainly at the level of providing a percentage or letter grade with no other comments, (d) teachers do not trust or use the assessment results of colleagues to help prepare classroom assessments, and (e) there is little focus on outcomes as speculation of teaching performance or as a means for critical reflection (Black, P. & Wiliam, D., 1998b; Cizek, et al., 1995; Senk, et al., 1997).

In the second analysis, schools were grouped into quartiles based on the schools Communication Arts and Mathematics MAP scores and were tested for significant differences. Significant findings were present for Factor 2 – Application of Formative Assessment, Scale 4 – Teacher Feedback to Students and Parents and Scale 5 – Teacher

Development of Self and Peer Assessments from teachers' responses to the TAPS-ML.

Tukey post hoc analysis documented the specific differences (Table 45).

Table 46

Differences among Student Achievement Quartiles for the TAPS-ML Factors

	Factor 2 – Application of Formative Assessment	Scale 4 – Teacher Feedback to Students and Parents	Scale 5 – Teacher Development of Self and Peer Assessments
CA1	>CA2 >CA3	>CA2 >CA3	>CA2 >CA3 >CA4
CA2	<CA1	<CA4	
CA3	<CA1	<CA4	<CA4
CA4			
MA1			
MA2			
MA3		<MA4	<MA4
MA4			

The results for the differences among quartiles revealed surprising information.

Quartile one, the lowest quartile in student achievement on the MAP, was shown to be significantly higher than the second and third quartiles for the variables, Application of Formative Assessment and Teacher Feedback to Students and Parents. For the variable, Teacher Development of Self and Peer Assessments, quartile one was higher than quartiles two, three, and four.

A possible interpretation as to why teachers from the lowest quartile are implementing classroom assessment strategies at a higher rate than the other quartiles may come from external pressures that are being placed on low achieving schools to improve. Schools that fail to meet the Adequate Yearly Progress (AYP) for the students in their schools are faced with corrective actions such as: (a) parents are given an option to transfer their child out of the school, (b) extend the school day or school year, (c) provide staff development in scientifically research-based professional development and (d) institute a new curriculum grounded in scientifically based research (Missouri Department of Elementary and Secondary Education, 2007). The last two corrective actions listed may be the influential factors if research-based professional development and curriculum on classroom assessment practices have been mandated. Additionally, school districts that fall in the lowest quartile may be more inclined to allocate money and time towards professional development on classroom assessment practices in an attempt to move out of the lowest achieving schools.

The question remains, if teachers in the lowest performing quartiles are significantly administering classroom assessments at a higher frequency than teachers in the other quartiles, then why are they still falling behind in student achievement? Three possible explanations may provide insight. The first explanation comes directly from the research. Studies have demonstrated that low achieving schools are highly correlated with schools with populations of low socioeconomic status (e.g. Sirin, 2005). Teachers in schools with low socioeconomic populations face a plethora of problems that impede student achievement beyond the corrective influential assessment practices that teachers use. The data from this study shows that teachers in the lowest quartile are clearly

implementing more strategies in the area of classroom assessment practices while working with the lowest socioeconomic group ($M = 57\%$ F/R).

Another explanation comes from the literature on change in schools. When schools implement change, in this case, classroom assessment practices, typically there is an initial drop in student achievement results (Porter & Snipes, 2006; Le, et al., 2006). When changes in practices are introduced, teachers tend to experience disequilibrium as they work their way through the learning process (Dewey, 1993). The low achievement scores from schools in the lowest quartile may show growth as teachers work their way through the change process and begin to implement classroom assessment practices at a higher level of confidence and comprehension.

A final explanation was revealed through the open-ended questions at the end of the TAPS-ML. Answers to the open-ended questions were divided by the quartiles based on the student achievement on the Missouri Assessment Program and analyzed for similarities and differences. When comparing the top two higher achieving quartiles with the bottom two lower achieving quartiles, there were similarities in teachers' responses about what types of assessments were most beneficial, but there were differences in the explanations as to why the assessments were beneficial.

Teachers' answers from the top two quartiles centered on the thinking process associated with classroom assessments. Teachers expressed that students need to "form their own thoughts" (Teacher from School #56) and gain "exposure to the process side of answering questions" (Teacher from School #154) as a way to "internalize information" (Teacher from School #287). Conversely, teachers' responses from schools in the bottom two quartiles focused their reasoning on the similarity of classroom assessment to the

MAP and how replication will help students score well on the test. Typical answers indicated that good assessments “allow the students to sample the MAP test before it arrives” (Teacher from School #134), “looked like the MAP test” (Teacher from School #165) and allow students to become “comfortable with the MAP format” (Teacher from School #58) “so you get a high score” (Teacher from School #38).

It is important to note that the survey items in this study measured the frequency, not quality of assessment strategies. While the teachers in the lowest quartile are employing classroom assessment practices at a higher rate, the quality of the practices remains untested except through the information provided through the open-ended questions.

Additionally, differences in responses may be tied to the external pressures that are being placed on the lowest achieving schools. With corrective actions such as extended school years and transfer policies a constant threat, teachers and principals may be drawn to focusing their energies more on “fixing the problem” of low test scores rather than focusing on making lasting changes in students’ learning. This illustrates the claim the Noonan and Renihan (2006) make that “some school professionals hold inappropriate or mistaken beliefs that constitute serious inhibitors to the effectiveness of the connections among teaching, testing, and learning” (p. 3).

The differences in teachers’ classroom assessment practices revealed another interesting facet among the quartile levels. Schools in the third quartile had the lowest means for the factors and scales of Application of Formative Assessment for Communication Arts, Teacher Feedback to Students and Parents for Mathematics, and Teacher Development of Self and Peer Assessment for both Communication Arts and

Mathematics. Socioeconomic status does not provide an explanation for the low assessment practices because the quartile is in the top half of the state for SES ($M = 37\%$ F/R). This study suggests that teachers in the third quartile are lacking compared to teachers in other quartiles in terms of implementing assessment strategies. One possible reason may be that teachers in this quartile may be reluctant to change their assessment practices because they are scoring in the top half of the state, quite possibly without any warnings or any notice at all.

Relationship of instructional leadership and student achievement

The factors for Instructional Leadership proved to be a minimal predictor of student achievement. One partial correlation occurred for Mathematics in grade seven with the Instructional Improvement factor ($pr = .220, p = .044$). When predicting instructional leadership on student achievement in Communication Arts, students receiving free and reduced priced lunch was the best predictor, but Instructional Improvement was also a significant predictor (Adjusted $R^2 = .504, F = 63.460, p = .000$).

Instructional leadership also had an impact on student achievement in Mathematics through Curriculum Improvement. Free and reduced price lunch was the strongest predictor. The factor, Curriculum Improvement also showed significance (Adjusted $R^2 = .535, F = 71.825, p = .000$). The results in this study are similar to findings in the literature.

Though many studies on instructional leadership and its impact on student achievement have failed to produce direct relationships most indicate that leadership impacts achievement indirectly (e.g. Hallinger, Bickman & Davis, 1996; Hallinger & Heck, 2000; Krug, 1992). In this study, strong relationships were found between

instructional leadership and teachers' assessment practices. Furthermore, teachers' assessment practices were found to be related to student achievement when divided into quartiles. Considering the influence of the findings from previous studies and adding the findings from this study on teachers' assessment practices, it is evident that instructional leadership indirectly impacts student achievement.

Implications

This research study provides information about teachers' assessment practices, principals' influence on assessment practices and the resulting impact on student achievement at the state-wide level. The notion of classroom assessment practices is relatively new to educators in Missouri. The information from this study can serve as baseline data for schools in Missouri, providing the opportunity for future studies to document changes in assessment practices.

Implications for Practice

Principals, teachers and students should benefit from training in assessment practices. Traditions in assessment in the United States are solidly established and engrained. Much work is needed to change these traditions and the thinking behind them. As Stiggins (2004) proclaimed, "we have inherited an assessment legacy that has actually prevented us from tapping the full power of assessment for school improvement" (p. 22). The main focus for practice for all should be that assessments are for learning and growth, not only summative judgments.

Principal preparation for assessment leadership

Principals need to understand and have the capacity to provide leadership for best assessment practices. Not only do principals have an obligation to their teachers, but they

are responsible to students to gain knowledge and skills about assessments. The Interstate School Leader Licensure Consortium (ISSLC) supports this through its standard that school leaders promote the success of all students by advocating, nurturing and sustaining an instructional program conducive to student learning (Green, 2001).

Chappuis, et al. (2006) provide necessary steps that principals need to take to implement solid assessment systems. Administrators must: (a) assist in the development of a clearly articulated and appropriate set of achievement standards for each student as the foundation for quality assessment, (b) have a commitment to providing accurate, understandable, and usable information about student achievement to all users of assessment results, (c) build an assessment-literate culture, (d) consider how best to collect, store, manage, and communicate information about student achievement, and (e) lay a foundation of assessment policy that supports quality practices (p. 51-52).

An important aspect of instructional leadership that was supported through this study is the employment of collaboration based on assessment data. This study added to previous studies on the importance of principals implementing time for collaboration for teachers. Dividing the teachers' perceptions of assessment practices into quartiles based on MAP scores revealed that the highest performing schools on the MAP were also the schools that used collaboration through assessment with the most frequency.

Principals play an important part in the assessment process by developing an assessment "vision" in their school and supporting it through solid collaborative discussions about student formal and informal assessment data. Considering the growing number of mandated assessments and growing number of groups interested in various assessment outcomes, the principal must have a vision regarding how the various

assessment parts fit together (Cizek & Fitzgerald, 1996). When establishing the assessment vision and goals in a school, it is essential that the principal have an appreciation of the integral relationships among teaching, learning and assessment and the resulting influence on student achievement (Noonan & Renihan, 2006; Girvin, 2005).

It is also essential to include the teachers in the assessment leadership of the school. The act of collaboration through assessment, or bringing together teachers to focus on assessment data lends itself to the research on professional learning communities (DuFour & Eaker, 1998). The common tenets of professional learning communities are: a) ensuring that students learn b) creating a culture of collaboration, c) focusing on results, and d) hard work and commitment (DuFour, Eaker, & DuFour, 2005). Classroom assessment data provides the stimulus for each of the tenets of a professional learning community (Stiggins, 2005).

Teacher Preparation

Sixty-two percent of the teachers in this study responded that they were required to take an assessment course as part of teacher certification. Assessment courses as part of teacher development are a step in the right direction; however the curriculum in assessment courses should include the current research on classroom assessments. Also teachers need to develop the capacity to construct various classroom assessments such as selected response, written response, performances assessments and direct personal communication. Teachers' philosophies about current classroom assessment practices should begin during their pre-service years if coursework was developed to support current research.

Teachers at the in-service level also need training in assessment development. Additionally, training needs to be focused on providing proper feedback to students, adjusting instruction to meet students needs, and developing students through teaching self and peer assessment strategies.

Student training

Teachers need to take the time to involve students in the assessment process. This study supports existing literature on students' use of self and peer-assessment to improve student achievement (Black & Wiliam, 1998a; Ellis, 2001; Marshall & Weinstein, 1984). The results from this study indicated that there was a relationship between student achievement in Mathematics and teachers' ability to develop students' self-assessment strategies. Students will benefit through being included in the assessment process.

Teachers need training in learning how to properly model self-assessment practices to their students (Popham, 2004). It cannot be assumed that students will know how to do this because, traditionally, assessments have been something "done to them," not a process in which they had an active role.

Implications for Future Research

This exploratory study of Missouri teachers' middle level assessment practices set the stage for additional research. For instance, a study in the elementary or high schools in the state of Missouri would provide an interesting comparison and possibly illuminate the differences in assessment practices across school levels. A grade level or content area study would possibly reveal specific strategies and strengths among the grades levels and different content area teachers in Missouri.

The analysis of instructional leadership and its impact on teachers' assessment practices model provides a conceptual frame for further research. . The relationship between leadership and collective assessment practices as mediated by individual practices, school climate and culture, teacher commitment or efficacy, are valid areas for future research to link leadership with school-wide assessment.

The findings from this study revealed disproportionate practices among schools divided by quartiles based on the results of the MAP test. The teachers' from the top performing quartile and the lowest performing quartile were found to be performing classroom assessment practices at a much higher level than teachers from the middle two quartiles. A qualitative analysis and additional quantitative studies would be valuable to confirm and explicate the cause of the phenomenon. A follow-up study utilizing qualitative methods to observe teachers in their classrooms would provide good insight about the perception data collected through this study. A qualitative study would not only test the accuracy of teachers' perceptions about their classroom assessment practices, but it would also provide information about the quality of teachers' classroom assessment practices.

Conducting a longitudinal study would be beneficial to Missouri educators as a means to show growth and provide information on the impact that classroom assessment practices have on student achievement over time. As educators in Missouri learn more about utilizing classroom assessment practices, a study measuring growth through the years might impact education in Missouri. Multiple opportunities for research are provided through the development of the Teachers' Assessment Practices Survey (TAPS). Data provided from future administrations of the survey will help to build on the

conclusions of this study, as well as refinement of the factors. In addition, the instrument can be revised to measure principals' assessment leadership. Moving the focus of the instrument to measure principals' perceptions of their assessment leadership would build on the findings from this study on teachers' perceptions of principals' instructional leadership. As the role of principal continues to change, particularly due to the growing number of mandated assessments, assessment leadership will become an important form of leadership (Cizek, 1995).

Closing Thoughts

Recent educational literature is replete with recommendations for change in the assessment practices across our nation's schools. From this study, we begin to see a glimmer of relationships between leadership and assessment, and assessment and achievement. Most educators today are convinced that the effective use of formative and summative assessment affect academic success, but few seem to understand the depth of the effect. While this study does not adequately answer that complex question, which would require a considerably larger data base to accommodate the inclusion of many additional variables, it does provide initial insight about the effect of some variables and adds to the literature four new factors for evaluating the degree of assessment practices in schools. Hopefully, quantitative research linking assessment with other variables and with achievement will blossom in the coming years.

REFERENCES

- Airasian, P. (2000). *Assessment in the classroom* (2nd ed.). Boston: McGraw-Hill.
- Airasian, P. (1991). *Classroom assessment*. Boston: McGraw-Hill.
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association (AFT/NCME/NEA) (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, 9(4), 30-32.
- Andrews, R., Basom, M., & Basom, M. (1991). Instructional leadership: Supervision that makes a difference. *Theory into practice*, 30(2).
- Appalachian Educational Laboratory (AEL), (2005). *Assessing the relationship between questioning and understanding to improve learning and thinking and student achievement in mathematics: A pilot study*. Retrieved November 1, 2006 from www.edvantia.org
- Arhar, J. M. (1992). Interdisciplinary teaming and social bonding of middle level students. In J.L. Irvin (Ed.), *Transforming middle level education: Perspectives and possibilities*. Needham Heights, MA: Allyn & Bacon.
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 93(2), 117-149.
- Barksdale-Ladd, M. & Thomas, K. (2000). What's at stake in high-stakes testing? Teachers and parents speak out. *Journal of Teacher Education*, 51(5), 384-397.
- Beatty, B.R. & Brew, C.R. (2003). Trusting relationships and emotional epistemologies: A foundational leadership issue. *School Leadership and Management* 24(3), 329-356.

- Beck, L. & Murphy, J. (1993). *Understanding the principalship: Metaphorical themes 1920s-1990s*, New York: Teachers College Press.
- Bergan, J., Sladeczek, I., Schwarz, R. & Smith, A. (1991). Effects of a measurement and planning system on kindergarteners' cognitive development and educational programming. *American Educational Research Journal*, 28, 683-714.
- Black, P. (1998). *Testing: Friend or foe?* London: Routledge-Falmer.
- Black, P. & Wiliam, D. (1998a). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 139-148.
- Black, P. & Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in Education*, 5, 7-74.
- Black, P. & Wiliam, D. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), 1-11.
- Black, P., Harrison, C., Lee, C. Marshall, B., & Wiliam, D. (2004). Working inside the black box. *Phi Delta Kappan*. 86(1), 13-22.
- Blase, J. & Blase, J. (1998). *Handbook of instructional leadership: How really good principals promote teaching and learning*, Thousand Oaks, CA: Corwin Press, Inc.
- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-on-one tutoring. *Educational Researcher*, 13(6), 4-16.
- Bloom, B. S., Hastings, J.T., & Madaus, G. (1971). *Handbook on the formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Bottoms, G. (n.d.). What school principals need to know about curriculum and

- instruction. Retrieved April 8, 2007 from www.sreb.org
- Bowman, M. & Valentine, J. (1984). Audit of principal effectiveness: A user's manual.
- Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education*, 10(2), 161-180.
- Brookhart, S. M. & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education*, 12(4), 409-425.
- Brown, A. (1994). The advancement of learning. *Educational Researcher*, 23(8), 4-12.
- Brown, G. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education*, 11(3), 301-318.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation of interest and performance. *British Journal of Educational Psychology*, 58, 1-14.
- Chappuis, S., Stiggins, R.J., Arter, J., & Chappuis, J. (2004). *Assessment FOR learning*. Portland: Assessment Training Institute.
- Cizek, G. (1995). The big picture in assessment and who ought to have it. *Phi Delta Kappan*, 77(3), 246-249.
- Cizek, G. & Fitzgerald, S. (1996). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159-179.
- Cohen, E., Lotan, R., Abram, P., Scarloss, B., & Schultz, S. (2002). Can groups learn? *Teachers College Record*, 104(6), 1045-1068.
- Congress of the United States. (1992). Lessons from the past: A history of educational

- testing in the United States. In Office of Technology Assessment (Eds.), *Testing in American schools: Asking the right questions*. Washington, DC: U.S. Office of Technology Assessment.
- Cronbach, L. J. (1977). *Educational psychology* (3rd ed.). New York: Harcourt Brace Jovanovich.
- Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-480.
- Cuban, L. (2004). Looking through the rearview mirror at school accountability. In K. Sirotnik (Ed.), *Holding accountability accountable: What ought to matter in public education*. New York: Teachers College Press.
- Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work*. New York: Teachers College Press.
- Darling-Hammond, L. & Wise, A. (1985). Beyond standardization: State standards and school improvement. *The Elementary School Journal*, 85(2), 315-336.
- Dewey, J. (1993). *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston: Houghton Mifflin Company.
- DuFour, R. & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement*. Bloomington, Indiana: National Educational Service.
- DuFour, R., Eaker, R., & DuFour, R. (2005). *On common ground: The power of professional learning communities*. Bloomington, Indiana: National Educational Service.
- Earl, L., Levin, B., Leithwood, K., Fullan, M., Watson, N., Torrance, N., Jantzi, D.,

- Mascall, B., & Volante, L. (2003). England's National Literacy and Numeracy Strategies: Final report of the external evaluation of the implementation of the strategies. Department of Education and Employment, England.
- Elawar, M. & Corno, L. (1985). A factorial experiment in teachers' written feedback on student homework: Changing teacher behavior a little rather than a lot. *Journal of Psychology*, 77(2), 162-173.
- Ellis, G. (2001). Looking at ourselves – self-assessment and peer assessment: Practice examples from New Zealand, *Reflective Practice*, 2(3), 289-302.
- Farr, R. (1992). Putting It All Together: Solving the Reading Assessment Puzzle. *Reading Teacher*, 46(1), 26-37.
- Fienberg, H. (2002). *Davis' failed test*. Retrieved December 12, 2006 from www.hfienberg.com/clips/ca_test.htm
- Fisk, E. & Resnick, L. (1999). *Developing principals as instructional leaders*. Philadelphia, Pennsylvania: University of Pittsburgh, HPLC Project, Learning Research and Development Center.
- Fortana, D. & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology*, 64, 407-417.
- Fox, D. (2003). Guiding instruction through assessment: What principals need to know. *Leadership*.
- Fullan, M. (1997). What's worth fighting for in the principalship? New York: Teachers College Press.
- Gerberich, J. R. (1963). The development of educational testing. *Theory into Practice*,

2(4), 184-191.

Gilles, C. (1995). Reflections. *Primary Voices K-6*, 3(1), 39-43.

Girvin, N. (2005). *The principal's role in k-12 professional development*. Retrieved from <http://www.askasia.org>

GOALS 2000: Educate America Act (1994). *The Clinton Administration Education Program*. Retrieved September 14, 2006 from www.ed.gov/G2K/

Graue, M. (1993). Integrating theory and practice through instructional assessment. *Educational Assessment*, 1(4), 283-310.

Green, L. (2001). *Practicing the art of leadership: A problem-based approach to implementing the ISSLC standards*. Columbus, OH: Prentice-Hall.

Guskey, T. R. (2005). Formative classroom assessment and Benjamin S. Bloom: Theory, research, and implications. *Man and Society*, 57(4), 219-233.

Hallinger, P. & Heck, R. (2000). *Exploring the principal's contribution to school effectiveness, 1980-1995*. Washington, D.C.: Institute for Educational Leadership.

Hallinger, P. & McCary, C.E. (1990) Developing the strategic thinking of instructional leaders. *The Elementary School Journal*, 91(2) 89-108.

Hallinger, P. & Murphy, J. (1985). Assessing the instructional management behavior of principals. *The Elementary School Journal*, 86(2), 217-247.

Hallinger, P., Bickman, L. & Davis, K. (1996). School context, principal leadership, and student reading achievement. *The Elementary School Journal*, 96(5), 527-549.

- Hargreaves, E. (2005). Assessment for learning? Thinking outside the (black) box. *Cambridge Journal of Education*, 35(2), 213-224.
- Hearne, J. (2001). Assessment as a tool for learning. *New Horizons for Learning*. Retrieved November 12, 2006, from www.newhorizons.org/strategies/assess/hearne.htm
- Henning, J. (2006). Teacher leaders at work: Analyzing standardized achievement data to improve instruction. *Education*, 126(4), 729-737.
- Herbert, G. (1997). Practical assessment and testing in a secondary school. In C. Cullingford (Ed.) *Assessment versus Accountability*. London: Cassell.
- Higgins, et al., (1994). Placing assessment into the hands of young children: A study of student-generated criteria and self-assessment. *Educational Assessment*, 2(4), 309-324.
- Hoy, A. & Hoy, W. (2003). *Instructional leadership*. 1st Edition. Boston, MA: Allyn and Bacon.
- Impara, J.C. & Plake, B. (1995). Comparing counselors', school administrators', and teachers' knowledge in student assessment. *Measurement and Evaluation in Counseling and Development*, 28(2).
- Jackson, A. & Davis, G. (2000). *Turning points 2000: Educating adolescents in the 21st century*. New York: Teachers College Press.
- Johnson, B. & Christensen, L. (2000). *Educational research: Quantitative and qualitative approaches*. Needham Heights, MA: Allyn & Bacon.
- Karhuse, A. (2007). House and senate have differing priorities for NCLB reauthorization, *Newsleader*, 54(5), 1, 11.

- Karp, S. (2004). No child left behind: The test. Rethinking Schools Online. Retrieved December 2, 2006, from http://www.rethinkingschools.org/archive/19_01/test.191.shtml
- Koretz, D. (2002). Limitations in the use of achievement tests as measures of educators' productivity. *The Journal of Human Resources*, 37(4), 752-777.
- Krug, S. (1992). Instructional leadership, school instructional climate, and student learning outcomes (Tech. Rep. No. 143). Urbana, IL: University of Illinois, National Center for School Leadership
- Lambert, D. & Lines, D. (2000). *Understanding assessment*. London: Routledge/Palmer.
- Larabee, D. (2005). Progressivism, schools and schools of education: An American romance, *Paedagogica Historica*, 41(1&2), 275-288.
- Lashway, L. (2002). Developing instructional leaders. *ERIC Digest*, 160.
- Le, V., Stecher, B., Lockwood, J., Hamilton, L. & Robyn, A. (2006). Improving mathematics and science education: A longitudinal investigation of the relationship between reform-oriented instruction and student achievement (ED494015). Washington, DC: RAND Corporation.
- Leithwood, K. & Duke, D. (1998). Mapping the conceptual terrain of leadership: A critical point of departure for cross-cultural studies. *Peabody Journal of Education*, 73, 31-50.
- Leithwood, K., Steinbach, R. & Jantzi, D. (2002). School leadership and teachers' motivation to implement accountability policies. *Educational Administration Quarterly*, 38(1), 94-119.
- Linn, R. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.

- Lukin, L., et al. (2004). Facilitating the development of assessment literacy. *Educational Measurement: Issues and Practice*, 26-36.
- Macmillan, R., Meyer, M., & Sherman, A. (2001). *The evolving role of educational administrators*. Paper presented for the 2001 Pan-Canadian Education Research Agenda Symposium Teacher Education/Education Training: Current Trends and Future Suggestions.
- Many, T. & Jakicic, C. (2006). A steadily flowing stream of information gives teachers much-needed data. *Journal of Staff Development*, 27(1), 46-51.
- Marks, H. & Printy, S. (2003). Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*, 39(3), 370-397.
- Marshall, H. & Weinstein, R. (1984). Classroom factors affecting students' self-evaluations: An interactional model. *Review of Educational Research*, 54 (3), 301-325.
- Marzano, R., Waters, T., & McNulty, B. (2005). *School leadership that works: From research to results*. Aurora, CO: Mid-Continent Research for Education and Learning
- McMillan, J. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, 34-43.
- McMillan, J.H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 125-136.
- McTighe, J. & O'Connor, K. (2005). Seven practices for effective learning. *Educational*

Leadership, 63(3), 10-17.

- Meisels, S., Atkins-Burnett, S., Xue, Y., & Bickel, D. (2003). Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores. *Education Policy Analysis Archives*, 11(9).
- Mertler, C. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, 33(1), 49-64.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18(2), 5-11.
- Missouri Department of Elementary and Secondary Education (2006). *Missouri School Directory*. Retrieved January 1, 2007, from <http://dese.mo.gov/directory/download.html>
- Missouri Department of Elementary and Secondary Education (2007). *2006 Schools in school improvement and restructuring*. Retrieved April 10, 2007, from <http://dese.mo.gov/divimprove/nclb>
- Morgan, G. (2006). *Images of Organization*. Thousand Oaks, CA: Sage Publications.
- Murphy, J. (1988). The instructional leadership role of the school principal: An analysis. *Educational evaluation and policy analysis*, 10(2), 71-79.
- Murphy, J. (1985). *Effective schools: What the research reveals*. APEX: Case Report, 1(5), 1-4.
- Myers, P. (2005). The princess storyteller, Clara Clarifier, Quincy Questioner, and the wizard: Reciprocal teaching adapted for kindergarten students. *The Reading Teacher*, 59(4), 314-324.
- Nagy, P. (2000). The three roles of assessment: Gatekeeping, accountability, and

- instructional diagnosis. *Canadian Journal of Education*, 25(4), 262-279.
- National Commission on Excellence in Education (1983). *A nation at risk*. Retrieved July 12, 2006 from www.ed.gov/pubs/NatAtRisk/index/html
- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22, 155-175.
- No Child Left Behind* (2002). Retrieved November 22, 2005 from <http://www.ed.gov/nclb/landing.jhtml>
- Noonan, B. & Renihan, P. (2006). Demystifying assessment leadership. *Canadian Journal of Educational Administration and Policy*, 56.
- North Central Regional Educational Laboratory (NCREL) (1991). *What does research say about assessment?* Retrieved April 27, 2006, from www.ncrel.org/sdrs/areas/stw_esys/4assess.htm
- Pederson, S. & Liu, M. (2003). Teachers' beliefs about issues in the implementation of a student-centered learning environment. *Educational Technology Research and Development*, 51(2), 57-76.
- Plake, B., Impara, J. & Fager, J. (1993). Assessment competencies of teachers: A national survey, *Educational Measurement: Issues and Practice*, 10-39.
- Pope, N. (2005). The impact of stress in self- and peer assessment. *Assessment and Evaluation in Higher Education*. 30(1), 51-63.
- Popham, W. J. (2004). Why assessment illiteracy is professional suicide, *Educational Leadership*, 82-83.
- Popham, W. J. (2002). *Classroom assessment: What teachers need to know* (3rd ed.). Boston: Allyn & Bacon.

- Popham, W. J. (2001). *The truth about testing: An educator's call to action*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Porter, K. & Snipes, J. (2006). *The challenge of supporting change: Elementary student achievement and the Bay Area School Reform Collaborative's focal strategy* (ED494666). New York: Manpower Demonstration Research Corp.
- Quinn, T. (2002). What principals should know and be able to do: Redefining leadership in the standards era. *Principal*, 82(1) 16-20.
- Resnick, D. (1980). Minimum competency testing historically considered. *Review of Research in Education*, 8, 3-29.
- Resnick, L. & Hall, M. (1998). Learning organizations for sustainable education reform. *Journal of the American Academy of Arts and Sciences*, 127(4), 89-118.
- Resnick, L. & Klopfer, L. (Eds.). (1989). Toward the thinking curriculum: An overview. In L.B. Resnick & L.E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Rodriguez, M. C. (2004). The role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education*, 17(1), 1-24.
- Ross, J.A., Rolheiser, C. & Hogaboam-Gray, A. (2002). Influences on student cognition about evaluation. *Assessment in Education*, 9(1), 81-95.
- Ruff, W. & Shoho, A. (2005). Understanding instructional leadership through the mental models of three elementary school principals. *Educational Administrative Quarterly*, 41(3), 554-577.
- Rutherford, F. (1997). *Reflecting on Sputnik: Linking the past, present, and future of*

- educational reform*. Retrieved October 13, 2006 from
www.nas.edu/Sputnik/ruther1.htm
- Sadler, D.R. (1998). Formative assessment: Revisiting the territory, *Assessment in Education: Principles, Policy & Practice*, 5(1), 1-7.
- Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Schafer, W. (1993). Assessment literacy for teachers. *Theory into Practice*, 32(2), 118-126.
- Schunk, (1996). Goal and self-evaluative influences during children's cognitive skill learning, *American Educational Research Journal*, 33(2), 359-382.
- Scriven, M. (1967). The methodology of evaluation. In R. F. Stake (Ed.), *American Educational Research Association monograph series: No. 1. Curriculum evaluation*. Chicago; Rand McNally.
- Senk, S., Beckmann, C., & Thompson, D. (1997). Assessment and grading in high school mathematics classrooms. *Journal for Research in Mathematics Education*, 28(2), 187-215.
- Sergiovanni, T. (1991) *The principalship: A reflective practice perspective* (2nd ed.). Needham Heights, MA: Allyn and Bacon.
- Shellard, E. (2005). How assessment data can improve instruction. *Principal*, 84(3), 30-32.
- Shepard, L. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, 66-70.
- Shepard, L. (2000a). The role of assessment in a learning culture. *Educational*

- Researcher*, 29(7), 4-14.
- Shepard, L. (2000b). The role of classroom assessment in teaching and learning. (CSE Technical Report, 517). Los Angeles: Center for the Study of Evaluation
- Shepard, L., Flexer, R., Hiebert, E., Marion, S., Mayfield, V., & Weston, T. (1996). Effects of introducing classroom performance assessments on student learning. *Educational Measurement: Issues and Practice*, 15, 7-18.
- Sheppard, B. (1996). Exploring the transformational nature of instructional leadership. *The Alberta Journal of Educational Research*, 42(4), 325-344.
- Sirin, S. R. (2005). Socioeconomic status and student achievement. A meta-analytic review of the research. *Review of Educational Research*, 75(3), 417-453.
- Sloane, F. & Kelly, A. (2003). Issues in high-stakes testing programs. *Theory into Practice*, 42(1), 12-17.
- Smith, W. & Andrews, R. (1989). *Instructional leadership: How principals make a difference*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Soles, D. (2001). *Sharing scoring guides*. Paper presented at the Annual Meeting of the Conference on College Composition and Communication in Denver, Colorado.
- Solmon, L. & Podgursky, M. (2000). *The pros and cons of performance-based compensation* (Opinion Paper No. 120). Santa Monica, CA: Milken Family Foundation.
- Starch, D. & Elliot, E. C. (1913). The reliability of grading high-school work in mathematics. *School Review*, 21, 254-259.
- Stiggins, R. J. (2005). Assessment for learning: Building a culture of confident learners.

- In DuFour, R., Eaker, R., & DuFour, R. (Eds.), *On common ground: The power of professional learning communities* (pp. 65-83). Bloomington, Indiana: National Educational Service.
- Stiggins, R. J., et al. (2005, September). *Assessment for learning defined*. Symposium conducted at the meeting of the ETS/Assessment Training Institute's International Conference: Promoting Sound Assessment in Every Classroom, Portland, Oregon.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86(1), 22-27.
- Stiggins, R. J. (2002). *Assessment crisis: The absence of assessment FOR learning*. Retrieved February 21, 2006, from <http://www.pdkintl.org/kappan/k0206sti.htm>
- Stiggins, R. J. (2001). *Student involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Merrill Prentice-Hall.
- Stiggins, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*.
- Stiggins, R. J. (1997). *Classroom assessment*. Upper Saddle River, NJ: Merrill Prentice-Hall.
- Stiggins, R. J. & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement*, 22(4), p. 271-286.
- Stiggins, R. J. & Conklin, N. (1992). *Investigating the practices of classroom assessment*. New York: State University of New York Press, Albany.
- Stiggins, R. J., Frisbie, R. J., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice* 8(2), 5-14.

- Stiggins, R. J., Griswold, M. M., & Wikelund, K. R. (1989). Measuring thinking skills through classroom assessment. *Journal of Educational Measurement*, 26, 233-246.
- Stiggins, R. J., Rubel, E., & Quellmalz, E. (1986). *Measuring thinking skills in the classroom*. Washington, DC: NEA Professional Library.
- System for Adult Basic Education Support (SABES) (2006). *Glossary of useful terms*. Retrieved April 11, 2007 from <http://www.sabes.org/assessment/glossary.htm>
- Thattai, D. (2001). *A history of public education in the United States*. Retrieved October 17, 2006 from <http://www.servintfree.net/~aidmn-ejournal/publications/2001-11/PublicEducationInTheUnitedStates>
- Valentine, J. & Bowman, M. (1988). Audit of principal effectiveness: A method for self-improvement. *NASSP Bulletin*, 72(508), 18-26.
- Valentine, J., Clark, D., Hackmann, D. & Petzko, V. (2004). *Leadership for highly successful middle level schools, Volume II: A national study of leadership in middle level schools*. Reston, VA: National Association of Secondary School Principals.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press
- Whiting, B., Van Burch, J. & Render, G. (1995). *Mastery learning in the classroom*. Paper presented at the Annual Meeting of the AERA. San Francisco, 1995 available from ERIC ED 382688.
- Wiggins, G. (2002). *Grant Wiggins on assessment*. Edutopia: The George Lucas Foundation. Retrieved December 17, 2006 from

http://www.edutopia.org/php/interview.php?id=Art_935&key=005#paragraph3

Wiliam, D. (2000). The meanings and consequences of educational assessments. *Critical Quarterly*, 42(1), 105-127.

Wiliam, D. & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions as assessment? *British Educational Research Journal*, 22(5), 1-11.

Wiliam, D., Lee, C., Harrison, C. & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education*, 11(1), 49-65.

Wolf, M., Crosson, A., & Resnick, L. (2006). *Accountable talk in reading comprehension instruction*. CSE Technical Report 670, Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and student testing.

Appendix A
Principal Cover Letter



Middle Level Leadership Center

University of Missouri-Columbia • 211 Hill Hall • Columbia, MO 65211

Director: Jerry Valentine (573) 882-0944

Assistants: Bernard Solomon (573) 882-0947; Kris Matthews (573) 882-0947; Greg Mees (573) 882-0947

Fax: 884-7922 • Email: ValentineJ@missouri.edu



Dear Middle Level Principal:

I am writing to seek your support of this comprehensive statewide study of Middle Level Education being conducted at the Middle Level Leadership Center at the University of Missouri. What I am asking of you is simple. Please ask your secretary to read the enclosed Secretary letter and complete the brief tasks listed in the letter. The envelope for your school secretary has the directions and materials necessary to distribute and collect the surveys. Essentially, we are asking each of the 10,000 plus Missouri middle level teachers to complete one survey, requiring 15-20 minutes of time. And we are asking 343 secretaries to handle the distribution and collection of the surveys. Please support that effort with a few kind words of encouragement to your teachers and your secretary.

To garner the support of the secretaries for this study and to express appreciation to those who go an extra mile to obtain as many returns as possible, I will give a \$50 gift card to the secretaries of the twenty schools with the highest percentage of returned completed teacher surveys. To make that distribution as fair as possible, I have divided the 343 middle level schools in the state into quartiles by enrollment. The five secretaries from each enrollment quartile with the highest percentage of returned completed teacher surveys within forty days from the original MU postmark will receive the gift cards.

As you can see, I consider this to be a very important study and am thus willing to recognize the efforts of the secretaries who help us make the study a success. The information gained from this study is extremely important to middle level education. This comprehensive study of all Missouri middle level schools is the largest of its kind ever conducted. From this study, more will be learned about specific successful practices in middle level education than from any previous study of middle level education. For Missouri, it will provide the opportunity to make policy recommendations to the Commissioner and other state policy makers in support of best practices for educating young adolescents. Also, it will provide the opportunity for each of the 343 middle level principals in the state to compare the aggregated statewide findings with practices used in individual schools. However, please be assured that at no time in this study will the names of schools or individual respondents be reported. All individual responses will be rendered anonymous and all data about schools will be aggregated and will be confidential.

What I am asking of you, therefore, is both important and straightforward. Please:

- a) Complete the brief Principal Survey (about 5-10 minutes)
- b) Sign the Principal Consent Form,
- c) Enclose both the Principal Consent Form and the completed Principal Survey in the envelope provided,
- d) Seal the envelope,

- e) Give the envelope to your school secretary so it can be returned to the Middle Level Leadership Center at MU,
- f) Encourage your teachers to complete their survey, and
- g) Encourage your secretary to follow through promptly with the process of collecting and returning surveys.

Please feel free to review any of the enclosed materials. The survey questions focus on a variety of key factors that affect student success, including programs and assessment practices, school culture and climate, principal and teacher leadership, and teacher commitment and efficacy. None of the materials ask sensitive questions but all are vital to build the necessary comprehensive picture of school effectiveness that allows us to study those variables that most directly affect student success.

Thank you for your time and for all of the energy and commitment you provide as a middle school leader. Please contact me if you have any questions about the study or the tasks I am asking you to complete. Look for the results of the study on the MLLC website so you can contrast your school with the statewide findings. I will also disseminate findings directly back to all participating Missouri schools through an email attachment.

Responses from your school and the other middle level schools across Missouri are vital to this effort as we work together to improve middle level education in our state and beyond. Thank you for your support.

Have a great school year!

Sincerely,

Jerry Valentine
Professor and Director, Middle Level Leadership Center
211 Hill Hall
University of Missouri
Columbia, MO 65211
(573) 882-0944
ValentineJ@missouri.edu
www.MLLC.org

PLEASE SUPPORT THIS STUDY IN YOUR SCHOOL BY:

- **GIVING THE PACKET OF MATERIALS TO YOUR SECRETARY**
- **ASKING THE SECRETARY TO FOLLOW THROUGH WITH THE APPROPRIATE TASKS**
- **ENCOURAGING YOUR TEACHERS TO TAKE THE FEW MINUTES NECESSARY TO RESPOND TO THEIR RESPECTIVE SURVEYS.**

THANK YOU!

Appendix B

Principal Demographic Survey

MISSOURI STUDY OF MIDDLE LEVEL EDUCATION

Principal Survey **School Name:** _____

Please take a moment to respond to the following demographic items. Please seal your responses in the envelope and give it to your secretary for mailing to my office at MU. Thank you for taking the time to complete this brief survey and supporting this important study in your school. What we learn will be extremely valuable to middle schools across the state and nation. The items are spaced for quick read and response. It will take you about 5-10 minutes to complete this survey. Thank you!

Jerry

Questions		Responses
1.	What is the current enrollment in your school?	
2.	How many full-time-equivalent teachers (staffing units for individuals who teach) do you have in your building (half-time teacher is .5, full-time teacher is 1.0, etc.)?	
3.	What grades are included in your school (e.g. 6-7-8, 7-8, etc.)?	
4.	What percent of your students are eligible for free or reduced lunch?	
5.	What do you estimate is the percent of average daily attendance for your student body?	
6.	What is your gender?	
7.	With what ethnic group do you identify yourself?	
8.	What is your age?	
9.	How many years were you a classroom teacher, counselor, or other non-administrator regardless of grade level?	
10.	How many years were you a classroom teacher, counselor, or other non-administrator at the middle level?	
11.	How many years were you an assistant principal, regardless of grade level?	
12.	How many years were you an assistant principal at the middle level?	
13.	How many years have you served as a principal, regardless of grade level, including this school year?	
14.	How many years have you served as a middle level principal, including this school year?	
15.	How many years have you served as the principal of this school?	

Continue on the Back -- Thank You.

16.	Overall, the culture of our school is positive, caring and collaborative.	1	2	3	4	5	6
17.	Overall, our school climate is positive, trusting, and respectful.	1	2	3	4	5	6
18.	Overall, our school is a healthy work environment.	1	2	3	4	5	6
19.	Overall, the teachers in our school trust their fellow teachers.	1	2	3	4	5	6
20.	Overall, the teachers in our school trust the clients (students and parents) they serve.	1	2	3	4	5	6
21.	Overall, the teachers in our school trust me (their principal).	1	2	3	4	5	6
22.	Overall, the teachers in our school trust the district-level leadership (district administrators and school board members).	1	2	3	4	5	6
23.	Overall, the teachers in our school are committed to the success of our school.	1	2	3	4	5	6
24.	Overall, the teachers in our school are committed to the personal and academic success of each and every student.	1	2	3	4	5	6
25.	Overall, our school faculty believes they can make a difference in the lives of our students.	1	2	3	4	5	6
26.	Overall, our school faculty effectively uses instructional strategies that enhance learning for each and every student.	1	2	3	4	5	6
27.	Overall, our school faculty effectively uses assessment strategies and data to improve instruction and achievement for each and every student.	1	2	3	4	5	6
28.	Overall, the teacher leaders in our school effectively lead other faculty in our schoolwide efforts toward excellence in teaching and learning.	1	2	3	4	5	6
29.	Overall, I provide effective leadership toward excellence in teaching and learning.	1	2	3	4	5	6
30.	Overall, I provide effective leadership for our school.	1	2	3	4	5	6
31.	Overall, the "leadership team" (principal, assistants, counselors, chairs, team leaders, etc.) provides effective leadership for our school.	1	2	3	4	5	6

Thank you for taking the time to complete this survey. Please seal this Survey and your Consent Form in the envelope provided and return it to the office secretary for mailing to the Middle Level Leadership Center.

Items in this survey not developed by the Middle Level Leadership Center were used with permission of the authors. Do not duplicate or use this survey without written permission from MLLC or the contributing authors. For information about the use of the survey or survey items, contact Jerry Valentine, Director, Middle Level Leadership Center.

Appendix C
Principal Consent Form

Principal Consent Form

As noted in your cover letter, the purpose of this study is to obtain information that will allow educators to better understand effective practices in middle level schools across the state. From that insight, policy and program recommendations will be made to state school leaders and an understanding of highly effective practices can be shared with middle level educators. The information your teachers provide about your school is extremely valuable in this effort. We are collecting data for this statewide study this winter and expect to analyze the data this spring.

Because our Center is a part of the University of Missouri, we must follow University guidelines when we conduct surveys. Therefore, even though there are no sensitive questions in this survey, we are obligated to inform you of your rights as a school leader whose school is participating in this study:

Your participation in this research study is voluntary and your responses, participation, or non-participation will not be used in any evaluative manner. You may choose not to participate for any reason; you may discontinue participation at any time, and you may refuse to answer any questions that may be uncomfortable for you. There are no foreseeable risks or discomforts because of participation. The responses from you and your teachers will be confidential and will go directly to the Middle Level Leadership Center. In compliance with University regulations, all data will be stored at the Middle Level Leadership Center in a locked storage cabinet for three years from the completion of the study. If you have any questions about the survey or its use, please contact Dr. Jerry Valentine, Center Director, at (573) 882-0944. If you have any concerns or questions regarding compliance with this statement, you may contact the University Institutional Review Board office for Human Subjects Research at (573) 882-9585.

Please (a) sign this Consent Form, (b) enclose both the Consent Form and the completed Survey in the envelope provided, (c) seal the envelope, (d) give the envelope to your school secretary so it can be returned to the Middle Level Leadership Center at MU, (e) encourage your teachers to complete their survey, and (f) encourage your secretary to follow through promptly with the process of collecting and returning surveys to MLLC. When we open your envelope at our Center, your consent form will be filed and the data about your school will remain confidential. Your teachers' responses will also be separated from their consent forms before they are entered into the data set for your school. Your school will be assigned a number code to maintain confidentiality of the school throughout the study. Thus, all individual teacher responses will be confidential and then anonymous; schoolwide data will be confidential; and, data will be analyzed and reported as aggregated group data. Neither your teachers' responses nor data about your school will be identified in any of the reports or recommendations from this statewide study.

Thank you for participating in this survey administered by the Middle Level Leadership Center. Your time and support are vitally important to middle level programs across the state.

Principal Signature

Date

Thank you,
Jerry Valentine
Director, Middle Level Leadership Center
211 Hill Hall, University of Missouri
Columbia, MO 65211

Appendix D
Secretary Letter



Middle Level Leadership Center

University of Missouri-Columbia • 211 Hill Hall • Columbia, MO 65211

Director: Jerry Valentine (573) 882-0944

Assistants: Bernard Solomon (573) 882-0947; Kris Matthews (573) 882-0947; Greg Mees (573) 882-0947

Fax: 884-7922 • Email: ValentineJ@missouri.edu



Dear School Secretary:

I am asking for your support of this comprehensive statewide study of Middle Level Education being conducted by the Middle Level Leadership Center at the University of Missouri. This study is the most comprehensive and significant study ever conducted in Missouri's middle level schools. For this study to be effective, I need your assistance. What I am asking of you is simple, but critical. Please complete the tasks listed below.

1. From state records we identified the approximate number of full-time classroom teachers in your school. Our numbers may be close, but not exact. Please distribute the enclosed surveys to your full-time classroom teachers in a random manner. Note that there are five different forms labeled "A" through "E" and that those five forms should be distributed evenly among the classroom teachers. Normally, this would be done by loading the surveys in a systematic manner in teachers' mailboxes. But if you prefer to use a teacher list, a faculty meeting, or some other form of distribution and collection, we will leave that to your judgment. We simply need to be sure that the surveys are distributed evenly to the full-time classroom teachers first and then any remaining surveys distributed to part-time teachers.
2. After two or three days, please offer a reminder to classroom teachers to complete and return the survey sealed in their appropriate envelope.
3. When you have about half of the teacher responses (or after one or two weeks), use the first of the two larger addressed, postage-paid envelopes I have provided to send the completed, sealed surveys to me at the Middle Level Leadership Center.
4. When you have all or nearly all of the classroom teachers' responses (or after two or three weeks), use the second larger addressed, postage-paid envelope to return the remainder of the completed, sealed surveys to the Middle Level Leadership Center.
5. If you have additional surveys returned to you after you have used the last large envelope, please call our office at (573) 882-0944 and I will be glad to send you an additional return, postage-paid envelope.
6. Please remember to ask your principal for the completed Principal Survey Form and include the principal's survey envelope when you return the teacher surveys.
7. If you need additional materials, including surveys or envelopes, please call my office at (573) 882-0944 and I will send the materials immediately.
8. Finally, please take a moment to complete the secretary contact form, providing us with a means to contact you in case you win one of the secretary gift cards described below.

To express appreciation to the individual designated to collect and return the survey packets to my office, I will give a **\$50 gift card** to the individual from each of the 20 schools that return the highest percentage of completed teacher responses. To make that distribution as fair as possible, I have divided Missouri's middle level schools into quartiles by enrollment. The five individuals from each enrollment quartile with the highest percentage of returned completed teacher responses will receive the gift cards. To be considered for this \$50 gift card, the returns must be received within 40 days from the original MU postmark.

Please contact me if you have any questions about the study or the tasks I am asking you to complete. As you can tell, responses from your schools and other middle level schools across Missouri are vital to this effort to support and improve middle level education in our state and beyond our state. Thank you for your time and effort to make this study a success.

Sincerely,

Jerry Valentine
Professor and Director
Middle Level Leadership Center
211 Hill Hall
University of Missouri
Columbia, MO 65211
(573) 882-0944
ValentineJ@missouri.edu
www.MLLC.org

Appendix E

Teacher Cover Letter (Short Version)



Middle Level Leadership Center

University of Missouri-Columbia • 211 Hill Hall • Columbia, MO 65211

Director: Jerry Valentine (573) 882-0944

Assistants: Bernard Solomon (573) 882-0947; Kris Matthews (573) 882-0947; Greg Mees (573) 882-0947

Fax: 884-7922 • Email: ValentineJ@missouri.edu



Teacher Instructions -- SHORT VERSION

There are two cover letters explaining this study and requesting your help. This is the “SHORT VERSION.” Please take a moment to read this first and if it does not convince you to complete this survey, please read the “LONG VERSION” on the last page. We need your help, but we also want to honor your time...thus this SHORT explanation.

WHAT ARE WE ASKING OF YOU?

1. Read the Directions (if you want to...or you can just begin the survey)
2. Sign the Teacher Consent Form...Please!
3. Complete the enclosed survey of Missouri Middle Level Teachers...Please!
4. Seal the completed items in the envelope provided and return them to your school office so the secretary can mail the sealed responses back to our office at the Middle Level Leadership Center at Mizzou...Please!

Your responses will be confidential and then anonymous (see below).

WHY IS THIS STUDY IMPORTANT TO MISSOURI AND TO YOU?

The fifteen minutes it will take you and fellow teachers to complete this survey will provide vital information about Missouri's middle school programs and practices. Your responses and those of fellow teachers across the state will:

1. Provide findings about key aspects of best practices in Missouri middle schools, including programs, culture, climate, leadership, commitment, efficacy, and assessment practices.
2. Provide findings that can shape policy changes in Missouri so we can better serve students.
3. Provide findings so you can contrast your programs and practices with those across Missouri.

WHO IS CONDUCTING THIS STUDY?

The Middle Level Leadership Center (MLLC) at the University of Missouri is conducting this study. Dr. Jerry Valentine (ValentineJ@missouri.edu), MU Professor and Director of the Center, is a nationally recognized researcher, author, and presenter in middle level education. MLLC is the only research and service Center in the nation devoted specifically to serving the needs of teachers and leaders in middle level schools. Learn more about the work of the Center at www.MLLC.org.

HOW IS THIS STUDY ORGANIZED?

This statewide study is so comprehensive that each teacher in your school will receive one of five randomly distributed surveys. The results will be returned directly to MLLC. When opened in our office, the Teacher Consent Form will be filed separately from the survey responses. Therefore, your responses will become anonymous and compiled with

other responses to create a confidential school-wide profile. At no time will those identities or responses be linked to individuals or schools.

**PLEASE PARTICIPATE...EDUCATION NEEDS YOUR
INSIGHT. THANKS!**

Appendix F

Teacher Cover Letter (Long Version)



Middle Level Leadership Center

University of Missouri-Columbia • 211 HILL Hall • Columbia, MO 65211

Director: Jerry Valentine (573) 882-0944

Assistants: Bernard Solomon (573) 882-0947; Kris Matthews (573) 882-0947; Greg Mees (573) 882-0947

Fax: 884-7922 • Email: ValentineJ@missouri.edu



Teacher Instructions – LONG VERSION

I know the tendency will be to toss this letter and the accompanying survey into the trash can...I have been a teacher and fully understand the urge, especially given the many tasks we are asked to complete each day.

However, before you toss these materials please take a moment to read this letter and consider this request. It will take you about five minutes to read this cover letter and the consent form, and about ten or fifteen minutes to complete the enclosed survey. Your time will be a significant contribution toward the development of a better understanding about the most effective practices for Missouri's middle level schools. Middle Level Schools are those that specifically serve students between grades five and nine, meaning with grade configurations such as 6-7-8, 5-6-7-8, 7-8, 8-9, and all combinations in between.

The importance of this study to Missouri...

Throughout our state middle level schools are coming under attack for low student achievement test scores. State officials repeatedly note that MAP test scores in elementary schools are improving significantly while MAP scores in middle schools are less impressive. Few policy makers understand the challenges of working with young adolescents who are intently focused on self and peers, attempting to adjust to new bodies, new emotions, and new social roles, while educators and policy makers continue to add "increased academic expectations" to their already stressful world. Simultaneously, societal changes in our nation are making early adolescence more challenging than ever. Little wonder young adolescents have difficulty developing the focus and commitment needed to achieve those academic standards. In this analysis of middle level education across Missouri, we are trying to understand the types of programs and environments that hold the greatest promise for supporting student success during these challenging years. Without effective middle level programs designed to address young adolescent development, we would have more dropouts and lower grades; with effective programs, we help students successfully navigate early adolescence so they can then be successful in high school. The dilemma is perplexing and we need your assistance to help us all understand what works and what does not work in our middle level schools.

The survey placed in your box by the secretary is one of five different survey forms randomly distributed to the teachers in your school and to the more than 10,000 teachers in each of the 343 middle level schools across Missouri. You will be able to respond to the items in your survey about as quickly as you can read them--so please take the 15-20 minutes to respond. With responses from all Missouri middle schools, we can carefully

study responses and offer to state policy makers the insight that will help them appreciate the hard and effective work that middle level educators do on a daily basis.

Responses will be confidential and anonymous...

You will note that your school's name is requested on the survey. To be able to analyze programs and practices across schools and compare those practices with other factors such as student achievement by school, I must know the name of your school so I can group responses by school. However, schools will not be identified by name for any part of this study. Names are needed to compute the types of analyses that generate findings and recommendations that can eventually influence middle school policies and practices across the state.

As mentioned previously, all responses to this survey will be confidential and will become anonymous as soon as the responses are separated from the Consent Form. All individual responses will be grouped by school for analyses. The names of all schools will remain confidential, with results reported by groups, not by individual schools, and certainly not by individual teachers. The results of this study will be made available through publications in professional journals, presentations at state and national professional meetings and via the Middle Level Leadership Center's website. Sharing the results will allow you and other educators to compare Missouri statewide findings with what you see occurring in your own school. I realize these explanations about confidentiality can be confusing, so I want to reassure you that when I say "share results and recommendations" I am talking about aggregated group data, not individual school or individual respondent data. Every response you make will be confidential and the names of each school will remain confidential.

The University of Missouri-Columbia is always sensitive to protecting the privacy and rights of respondents. So if you have any questions about this email or the survey, please contact me at ValentineJ@missouri.edu or by phone at (593) 882-0944 or contact our University Institutional Review Board office for Human Subjects Research at umcresearchcirb@missouri.edu or by phone at (573) 882-9585. All data from this study are confidential and stored on a secure, password protected hard drive here in our Middle Level Leadership Center office. The University requires that we maintain data from studies such as this for three years after the completion of the research project to ensure protection of your rights as a respondent, even when the data have become anonymous by separation of the Consent form. I assure you that we go to great efforts in our Center and at MU to be sure that responses are confidential and anonymous.

Please complete the survey today and if you cannot do it today, please try to do it within the next two or three days. Use the envelope provided to return your completed Survey and Consent Form to the school secretary, who will then mail all responses to our Center in Columbia.

A final perspective...

This is the largest, most comprehensive study of middle level schools ever undertaken in the U. S. More importantly to us in Missouri, it has the potential to provide valuable insight for all educators and policy makers. To make this study successful for all Missouri educators, I am asking that you make an effort to find the 15-20 minutes to respond.

Thank you so much for your consideration and time on this important task for our state.

Jerry Valentine, Professor and Director
Middle Level Leadership Center
211 Hill Hall
University of Missouri
Columbia, MO 65211
Phone (573) 882-0944

Appendix G

Middle Level Teacher's Assessment Practices Survey

MISSOURI STUDY OF MIDDLE LEVEL EDUCATION

Teacher Survey Form E

School Name: _____

Please use the following key to indicate the degree to which each statement applies to your classroom assessment practices. This survey may look lengthy, but the items have been spaced for easy reading to save you time. The typical time to complete this survey is about 15-20 minutes.

1=Never (0%) (25 to 50%)	2=Infrequently (1 to 25%)	3=Sporadically
4=Occasionally (50 to 75%) (100%)	5=Frequently (75 to 99%)	6=Always

(Circle, check, or darken the number that applies)

1.	I inform students about the expectations for their learning at the very beginning of the learning process.	1	2	3	4	5	6
2.	I tell my students, in language that they understand, the objectives they are going to be learning.	1	2	3	4	5	6
3.	I provide a detailed scoring guide to students before engaging them in the learning experience(s), so they understand the learning expectations for the content being studied.	1	2	3	4	5	6
4.	Before instruction, I provide previous work examples for students to analyze to guide their thinking about what a successful product looks like.	1	2	3	4	5	6
5.	I spend more time on planning my instruction than I do grading my students' work.	1	2	3	4	5	6
6.	I truly think about students' responses during class discussions to ensure that I am delivering the material in a manner that they understand.	1	2	3	4	5	6
7.	I provide a detailed scoring guide as a means for students to engage in peer and self assessment.	1	2	3	4	5	6
8.	I carefully plan the questions that I ask in class so responses to questions give me good insight about student understanding.	1	2	3	4	5	6
9.	When I ask a question, I provide adequate wait time and look for visual cues to ensure that all students have formulated an answer.	1	2	3	4	5	6
10.	When I have finished teaching a unit or concept and I have assessed my students' understanding, I move on if most of my students have mastered the content.	1	2	3	4	5	6

11.	When I teach a unit or concept and I have students who don't thoroughly understand the content, I re-teach the content using the same or similar instructional materials or strategies.	1	2	3	4	5	6
12.	When I teach a unit or concept and I have students who don't thoroughly understand the content, I design and implement different instructional strategies and experiences.	1	2	3	4	5	6
13.	When I teach a unit or concept and some of my students don't thoroughly understand the content, I provide individualized learning experiences so those who don't understand have more time to develop that understanding and those who do understand have time for enrichment experiences.	1	2	3	4	5	6
14.	I design my instruction based on the outcomes of previous classroom assessments.	1	2	3	4	5	6
15.	I teach, re-teach and continue to work with students until each student has truly mastered the learning concept.	1	2	3	4	5	6
16.	I primarily use a letter grade or percentage to inform students of how they performed on an assignment or test.	1	2	3	4	5	6
17.	I write detailed comments on students' work to identify areas of strength as well as areas for improvement.	1	2	3	4	5	6
18.	I use a detailed scoring guide to assess students' performance on a learning task.	1	2	3	4	5	6
19.	I work to ensure that my students clearly understand their level of performance on a classroom assessment.	1	2	3	4	5	6
20.	I work to ensure that my students clearly understand the desired goal of a classroom assessment and how to achieve the desired goal.	1	2	3	4	5	6
21.	I meet with students individually at least once a week to discuss their performance.	1	2	3	4	5	6
22.	My students' parents know the level of their child's performance in my class throughout the span of the course.	1	2	3	4	5	6
23.	I explain to parents the meaning of their child's assessment scores and how those scores relate to their child's growth as a learner.	1	2	3	4	5	6
24.	Students in my class are comfortable talking with each other and identifying strengths as well as areas to improve.	1	2	3	4	5	6

25.	My students write reflections about their growth and learning over a period of time for classroom objectives.	1	2	3	4	5	6
26.	I use conversation or written correspondence with my students to discuss, back and forth, the student's level of performance in my class.	1	2	3	4	5	6
27.	I teach self-assessment strategies so students can use them independently.	1	2	3	4	5	6
28.	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused me to implement changes in instructional delivery .	1	2	3	4	5	6
29.	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused me to implement changes in curriculum .	1	2	3	4	5	6
30.	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused me to implement changes in the programs my school supports.	1	2	3	4	5	6
31.	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused my school to implement changes in instructional delivery .	1	2	3	4	5	6
32.	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused my school to implement changes in curriculum .	1	2	3	4	5	6
33.	Formal assessment data, like the scores from an end of the unit test, common unit assessments, or district-mandated tests (e.g. writing assessments, SRI, DRA, math objective tests), have caused my school to implement changes in the programs my school supports.	1	2	3	4	5	6
34.	I meet with my principal, colleagues, and/or team to analyze formal assessment data (e.g. writing assessments, SRI, DRA, math objective tests), to plan for further instruction .	1	2	3	4	5	6

35.	I meet with my principal, colleagues, and/or team to analyze informal classroom assessment data (e.g. weekly assignments, daily quizzes, performance assessments) to plan for further instruction .	1	2	3	4	5	6
36.	I meet with my principal, colleagues, and/or team to analyze and discuss formal assessment data (e.g. writing assessments, SRI, DRA, math objective tests), to look for individual students that need extra help.	1	2	3	4	5	6
37.	I meet with my principal, colleagues, and/or team to analyze and discuss informal classroom assessment data (e.g. weekly assignments, daily quizzes, performance assessments) to look for individual students that need extra help.	1	2	3	4	5	6
38.	Overall, our school faculty effectively uses assessment strategies and data to improve instruction and achievement for our students.	1	2	3	4	5	6

39. What assessment strategies do you voluntarily use because you believe they are truly beneficial in raising student achievement on the MAP test?

40. What assessment strategies are mandated by your district or principal with the intention of raising student achievement on the MAP test?

41. Based upon your responses to questions 39 and 40, which forms of assessment do you believe are the **most** beneficial in raising student achievement on the MAP test? Why?

42. Based upon your responses to questions 39 and 40, which forms of assessment do you believe are the **least** beneficial in raising student achievement on the MAP test? Why?

Appendix H

Audit of Principal Effectiveness

MISSOURI STUDY OF MIDDLE LEVEL EDUCATION

Teacher Survey Form A

School Name: _____

Please use the following key to indicate the degree to which each statement applies to your principal. Respond specifically about your principal, not an assistant principal or other school leader. This survey may look lengthy, but the items have been spaced for easy reading to save you time. The typical time to complete this survey is about 15-20 minutes.

1=Strongly Disagree 2=Disagree 3=Somewhat Disagree 4=Somewhat Agree 5=Agree
6=Strongly Agree

(Circle, check, or darken the number that applies)

1.	The principal assists faculty in developing an understanding of, and support for, the beliefs and attitudes that form the basis of the educational value system of the school.	1	2	3	4	5	6
2.	The principal provides for the identification of, and the reaching of consensus on, the educational goals of the school.	1	2	3	4	5	6
3.	The principal has high, professional expectations and standards for self, faculty, and school.	1	2	3	4	5	6
4.	The principal helps the faculty develop high, professional expectations and standards for themselves and the school.	1	2	3	4	5	6
5.	The principal envisions future goals and directions for the school.	1	2	3	4	5	6
6.	The principal encourages changes in school programs that lead to a better school for the students.	1	2	3	4	5	6
7.	The principal communicates to teachers the directions the school's programs need to take for growth.	1	2	3	4	5	6
8.	The principal develops plans for the cooperation and involvement of the community, individuals, and agencies of the school.	1	2	3	4	5	6
9.	The principal utilizes resources from outside the school to assist in the study, development, implementation, and/or evaluation of the school.	1	2	3	4	5	6
10.	The principal provides for the gathering of information and feedback from individuals and agencies in the community.	1	2	3	4	5	6

11.	The principal provides for the dissemination of information to individuals and agencies in the community.	1	2	3	4	5	6
12.	The principal is supportive of, and operates within, the policies of the district.	1	2	3	4	5	6
13.	The principal maintains good rapport and a good working relationship with other administrators of the district.	1	2	3	4	5	6
14.	The principal invests time with the district office and other external agencies to obtain support and resources from the agencies.	1	2	3	4	5	6
15.	The principal strives to achieve autonomy for the school.	1	2	3	4	5	6
16.	The principal develops and implements school practices and policies that synthesize educational mandates, requirements and theories, e.g. legal requirements, social expectations, theoretical premises.	1	2	3	4	5	6
17.	The principal understands and analyzes the political aspects of education and effectively interacts with various communities, e.g. local, state, national, and/or various subcultures within the local community.	1	2	3	4	5	6
18.	The principal informs the staff of new developments and ideas in education.	1	2	3	4	5	6
19.	During the identification of needed change, the principal's style is more supportive and participative than directive and authoritative.	1	2	3	4	5	6
20.	During evaluation of change, the principal's style is more supportive and participative than directive and authoritative.	1	2	3	4	5	6
21.	The principal anticipates the effects of decisions.	1	2	3	4	5	6
22.	The principal fairly and effectively evaluates school personnel.	1	2	3	4	5	6
23.	The principal employs new staff who enhance the overall effectiveness of the school and complement the existing staff.	1	2	3	4	5	6
24.	Through discussion with teachers about concerns and problems that affect the school, the principal involves teachers in the decision-making process.	1	2	3	4	5	6
25.	The principal discusses school-related problems with teachers, seeking their opinions and feelings about the problem.	1	2	3	4	5	6
26.	The principal utilizes a systematic process for change that is known and understood by the faculty.	1	2	3	4	5	6

27.	The principal has the patience to wait to resolve a problem if the best solution to that problem is not yet readily evident.	1	2	3	4	5	6
28.	The principal is willing to admit to making an incorrect decision and corrects the decision if feasible.	1	2	3	4	5	6
29.	The principal is perceptive of teacher needs.	1	2	3	4	5	6
30.	The principal gives teachers the support they need to be effective.	1	2	3	4	5	6
31.	The principal diagnoses the causes of conflict and successfully mediates or arbitrates conflict situations.	1	2	3	4	5	6
32.	Teachers feel at ease in the presence of the principal.	1	2	3	4	5	6
33.	When deserving, teachers are complimented by the principal in a sincere and honest manner.	1	2	3	4	5	6
34.	The principal is receptive to suggestions.	1	2	3	4	5	6
35.	The principal is accessible when needed.	1	2	3	4	5	6
36.	The principal takes time to listen.	1	2	3	4	5	6
37.	Teachers feel free to share ideas and concerns about school with the principal.	1	2	3	4	5	6
38.	When teachers discuss a problem with the principal, the principal demonstrates an understanding and appreciation of how teachers feel about the problem.	1	2	3	4	5	6
39.	When talking to the principal, teachers have the feeling the principal is sincerely interested in what they are saying.	1	2	3	4	5	6
40.	Through effective management of the day-by-day operation of the school, the principal promotes among staff, parents, and community a feeling of confidence in the school.	1	2	3	4	5	6
41.	The principal finds the time to interact with students.	1	2	3	4	5	6
42.	Students feel free to initiate communication with the principal.	1	2	3	4	5	6
43.	Students in the school view the principal as a leader of school spirit.	1	2	3	4	5	6
44.	The principal encourages student leadership.	1	2	3	4	5	6
45.	The principal helps develop student responsibility.	1	2	3	4	5	6
46.	The principal is highly visible to the student body.	1	2	3	4	5	6

47.	The principal positively reinforces students.	1	2	3	4	5	6
48.	The principal enjoys working with students.	1	2	3	4	5	6
49.	The principal keeps teachers informed about those aspects of the school program of which they should be aware.	1	2	3	4	5	6
50.	When the principal provides teachers with the information about school operations, the information is clear and easily understood.	1	2	3	4	5	6
51.	When teachers are informed of administrative decisions, they are aware of what the principal expects of them as it relates to the decision.	1	2	3	4	5	6
52.	The principal is able to organize activities, tasks, and people.	1	2	3	4	5	6
53.	The principal develops appropriate rules and procedures.	1	2	3	4	5	6
54.	The principal uses systematic procedures for staff appraisal, e.g. retention, dismissal, promotion procedures.	1	2	3	4	5	6
55.	The principal establishes the overall tone for discipline in the school.	1	2	3	4	5	6
56.	The principal establishes a process by which students are made aware of school rules and policies.	1	2	3	4	5	6
57.	The principal communicates to teachers the reasons for administrative practices used in the school.	1	2	3	4	5	6
58.	The principal works with other leaders of the school in the implementation of a team approach to managing the school.	1	2	3	4	5	6
59.	The principal encourages faculty to be sensitive to the needs and values of other faculty in the school.	1	2	3	4	5	6
60.	The principal helps teachers clarify or explain their thoughts by discussing those thoughts with them.	1	2	3	4	5	6
61.	During meetings, the principal involves persons in the discussion who might otherwise not participate.	1	2	3	4	5	6
62.	The principal shares personal feelings and opinions about school issues with teachers.	1	2	3	4	5	6
63.	Humor used by the principal helps to improve the school environment by creating a more congenial working climate.	1	2	3	4	5	6

64.	Personal thoughts shared by the principal about school help teachers develop a sense of pride and loyalty as members of the school.	1	2	3	4	5	6
65.	The principal is knowledgeable of the general goals and objectives of the curricular areas.	1	2	3	4	5	6
66.	The principal is knowledgeable of the varied teaching strategies teachers might appropriately utilize during instruction.	1	2	3	4	5	6
67.	The principal possesses instructional observation skills that provide the basis for accurate assessment of the teaching process in the classroom.	1	2	3	4	5	6
68.	The principal actively and regularly participates in the observations and assessment of classroom instruction, including teaching strategies and student learning.	1	2	3	4	5	6
69.	The principal has effective techniques for helping ineffective teachers.	1	2	3	4	5	6
70.	The principal maintains an awareness and knowledge of recent research about the learning process.	1	2	3	4	5	6
71.	When criticizing poor practices, the principal provides suggestions for improvement.	1	2	3	4	5	6
72.	The principal is committed to instructional improvement.	1	2	3	4	5	6
73.	The principal promotes the development of educational goals and objectives that reflect societal needs and trends.	1	2	3	4	5	6
74.	The principal promotes the diagnosis of individual and group learning needs of students and application of appropriate instruction to meet those needs.	1	2	3	4	5	6
75.	The principal administers a school-wide curricular program based upon identification of content goals and objectives and the monitoring of student achievement toward those goals and objectives.	1	2	3	4	5	6
76.	The principal participates in instructional improvement activities such as program and curriculum planning and monitoring of student learning outcomes.	1	2	3	4	5	6
77.	The principal uses objective data such as test scores to make changes in curriculum and staffing.	1	2	3	4	5	6
78.	The principal has a systematic process for program review and change.	1	2	3	4	5	6

79.	The principal encourages articulation of the curricular program.	1	2	3	4	5	6
80.	Overall, our principal is an effective leader for our school.	1	2	3	4	5	6

Appendix I
Consent Form

Teacher Consent Form

As noted in the cover letter, the purpose of this survey is to obtain information that will allow educators to better understand effective practices in middle level schools across the state. From that insight, policy and program recommendations will be made to state school leaders and an understanding of highly effective practices can be shared with middle level educators. The information you provide about your school is extremely valuable in this effort. We are collecting data for this statewide study this winter and expect to analyze the data this spring.

Because our Center is a part of the University of Missouri, we must follow University guidelines when we conduct surveys. Therefore, even though there are no sensitive questions in this survey, we are obligated to inform you of your rights as a survey respondent:

Your participation in this research study is voluntary and your responses, participation, or non-participation will not be used in any evaluative manner. You may choose not to participate for any reason, you may discontinue participation at any time, and you may refuse to answer any question that might be uncomfortable for you. There are no foreseeable risks or discomforts because of participation. Your responses will be confidential and will go directly to the Middle Level Leadership Center when you complete the survey. In compliance with University research regulations, all data will be stored at the Middle Level Leadership Center in a locked storage cabinet for three years from the completion of the study. If you have any questions about the survey or its use, please contact Dr. Jerry Valentine, Center Director, at (573) 882-0944. If you have any concerns or questions regarding compliance with this statement you may contact the University Institutional Review Board office for Human Subjects Research at (573) 882-9585.

Please (a) sign this Consent Form, (b) enclose both the Consent Form and the completed Survey in the envelope provided, (c) seal the return envelope, and (d) give the sealed envelope to the school secretary for return to our Center at MU. When we open your envelope at our Center, we will separate your Consent Form from the survey and keep both on file, thus rendering your responses anonymous. We will then enter your anonymous responses into the data set for your school and then issue a number code for the school and maintain confidentiality of the school name throughout the study. Thus, all individual responses will be confidential and then anonymous; school wide data will be confidential; and, data will be analyzed and reported as aggregated group data. Neither your responses nor data about your school will be identified in any of the reports or recommendations from this statewide study.

Thank you for participating in this survey administered by the Middle Level Leadership Center. Your time and responses are vitally important to our efforts to support middle level programs across the state.

Respondent Signature

Date

Thank you,
Jerry Valentine
Director, Middle Level Leadership Center
211 Hill Hall, University of Missouri
Columbia, MO 65211

Appendix J

Campus Institutional Review Board Approval



Campus Institutional Review Board

University of Missouri-Columbia
483 McReynolds Hall
Columbia, MO 65211-1150

PHONE: (573) 882-9585

FAX: (573) 884-0663

Project Number: **1055838**

Project Title: Missouri Middle School Programs and School-wide Student Achievement

Approval Date: 12-11-2006

Expiration Date: 11-07-2007

Investigator(s): Goodman, Matthew Donald
Klinginsmith, Elmo Nyle
Matthews, Kristin Weiser
Mees, Gregory William
Solomon, Cameron Bernard
Valentine, Jerry Wayne

Level Granted: Expedited

Your Amendment was reviewed and we have determined that you are APPROVED to continue to conduct human subject research on the above-referenced project.

Federal regulations and Campus IRB policies require continuing review of research projects involving human subjects. Campus IRB approval will expire one (1) year from the date of approval unless otherwise indicated. Before the one (1) year expiration date, you must submit a Campus IRB Continuing Review Report to the Campus IRB. Any unexpected events are to be reported at that time. The Campus IRB reserves the right to inspect your records to ensure compliance with federal regulations at any point during your project period and three (3) years from the date of completion of your research.

Any additional changes to your study must be promptly reported and subsequently approved. If you have any questions, please contact the Campus IRB office at (573) 882-9585.

Vita

Kristin Ramey Weiser Matthews was born on April 17, 1966, in Denver, Colorado. She was raised in St. Joseph, Missouri, where she graduated from Central High School in 1984. She received her Bachelor of Science in Elementary Education with an emphasis area in Social Studies from Culver-Stockton College in 1990. She also received her Masters of Science in Elementary Administration from Northwest Missouri State University in Maryville, Missouri in 1998, an Educational Specialist in Administration from the University of Missouri – Columbia in 2007, and a Doctorate of Philosophy in Educational Leadership and Policy Analysis from the University of Missouri – Columbia in 2007.

Kristin taught for twelve years as an upper elementary teacher in the Savannah Public School District, St. Joseph Public School District and Columbia Public School District. During the pursuit of her doctorate, she worked as a Graduate Research Assistant at the Middle Level Leadership Center at the University of Missouri. She also worked as a Graduate Research Assistant for the St. Louis New Leaders Project at the University of Missouri, and as a Graduate Instructor for the Teacher Develop Program at the University of Missouri. She currently works for the Columbia Public School District as an assistant principal at Gentry Middle School.

Kristin is married to Chris Matthews of Edgerton, Missouri. They have two children, a daughter, Brooke Elizabeth, and a son, Zachary Christopher. The Matthews currently reside in Columbia, Missouri.