FORMATIVE ASSESSMENT IN INSTRUCTION AS PERCEIVED BY TEACHERS

BY

Neil O. Friesland

Submitted to the graduate degree program in Special Education and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Education.

_____

Chairperson

_____

_____

_____

_____

Committee members

Date Defended: _____

The Dissertation Committee for Neil O. Friesland
certifies that this is the approved version of the following dissertation:

FORMATIVE ASSESSMENT IN INSTRUCTION AS PERCEIVED BY TEACHERS

_____
Chairperson

Date approved:_____

ABSTRACT

The purpose of this study was to determine the perceptions of elementary-level teachers who utilize formative assessment scores to make curricular and pedagogical adjustments for students who perform poorly on those assessments. The research question for this study was, "Do intermediate elementary-level teachers who use formative assessments use the data to adjust teaching?" This study presents data collected from 3rd, 4th, 5th and 6th grade teachers about their perceptions of their impact on student performance, the efficacy of formative assessments, and administrator support.

The literature reports that formative assessment can be used to obtain information about students throughout the school year in relation to the curriculum addressed at a particular time. The literature goes on to suggest that the formative assessment process is a cycle that that takes place in the classroom and is part of the natural progression of teaching.

For this descriptive study, teachers across one midwestern state voluntarily completed a survey of teacher perceptions and use of formative assessments. Under the auspices of the state's testing agency, the author sent an initial invitation to participate to 5,900 teachers who use the formative assessment program. Once data collection was complete, the survey items were grouped using factor analysis and conclusions were made in relation to the teachers' responses.

The main goal of this study was to determine whether or not teachers change their methods or strategies if their students perform poorly on formative assessments. According to the data, teachers responded that they make instructional adjustments or changes when students perform poorly on the formative assessment. They also responded that they seek professional development opportunities to aid in instruction after formative assessments are given.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1

Introduction

This study focuses on determining the extent to which elementary school teachers in a midwestern state utilize formative assessment results to make pedagogical changes for students who perform poorly. The purpose of this chapter is to provide background on the strategies many states are implementing to meet the instructional outcome requirements of the *No Child Left Behind* (NCLB) legislation enacted in 2001. A major goal of NCLB is to increase the accountability of schools and teachers by requiring students to demonstrate Adequate Yearly Progress (AYP) [Sec. 111(b)(2)(J)].

Because teachers have high stakes in the AYP requirement, it is important to know what states are doing to assist teachers in preparing students to meet this requirement. AYP applies to all students, including students who qualify for special education services. *The Individual with Disabilities Education Act* (IDEA), 2004, states "Children with disabilities must be included in general state- and district-wide assessment programs, with appropriate accommodations and modifications in administration, if necessary" [Sec.300.138(a)].

Background

A popular strategy employed by states is to provide items from prior state assessments for the purpose of creating new formative assessments (Gallager & Worth, 2008). Teachers are encouraged to develop formative tests they can administer to their students. Formative assessments are intended to provide a snapshot of how students are performing at a particular time in the school year. This snapshot allows teachers to review the performance of individual students, and to determine how best to modify their

instruction to assist students whose performance indicates they need additional help in selected content areas. Failure of schools to meet AYP requirements is a consequence of the collective low performance of students. There can be multiple reasons for low performance, which may or may not include the ability of students and the instructional skills of teachers. The resources available to teachers, combined with the professional preparation of teachers, may also be a contributing factor (William, 2006).

Periodically it can be heard in schools and institutions of higher education that teachers "teach to the test" and do not utilize formative assessments as they are intended to be used. One example of teaching to the test is looking at past exam content and focusing instruction on those curricular areas assessed in state assessments. There are differing viewpoints on teaching to the test. Bushweller (1997) states that as long as the test measures what students are supposed to learn, it is fine to teach to the test. However, Brehm and Gates (2008) question the appropriateness of teaching to the test, and suggest that this practice does not foster learning, but instead becomes the year's entire focus. Herein lies a fundamental question: do students whose teachers teach to the test perform better on comprehensive exams, such as state assessments?

This study will focus on identifying teachers' perceptions of how they make changes in instruction on the basis of formative assessments. Another issue addressed in this study is whether classroom teachers believe they possess the skills needed to utilize formative assessment results in instructional planning. According to one blue print for state assessment, the formative process should "provide students with timely and useful feedback on how to improve their work" (Wiggins, 2002, p.1). For this to happen, the

teacher needs to know how to effectively use formative assessment results for the benefit of individual students.

Classroom Assessment

Gipps (1994) suggested that the role of traditional schooling is for the classroom teacher to know exactly what students are learning. Traditional schooling may be described as "teach, then test." Assessment is the natural process by which teachers find out what students know. Socratic questioning is a prime example of how good teachers use schema and paradigms to determine what students know in relation to what subject matter has been taught (Merrits & Walter, 2006).

School reform has focused extensively on testing and assessment (Sunderman, 2008). When NCLB was enacted in 2001, many educators saw this legislation as a way to hold schools accountable for student learning. Lagemann (2000) suggested that education has failed to develop a strong, self-regulating, professional community. This view of the education profession lends itself to the idea that schools across the country are renegade entities and tend to do whatever they want. However, one could argue that testing and accountability are some of the specific ways states control schools to determine whether learning takes place (Beare & Boyd, 1993).

Tests and assessments evolved to gauge whether a student has mastered a specific subject. Today, assessment tools have been developed to measure many forms of behavior and areas of human performance. In contemporary education, these tools assess performance at specific ages and grade levels.

The role of assessment in teaching is focused on how teachers can improve student learning (Popham, 2003). Teachers must be highly qualified in order to improve

student learning. This process of teacher training and development was integral to the reauthorization of IDEA legislation in 2004, as well as the enactment of NCLB in 2001, and has become a major approach to accountability. According to McMillian (2006), four components describe classroom assessment (See Figure 1.1):

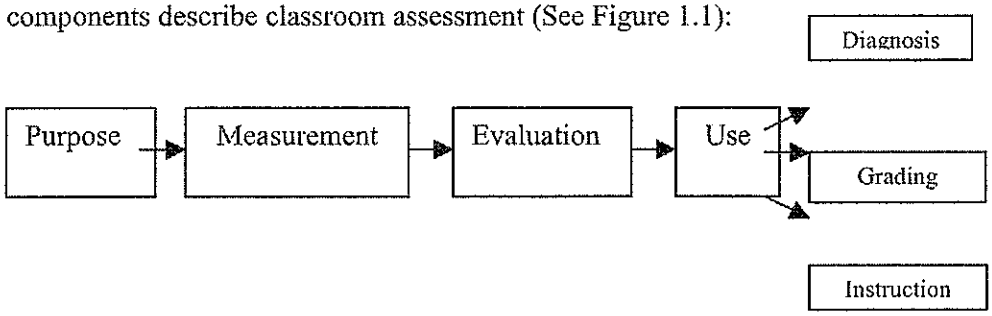


Figure 1.1. Flowchart of classroom assessment components.

Popham (2003) reinforced this process of classroom assessment with references to meeting the needs of students. McMillian (2006) says that "a clear vision is needed" when one looks at the purpose of any type of method to measure student learning (p. 9). Additionally, teachers should consider the act of measurement itself, as well as what techniques should be used to gather information. How teachers use assessment results is one of the most important ideas to consider in regard to assessment. McMillian (2006) stated three uses of classroom evaluation that can be beneficial: diagnosis, grading, and instruction. These three uses align with the focus of this study.

In addition to McMillan's uses of classroom assessment, there have been other roles identified, including the monitoring of learner mastery benchmarks and standards. Butler and McMunn (2006) stated that teachers must ask themselves these questions: (a) what are learning targets, and how do I unpack, or analyze, and reflect on those targets, and (b) what are my goals for my students?

Building Level Instruction →	Grade level Instruction →	Classroom instruction →	Individual Instruction

Figure 1.2. Formative assessment flowchart.

In the figure above, individual schools use assessment data to make instructional decisions at the building level; separate grade levels make instructional decisions based on data for each grade level; teachers make decisions for their classrooms and also for each individual student (Justice & Vukelich, 2008).

Statement of the Problem

NCLB (2001) has resulted in a level of instructional accountability that places increased expectations on teachers. Teachers are expected to ensure their students meet performance expectations specific to curriculum standards set by their State Educational Agency (SEA). This makes access to assessment data important in instructional decision-making, especially for students who are eligible for IDEA services. Once a child has been identified as qualifying for IDEA services, funding can then be used to help meet the student's needs. Several states have invested in making the assessment and instructional resources that are aligned with curriculum and/or assessment standards available to teachers. One such resource takes the form of formative assessments. Formative items are derived from previously administered state assessments. Teachers may access these items and organize them as tests for their students. The results become additional data teachers may use to make decisions regarding instruction and to enhance the performance of their students. The underlying assumption is that students will achieve the curriculum standards, and that their learning will be reflected in their performance on state assessments. The challenge, however, is that teachers are provided little or no direction

in how to use the results to improve instruction for their students who do not perform well on formative tests (Tienken & Wilson, 2001).

Little is known about how teachers use formative test results (Wilson & Barenthal, 2006). Formative assessments represent timely snapshots of student performance, and are intended as useful feedback to teachers. If used properly in making instructional decisions, formative assessments can increase success in the classroom and performance on state assessments (Wilson & Barenthal, 2006). The intent of this study is to determine how formative assessments, and the results of such assessments, are used in elementary classrooms in a midwestern state. The question addressed in this study is whether teachers change their teaching strategies based on formative test results in order to enhance the performance of students scoring low on the formative tests.

Purpose of the Study

The purpose of this study is to determine whether teachers at the third-, fourth-, fifth-, and sixth-grade levels, in districts that utilize formative data in reading and in math, plan and implement adjustments to instruction for individual students. Because a number of states do not provide a formative assessment program, an additional purpose of this study is to determine whether the use of formative assessments by teachers helps students perform better on state assessments after they make instructional adjustments.

Research Question

This study will focus on whether teachers utilize formative assessment data in improving instruction for learners. Specifically, do classroom teachers at the third-, fourth-, fifth-, and sixth-grade levels utilize formative assessment data in reading and math to plan and implement adjustments to curriculum and instruction for their pupils?

Definition of Terms

Accommodations: Accommodations are techniques and materials that make learning easier without changing the basic curriculum. They also help teachers provide access to the general education curriculum for students who are in need of accommodations.

Adequate Yearly Progress (AYP): AYP is required under NCLB (2001) and provides a way to measure school performance. Each state is required to establish challenging content and performance standards, and to implement assessments that measure students' performance against those standards.

Alternate assessment: This type of assessment is used for a small number of students who cannot participate in large-scale assessments, even with accommodations. However, states may include proficient scores from those assessments, and those scores will be capped at 2.0% of the total tested population. Additionally, states may continue to include proficient scores from such assessments in making AYP decisions and those scores will still be capped at 1.0% of the total tested population (U.S. Department of Education, 2006). Alternate assessments involve evaluative procedures that vary from the usual, customary, or standardized way a measurement is conducted, either by some special accommodation made for the assessee, or by alternate methods designed to measure the same variable.

Assessment: The process of documenting, in measurable terms, the knowledge, skills, attitudes and beliefs of students in relation to the curricular requirements of their teachers, as well as state and local educational agencies. Assessment can be formative or summative.

Computerized assessment (*CA*): An online, computer-delivered version of state assessments for the mandated NCLB state-testing program in mathematics and reading. The online testing formats include the same content as the existing paper and pencil test.

Cut score: Cut scores are selected points on the score scale of a test. The points are used to determine whether a particular test score is sufficient to reach a particular performance level. Cut score performance levels are narrative descriptions of the skills and knowledge that students must demonstrate in order to reach each performance level.

Formative assessment: A form of assessment designed to provide immediate evidence of student needs regarding a particular skill or content area, often conducted at the beginning of, or during, the school year.

Formative assessments: *TestBuilder* is an example of a computer-based vehicle for delivering formative assessments to local schools.

Individualized Education Program (*IEP*): A written plan and legal document that states a child's present level of functioning, specific areas in which the child needs special services, annual goals, short-term objectives, services to be provided, and the method of evaluation to be implemented for children 3 to 21 years of age who have been determined to be eligible for special education.

Summative assessment: A comprehensive assessment of the level of student learning at the end of a program. Summative data are often used to provide evidence of mastery and progress toward achieving the goals of a program, about which the summative assessment measures knowledge and/or performance.

Chapter 2

Review of Literature

Introduction

This study focuses on determining the extent to which teachers in grades 3, 4, 5, and 6 in a midwestern state utilize formative assessment results to make pedagogical changes for students who perform poorly on those assessments. Formative assessment is an instruction-related procedure teachers use to take periodic snapshots of student progress across the curriculum. This review of literature is a description of relevant and current research pertaining to the goals of this study. Topics covered in this chapter are: (a) the history of assessment, (b) school and classroom assessments, (c) formative assessment, (d) summative assessment, (e) formative assessment policies and practices in one midwestern state, and (f) changes in pedagogy based on formative assessment.

The need for assessment has existed ever since the need to determine whether a person is ready to move on in a particular area, or to accept a particular role in life. Gipps (1994) observed that the traditional role of assessment is to help teachers better instruct poorly performing students. While traditional schooling may look like "teach-then-test," Gipps (1994) would argue that schooling should feature a teach-then-reteach approach, as necessary. Specifically, assessment is the natural process by which teachers find out what students know (Popham, 2010). Socratic questioning is an example of how good teachers use schema and paradigm to determine what students know in relation to what has been taught (Merrits & Walter, 2006). Schema, for teachers, is a mental view of teaching, philosophy, and assessment. This mental view is developed over time from others in the teaching field. Kuhn (1962) defined "paradigm" as a set of practices that define a

scientific discipline during a particular period of time. These sets of practices are what drive teachers in pedagogy and assessment.

School reform has focused on the need for testing and assessment to reflect the success or failure of new policies and practices. From early normal schools to modern institutions of higher education, teacher training was built on science and the discipline of looking for truth (Yung, 2006). This truth, ultimately, was reflected in ways to measure what students learn. Over the years, those efforts led to the widespread practice of systematic "checks for understanding." These checks have become, and are manifested today, as state assessments (Lagemann, 2000).

When NCLB was enacted in 2001, many educators saw this legislation as a way to attempt to hold schools accountable for student learning. Lagemann (2000) suggested that education had failed to develop a strong, self-regulating professional community; thus, accountability was externally imposed by the federal government and is the lynchpin of NCLB.

History of Assessment

The earliest known written tests were developed in China around 2200 B.C. The earliest standardized tests were used in China around 537 B.C. (Miyazaki & Schirokauer, 1981). These tests were developed to determine whether government officials could perform and bring honor to their particular offices. Unlike aristocrats, government officials were appointed, rather than born, into a position. Special areas such as music and art were subject to particular scrutiny, due to the high honor of those positions (Hashaway, 1998). This type of assessment operated for over 2000 years. Hashaway (1998) goes on to say that in the United States tests were used to determine merit for

positions and advancement until the 1850s. At this point in history, tests did not have reliability data or validity to guide administration. This type of testing ended near the 1850s because of a general lack of accountability and standardization, as well as increased immigration of limited English speakers to the United States (Hashaway, 1998).

In the nineteenth-century, schools began to test their students to see if they had mastered what they were taught. Students who did not pass the tests were literally left behind, and were required to repeat what they had not mastered (Boaler, 2000). Teachers in the nineteenth century were often required to pass a test of their knowledge, and could be interviewed by members of the local school board to ensure they measured up to community, secular, and most often, religious standards, and to ensure the teachers were "fit" to serve in that role. Once a candidate was hired for a teaching position, he or she faced no more tests of proficiency or pedagogical delivery (Fiske, 1987). Ravitich (2002) stated that the feeling in schools at that time was that if a student failed to learn, it was the fault of the student.

Changes implemented at the turn of the 20th century increased student success as well as the standing of the United States in the global market. Shepard (2000) stated that the social efficiency movement of the 1900s sought to use science to solve problems of industrialization and urbanization. However, this movement also sought to improve education in order to address the larger issues of industrialization and urbanization.

Modern assessments were born of necessity to gauge student mastery of specific subjects (Parsons, Hinson & Sardo-Brown, 2001). Because there were no standards or standardization processes in place in at the turn of the 20th century, assessments may have

been churned out haphazardly and randomly (Watkins & Campbell, 2000). Today many assessment tools have been developed to measure a wide range of contracts, from IQ to attitudes about one's career choice. In the contemporary education setting, assessment tools are developed to show specific performances at specific ages and grade levels. These tools have been developed with stringent validity and reliability measures in place. Stiggins and Conklin (1992) stated that, over time, assessment tools became better at what they measured because of the nature of school reform and improvement.

From the mid 1830s to 1852, Horace Mann became a revolutionary educator through his advocacy for the standardization of schooling, particularly school evaluation and measurement (Hashaway, 1998). Mann measured the progress of students in Boston schools and provided an accountability tool for school boards. In this way, school boards could hold administrators accountable for student learning. Roles and responsibilities that focused on student learning and achievement, for administrators and teachers alike, came to light as various educational reforms gained a foothold in the United States' educational system. Historically, assessment practices shifted the responsibility for learning slightly away from the learner and somewhat toward the school and its leaders. No longer was the responsibility for failure on the student alone; responsibility was placed on administrators and subsequently on the classroom teacher (Hashaway, 1998).

Assessment Practices in Schools and Classrooms

Sternberg and Williams (1998) observed that students in all disciplines, and at all levels, confront exactly the same problem: they are being taught by methods that fit poorly with patterns of abilities. Sternberg and Williams (1998) went on to ask whether teachers should teach to student strengths rather than weaknesses. If so, teachers and

students need to understand the student's strengths and weaknesses, and then direct instruction to the strengths in order to minimize the weaknesses. This is related to classroom assessment because teaching to the test is one of those ideas teachers use to compensate for perceived deficits in students. Firestone, Monfils and Schorr (2004) indicated that teachers might use the approach of teaching to the test as a way of predicting in what areas the students will perform poorly on actual tests.

Weimer (2002) stated assessment can be a process of gathering information from multiple sources, including student self-assessment, to provide a snapshot of what students know and can do. This implies that assessment may be as simple as the teachers' and students' subjective judgment of performance, or as complex as a comprehensive, standardized test of several content areas. Wormeli (2006) pointed out that assessment affects a variety of decisions, including grades, advancement, placement, instructional needs, and curriculum. Teachers must be diligent in fair and useful grading practices. This perspective of assessment as the gathering of information allows teachers to make adjustments to teaching methodology as well as assign grades for course content.

Testing increases the pressure placed on teachers and administrators to devote considerable time to prepare students to do well on tests. Damage to students and schools can be an unintended product of this increased pressure (Thomas, 2005). As a consequence, narrowly focused tests that emphasize recall have led to a similar narrowing of curriculum and emphasis on rote memorization of facts, with little opportunity to practice higher-order thinking skills. Such a practice essentially involves teaching to the test (Thomas, 2005). When teachers teach to traditional tests by providing daily skill instruction in formats that closely resemble tests, their instructional practices

are both ineffective and potentially detrimental, due to their reliance on outmoded theories of learning and instruction (Thomas, 2005). According to Firestone, Monfils and Schorr (2004), teaching to the test has been linked to intense preparation for an exam, in the form of end-of-the-year summative assessments. Firestone, Monfils and Schorr (2004) went on to say teachers know more about teaching to the test than they do about what items might be included on an assessment.

According to Butler and McMunn (2006), like all assessments, high stakes tests pursuant to the requirements of NCLB (2001) are used to collect information about students and make comparisons. Unlike formative assessment, high stakes tests are used to gauge knowledge of students without regard to remediation. One of the purposes of high stakes tests is for teachers, administrators, parents, and students to discern comparative similarities and differences between and among students (Madaus, Russell, & Higgins, 2009). Madaus, Russell and Higgins (2009) observed that, in preparing students for high stakes testing, teaching to the test becomes an incentive to help students succeed on summative assessment.

Both schools and parents depend on assessment. Teachers and administrators use assessment to perform individualized diagnostic routines for students, as well as to monitor their advancement in the curriculum before, during, and/or after the school year. This process is formulated to assess, provide feedback, and ultimately motivate students to perform. In essence, teachers and administrators use assessment to make informed and educated estimates about students' performance in relation to mastery of goals and benchmarks of the curriculum (Butler & McMunn, 2006). Parents use assessment to assess strengths and weaknesses in educational programs, as well as to determine and

build upon the individual interests and strengths of their children. Parents also use assessment as NCLB intended—that is, to hold schools accountable to standards set by state and local educational agencies. Like teachers and administrators, parents also use assessment to make informed and educated decisions about their children's education.

Butler and McMunn (2006) identified several key features of assessment in the classroom. First, classroom assessment should be authentic (see Wiggins, 2002). This type of assessment should provide relevance in regard to task and environment. Authenticity can increase student ownership of the assessment process by showing students the real-life application of what is being measured. An example of authentic assessment could involve assessing the accuracy of making change at a student-run school store.

Classroom assessment should also be characterized as quality assessment (Angelo & Cross, 1993). Quality, in this case, involves a clear and concise purpose regarding what data the teacher is trying to obtain by way of assessment. Teachers can directly measure performance in relation to expectations and provide instructive feedback to students.

A third feature is that tests should be a routine part of on-going classroom assessment and instruction. Typical tests measure content, knowledge, and the skill-performance of students. In recent years, teachers have devised classroom tests to measure students' progress toward general education curriculum standards (Tindal & Haladyna, 2002).

Finally, in general, classroom assessments serve a diagnostic-prescriptive purpose, but they can also be used to inform policy and drive accountability provisions of policy. According to the United States Department of Education (2006), the NCLB

accountability provision requires states to describe how they will close the achievement gap and make sure all students—including those experiencing economic disadvantage, those among ethnic minority groups, those with disabilities, and those whose primary language is not English—achieve academic proficiency. Schools must report data to the state departments so the latter can produce annual state and school district report cards that inform parents, community, and policy makers about the status of student proficiency and school progress. Schools that do not make adequate progress must provide supplemental services, such as free tutoring or after-school assistance, take corrective actions, and, if still not making adequate yearly progress (AYP) after five years, make dramatic changes to the way the school is run (Walsh, Kemerer, & Maniotis, 2005).

The role of formative assessment.

The role of formative assessment is to assess students periodically in order to determine the extent to which they have learned what has been taught. Dell'Olio and Donk (2007) stated that the most important product of formative assessment is immediate and frequent feedback. This is important for teachers, because they can focus specifically on re-teaching skills students did not master, and give students experiencing difficulties more information. This is important for the student, not only because he or she will be left behind, but also because re-teaching can focus on learning, rather than merely finding the right answers.

Formative assessment is an approach to assessment that enables "snapshots" of student progress within the curriculum; "Formative assessment has become a buzzword in the field of educational measurement" (McMillian, 2006, p. 117). Student achievement and success are enhanced, in regard to pressing toward benchmarks, when effective

formative assessment strategies are used (McMillian, 2006). Formative assessment can be used to take a quick look at how the students are progressing toward those benchmarks. Quenemoen and Thurlow (2004) state teachers must put time, effort, and thought into improving how they determine what students are learning. This can improve all forms of assessment, in varying contexts and for varying purposes, including formative assessment. Formative assessment, when implemented as intended, operates as a cyclical and continual process involving student work and behavior, feedback to students, and instructional correctives. McMillian's (2006) Formative Assessment Cycle illustrates this process:

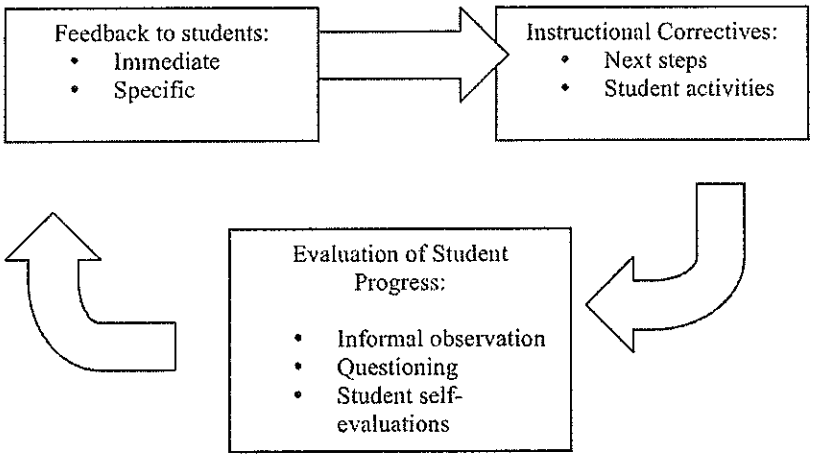


Figure 2.1. The formative assessment cycle.

Ainsworth and Viegut (2006) observed that formative assessment could provide immediate feedback to students regarding their current levels of understanding. When teachers have information on which to base educated decisions about their students, they can plan accurately and with necessary precision in relation to curriculum standards. If the teacher knows in advance what knowledge and skills are required for students to master a particular subject, the use of formative assessment, such as a sight word

recognition activity in reading, can help teachers set grade level mastery benchmarks (Ainsworth & Viegut, 2006).

Hall and Burke (2004) stated that a prominent use of formative assessment is to engage in permanent interaction with the curriculum, providing evidence that learning is occurring. Teachers can use formative assessment to examine student learning over time and in multiple contexts. Moreover, these authors regard formative assessment as a process that takes place in the classroom, beginning with eliciting a response from the student, revising instruction, and as a result, improving student learning.

Having students score and analyze their own formative assessments furthers the benefit to students (Ainsworth & Viegut, 2006). This process can help students set goals, as well as establish a buy-in for the assessment process as a whole. Student experience in formative assessment allows teachers to make pertinent adaptations and adjust for those students who may not be learning what the teacher is teaching. Formative assessment enables teachers and students alike to respond constructively to instructional data and increase student performance (Wormeli, 2006).

The role of summative assessment.

Unlike formative assessment, summative assessment involves a comprehensive evaluation of people and/or a program. It is used to evaluate the level of learning at the end of a program (Popham, 2003). Summative data are often used to provide evidence of mastery and/or progress toward achieving the goals of the program for which the summative assessment measures knowledge and/or performance. In schools, summative assessment is most often used to make judgments regarding student performance. These judgments are then represented as marks, scores, or statements such as a "B+", or "shows

improvement." Summative assessments can then be defined as a measure of performance at the end of a sequence of study (Miller, Imrie, & Cox, 1998).

Standardized assessment, pursuant to the requirements of NCLB (2001), is summative in nature. As such, its results provide data for local administrators, state education officials, and, ultimately, state and federal policy makers. However, given the timing of the release of test scores, and the summative nature of these assessments, teachers are unable to use the test scores to make adjustments in specific curriculum content or teaching methods for the class of assessed students. Realizing the potential benefit of formative assessments for teachers and students, many states provide schools the option to take part in formative assessment procedures periodically throughout the school year, as part of the state's (summative) standardized assessment programs. One such midwestern state served as the focal point for this study.

Formative Assessment Practices in One Midwestern State

In the Midwestern state where this study was conducted, formative assessment comprises a computer-based assessment delivery system, developed at one of the state's universities. This computerized formative assessment is a tool designed to enable teachers to monitor students' progress in reference to the state's standards-based curriculum.

This state's education department does not track the identities of those who take the formative assessment (D. Bowman, personal communication, January 7, 2010). However, the state's Board of Education provides the formative assessment tool as a suggestion for supplementing instruction, prior to the summative assessments at the end of a school year. Not all schools in this state use the formative assessment tool, nor are

they required to do so. Any student is eligible to sit for this computerized assessment as their "official" state assessment. However, students may complete the formative assessment, via the formative assessment tool, if their school participates in the statewide program.

Test Builder is a computer program developed by the state university for teachers to develop and implement formative assessments. To use Test Builder, a teacher subscribes to the program via on-line registration; he or she is able to build specific assessments to meet his or her needs. Two processes are available: an existing, pre-built test, or the option to build one's own test by selecting individual questions. Formative assessments provide feedback that can be used to modify teaching and learning activities.

Test Builder involves the following processes: (a) selecting and assigning formative tests, (b) modifying and creating formative tests, (c) administering formative assessments, (d) viewing and interpreting students' results after giving formative assessments, and (e) accessing the teacher formative testing system. Once formative assessments are given, teachers can then look to improve the ways they introduce and teach the curriculum.

Change in Teacher Pedagogy with Assessment

The term "pedagogy" means the manner in which something is taught (McLoughlin & Taji, 2005). Pedagogy can also be described in terms of a deep-rooted values system regarding how teachers approach the process of teaching. McLoughlin and Taji (2005) stated that teachers must be given time to think about the change in pedagogy that may come after formative assessments are given, because their values systems are developed over time and cannot be changed quickly. Several authors have indicated a

need for change in how teachers view adjustments in pedagogy. According to Appleton (2006), pedagogy is based on a view of learning held by the individual teacher. This view is shaped over time, and often is teacher-centered. Activities can be accepted or rejected according to how the teacher sees his or her learning environment.

Gordon and Bridglall (2006) observed that pedagogy is an interaction between students and teachers in relation to learning; thus, this relationship may be fluid rather than static. Formative assessment is a process that takes place at periodic intervals in the classroom, to provide pertinent information to guide the teacher to make appropriate adjustments to pedagogy.

The question may be posed, "Are teachers literate in assessment?" Stiggins (1999) used the term, "assessment literacy," as a way of showing the types of assessment skills teachers possess and utilize for the benefit of students. Stiggins (1999) also noted that many teachers did not have experience developing these skills in their teacher education programs. Olson (2002) reported that only 14 states required teachers to demonstrate competence in assessment to earn a teaching license, and only three states required that principals show expertise in assessment.

Shepard (2000) expressed the need for teachers to develop classroom assessment skills to increase student success, based on emerging research and discipline-based standards documents. Shepard (2000) went on to suggest that teachers need to make the following changes in their assessment practices: first, teachers must change their approach to teacher-student communication so that students can develop greater knowledge of, and responsibility for, learning goals; second, teachers must discover and develop students' prior knowledge and use that information in planning better instruction

to meet student needs; next, teachers must provide feedback to students that enhances the educational process, as opposed to information that focuses solely on deficits.

Building teachers' assessment skills and tools should be a focus of teacher education programs; new teachers should expect to share their knowledge of assessment with school administrators in screening interviews (Wilson, 2004). Because the stress of high stakes testing is so apparent in schools today, teachers may enter the profession at a disadvantage, due to lack of training in assessment. Butler and McMunn (2006) provided a phase system for implementation of assessment training and skills development. Phase one is the goal setting phase. This phase allows teachers to access and learn district benchmarks and standards. This helps teachers to develop standards and expectations for their students, as well as collaborative relationships with veteran teachers, in order to begin building expertise in assessment. Phase two includes the processes by which grading and reporting practices are examined. This phase focuses on principles of good classroom assessment, developing the ability to distinguish between formative and summative assessments, and learning how to interpret and analyze assessment scores. Phase three is the implementation of effective classroom assessment practices. In this phase, teachers regularly examine assessment practices and pedagogy, and continue to develop collaborative relationships with veteran teachers and the administration.

Butler and McMunn (2006) stated that teams of teachers can enhance assessment skills. Butler and McMunn (2006) have also shown that, through the process of collaborative learning and organizing teachers around learning teams, an educational institution can address and potentially improve the organization's capacity for change. Collaborative learning among teachers can improve capacity for change by allowing

teachers to develop skills that focus on each student's individual learning needs. Cunningham (2003) advocated for a learning-centered model that utilizes assessment tools matched to individual teaching styles. The move from teacher-centered models of learning to student-centered models has led to changes in teacher education programs across the United States (Appleton, 2006).

Sunderman (2008) showed teachers might need to go back to the basics with regard to assessment. According to these authors, schools need to determine a more efficient and manageable way to implement assessment for changes in pedagogy to occur. They also showed that effective change could come from re-thinking the assessment process as a whole to allow for change in pedagogy with effective data collection and interpretation.

Formative and summative assessments play an integral role in measuring student performance. Teachers use data from these assessments to make decisions regarding mastery of content, grading, and movement from one grade to another. Teachers also use formative assessments as a way to determine student performance at periodic intervals throughout the school year. Data from formative assessments provide a rationale for change in pedagogy, in order to reach students who are performing poorly. Chapter 3 will review the methodology and data collection processes of this study.

Chapter 3

Methodology

Study Purpose and Procedures

The purpose of this study was to determine the perceptions of intermediate elementary-level teachers who utilize formative assessment scores to make curricular and pedagogical adjustments for students who perform poorly on those assessments. The research question for this study was, "Do intermediate elementary teachers who administer formative assessments use data from these assessments to adjust their teaching?"

Participants

Participants were 164 intermediate elementary-level teachers in a midwestern state who voluntarily completed a survey regarding use of formative assessments. Under the auspices of the state's testing agency, the author sent an initial email (January 2008) to 5900 teachers who had used the formative assessment program. Of the 5900 email messages distributed, 566 were returned as "undeliverable." One hundred six emails had been sent to teachers who did not fit the sample parameters of this study. Thus, the total number of emails successfully delivered to teachers of grades 3, 4, 5, and 6 was 5228. When data collection started, 180 of the teachers responded via email that they would be willing to participate in the study. When data collection ended (late February 2008), 164 of the original 180 teachers had completed the survey. As of February 2008, 164 of the 5228 (3.1%) teachers who reported utilizing formative assessments had participated in the study in part or whole. Consequently, email links and reminders were sent only to those 180 teachers who had agreed to participate in this study.

To maximize participation rates for this study, the incentive of a pencil was promised to all 180 participants on the condition that they provide their school address so the pencil could be mailed.

Development of the Survey

To begin the process of developing this survey, the researcher consulted two administrators: an assistant superintendent and a building principal, each employed in a different local education agency (LEA). These administrators were chosen because of their extensive experience with state assessments at the district level, as well as their work experience in assessment. Conversations with these administrators involved the feasibility and appropriateness of survey formatting and content. For example, both administrators were asked about the appropriateness of questions related to the skills, instructional responsiveness, and efficacy of elementary teachers.

Pilot testing began in the fall of 2007 with 21 intermediate elementary-level teachers and administrators in these same two LEAs. Administrators were chosen to participate in the pilot study for two main reasons: first, because building administrators could provide face validity for the study; secondly, because building administrators at the elementary level are often former elementary classroom teachers; their perspectives might, therefore, reflect the general population of the participants in this study. Of the 115 surveys distributed, 21 were completed and returned. The return rate of 12.7% from the entire pilot study was lower than expected. A reminder was sent out two weeks after the pilot surveys were distributed. At one LEA, reminders were mailed to individual building administrators and were distributed accordingly. The Assistant Superintendent was also contacted and a reminder was sent to the participating teachers. This reminder

produced an additional two responses. At the other LEA, the reminders were taken to individual schools and placed in the teachers' mailboxes. This reminder produced an additional five responses. See Appendix B for an additional explanation of the results of the pilot study.

The study survey was formatted using a four-point Likert scale, which removes the neutral response and forces participants to agree or disagree with survey items (Lancaster, 2005). Statements were posed in such a way that participants' answers could be labeled as follows: 1=strongly disagree, 2=disagree, 3=agree, and 4=strongly agree. Groups of statements were used to determine each individual's perceived awareness of multiple areas of teaching. These areas were: a) effectiveness of teaching as reflected in the formative assessment process, b) administrator support in state assessments, c) changes in teaching strategies and methods with formative assessment score receipt, and d) professional development in knowledge and skills related to formative assessment score utilization.

Because the use of electronic surveys can be a way to reach potential respondents and possibly increase response rates, the Internet data collection site, http://www.surveymonkey.com, was used to host the survey and collect responses. Best and Krueger (2004) indicated that once participants are identified, an instrument can often be administered by email. The survey link was emailed four times to all participants within the data collection period. The initial email for the link to Surveymonkey.com was sent to 180 participants in December 2007. Another email was sent in early- and mid-January 2008. Finally, a reminder email was sent out in mid-February 2008. Kaplowitz, Hadlock, and Levine (2004) observed that using the web to collect research data achieves

a comparable response rate to mailing hardcopy questionnaires. However, emailing a link to the survey saved time and money in this study; thus, the response rate was only one reason for using the Internet to collect data. The benefits of using an online survey include speed of data collection, accuracy, and relatively low costs. Speed, cost, and efficiency are important for collecting reliable and accurate data in a short amount of time (Burke & James, 2006).

Data Analysis

Data analysis consisted of a principal component factor analysis of the individual survey statements to determine whether specific groupings of statements would emerge. Emergent clusters, regarding participants' use of formative assessments for adjusting their instruction, assisted interpretation of the data. In addition, simple correlations among items demonstrated preliminary relationships as perceived by participants. Finally, basic descriptive statistics pertaining to item responses and participants' demographic information are presented.

Chapter 4

Results and Discussion

The purpose of this study was to determine the perceptions of intermediate elementary-level teachers who utilize formative assessment scores to make curricular and pedagogical adjustments for students who perform poorly on those assessments. The research question for this study was, "Do intermediate elementary teachers who administer formative assessments use the data to adjust their teaching?" The survey used in this study consisted of 45 items; participant responses were based on a four-point Likert scale (1=strongly disagree and 4=strongly agree; see Appendix A). Participants were also asked to provide responses to five demographic items.

When the original request to participate in the study was sent out via email to teachers across the state, 180 teachers replied via email that they would participate. Thus, a convenience sample was drawn from a population of all teachers who had been alerted to the opportunity for participation. As of February 2008, 164 of the 5228 (3.1%) teachers had participated in the study by completing the survey hosted at http://www.surveymonkey.com.

Survey Characteristics

As mentioned in Chapter 3, an assistant superintendent and a building principal of a local education agency (LEA) and staff from the state's Department of Education were consulted before drafting the survey tool. A reliability analysis was completed using SPSS and yielded a Chronbach's alpha of .71. This suggests that there was adequate survey reliability (Moore, 2009). Pearson correlations were obtained between individual survey items. The definition of "higher correlation" for the pilot study was a correlation

of .600 or above; thus, from these data one can conclude the answers given by participants were not made in error, nor did the participants misunderstand the statements. Seven pairs of statements correlated highly according to this criterion. Table 4.1 displays the pairs of statements with high correlations.

Table 4.1

Highly Correlated Item Pairs

Items	Coefficients
Q1: I use the Computerized Assessment (CA) and my students score well every year. Q40: I regularly actively engage all of my students.	.60
Q4: I feel like I am a more effective teacher because I attend professional development in the areas of assessment, personal research, and individual and group instructional adjustment with the CA. Q5: My students are better prepared due to my personal preparation because I attend professional development in the area of CA.	.71
Q5: My students are better prepared due to my personal preparation because I attend professional development in the area of CA. Q30: I use a number of various teaching strategies in my class to prepare students for the CA.	.62
Q8: My students are better prepared for the state assessments because of the teaching strategies and methods I incorporate after using the CA. Q9. I reflect on my effectiveness as a teacher if I have students perform poorly on the CA.	.68
Q6: I use research-based strategies to remediate problems after formative assessments are given. Q28: I provide feedback to my students as to what areas they performed poorly on the state formative assessments.	.75
Q21:I don't make many changes in my teaching when students perform poorly on the CA. Q35: I have had professional development training in mathematics and utilization of formative assessment scores.	.77
Q43: I feel that my students are NOT prepared for the CA. Q37: I don't feel the need to attend additional professional development to teach my students more effectively; I need more resources and support.	.60

Scale analysis of the survey provided higher negative correlations for the following comparison (see Table 4.2).

Table 4.2.

Highly Negatively Correlated Item Pairs

Item	Coefficient
Q1: I use the CA every year and my students score well. Q2: I use the CA every year and my students perform poorly.	-.74
Q4: I feel like I am a more effective teacher because I attend professional development in the area of assessment, personal research, and individual and group instructional adjustment with the CA. Q21: I don't make many changes in my teaching when students perform poorly on the CA.	-.64
Q21: I don't make many changes in my teaching when students perform poorly on the CA. Q9: I reflect on my effectiveness as a teacher if I have students perform poorly on the CA.	-.70
Q8: My students are better prepared for the state assessments because of the teaching strategies and methods I incorporate after using the CA. Q18: When I use the CA, my students don't perform any better on the spring summative assessments.	-.64
Q18: When I use the CA, my students don't perform any better on the spring summative assessments. Q40: I regularly actively engage all of my students.	-.60
Q7: I use the CA as it is intended to be used, not as a practice for the spring summative assessments. Q16: I use the CA as a practice run for the summative assessment in the spring.	-.76

Demographic Findings

Demographic data were collected for respondents of this study (see Table 4.3).

Table 4.3

Descriptive statistics for Demographic Variables

Demographic Variable	Mean (in years)	Range	Standard Deviation
Age	43.23	22-64	10.095
Experience (in years)	16.70	1-36	9.277
Longevity (in current role)	12.06	1-36	8.606
Number of Students	28.36	18-100	31.732

The average age of respondents to the survey was 43.2 years, with a range of 22 years to 64 years. The average years of experience the teachers had in the classroom was

16.7 years, with a range of 1 year to 36 years. Teachers reported an average longevity in their current roles of 12.06 years, with a range of 1 year to 36 years. They reported 28.36 as their average number of students, though the standard deviation is indicative of substantial variation.

As shown in Table 4.4, a majority (59.1%) of the respondents work in rural school districts. Of the remaining teachers, 27.4% of the respondents teach in suburban school districts, and 12.8% teach in urban schools.

Table 4.4

Participants' Geographic Locations

Locale Type	Total n	Frequency	% of study	% of n
	13,616			
Rural (non-metropolitan)	5454	97	59	.007
Suburban (Metropolitan)		45	27.4	.003
Urban (Metropolitan)	8163	21	12.8	.001
Total*		163	99.3	.012

* 1 participant did not respond to item

The total number of elementary school teachers in the Midwestern state at the time of this study was 13,616 (W.Wallis, personal communication, January 15, 2010). This state divides the schools into two groups: a) metropolitan (urban and suburban) and b) non-metropolitan (rural). According to the state report, 60% (8163) of the elementary school teachers in this midwestern state teach in metropolitan school districts, whereas 40% (5454) teach in non-metropolitan districts. This study showed the majority of participants

taught in rural (non-metropolitan) school districts and are not indicative of the population of teachers in the state.

Item Responses and Patterns

The following sections will provide the results of this study on the basis of common statements and common factors determined via factor analysis. Table 4.5 provides the survey's item means and standard deviations.

Table 4.5.

Item Means and Standard Deviations

Item	Mean Response	Standard Deviation	Item	Mean Response	Standard Deviation
1	2.96	.675	24	2.63	.718
2	1.69	.643	25	3.05	.617
3	1.80	.643	26	1.99	.823
4	2.74	.857	27	2.90	.593
5	2.71	.824	28	3.19	.541
6	3.14	.635	29	2.29	.711
Item	Mean Response	Standard Deviation	Item	Mean Response	Standard Deviation
7	2.74	.725	30	3.37	.534
8	3.27	.545	31	2.03	.696
9	3.10	.734	32	2.37	.844
10	1.42	.543	33	3.13	.702
11	3.39	.613	34	2.53	.650
12	3.31	.707	35	2.60	.775
13	3.06	.791	36	3.21	.515
14	3.22	.608	37	2.45	.908
15	2.35	.624	38	1.76	.585
16	3.13	.642	39	2.36	.607
17	2.72	.768	40	3.19	.549

18	1.85	.540	41	1.52	.525
19	2.63	.706	42	1.40	.484
20	2.43	.753	43	1.82	.678
21	1.52	.570	44	2.46	.536
22	3.20	.577	45	2.80	.625
23	3.39	.488			

* Scale: 4=Strongly Agree 3= Agree 2=Disagree 1=Strongly Disagree

As reported in Table 4.5, the highest average agreement responses on the Likert scale were Q11, "I re-teach the big ideas" (3.39), and Q23, "I take time to reflect on poor performers" (3.39). The largest standard deviations for the survey respondents was Q37 at .908, "I don't feel that I need additional professional development," and Q4 at .857, "I feel like I am a more effective teacher because I attend professional development." The large standard deviations reported are important because these indicate the data are "spread out" over a large range of values; thus, teachers responded differently to these items.

Initial data analysis was conducted using bivariate Pearson correlations (Moore, 2009) to determine the degree to which all item pairs related to one another (Appendix C provides a complete correlations matrix of Pearson Coefficients). The following tables provide the correlations for some of the logically (not empirically) clustered common statements on the survey.

Table 4.6.

Correlations of Q10, Q21, and Q39, "I Don't Make Changes"

	Q10	Q21	Q39
Q10	1.0	.439*	.258*
Q21	.439*	1.0	.409*
Q39	.258*	.409*	1.0

*2-tailed significance at the .05 level

The highest positive correlation for this common set of statements was Q10, " I don't make instructional adjustments or changes when students perform poorly on the CA," and Q21, "I don't make many changes in my teaching when students perform poorly on the CA." The lowest positive correlation for this set was Q10, "I don't make instructional adjustments or changes when students perform poorly on the CA," and Q39, " I don't need to change my teaching style; I am an effective teacher."

Table 4.7.

Correlations of Q6, Q8, Q11, Q14, and 23; "I Make Changes"

	Q6	Q8	Q11	Q14	Q23
Q6	1.0	.291*	2.11*	.460*	.428*
Q8	.291*	1.0	.261*	.372*	.416*
Q11	.211*	2.61*	1.0	.257*	.058
Q14	.460*	.372*	.257*	1.0	.368*
Q23	.428*	.416*	.058	.368*	1.0

*2-tailed significance at the .05 level

For this set of common statements, the highest correlation was Q14, " I make changes in my teaching methodology when my students perform poorly on the CA," and Q6, "I use research based strategies to remediate problems after formative assessments are given." The lowest correlation for this set was Q11, "I re-teach the 'big Ideas' when I have students perform poorly the CA," and Q23, "I take time to reflect on what I can do to help students who perform poorly the CA."

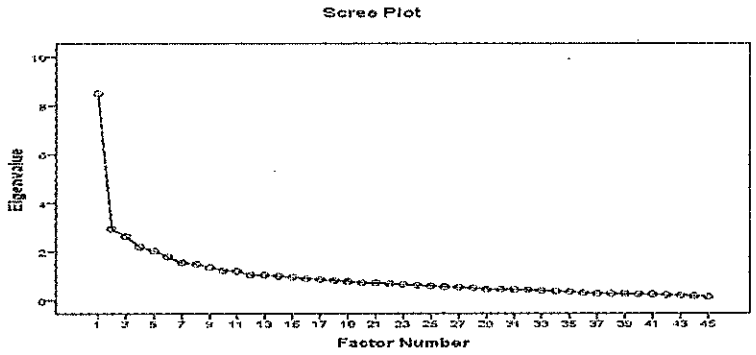
Principal Component Factor Analysis

Additional analyses were completed on the data using simple factor analysis procedures via SPSS (v. 18, 2009) methodology was the principal axis factor analysis (Moore, 2009), using squared multiple correlations of communalities. These correlations were used to group responses to individual item statements from the survey into factors with common themes. The original survey posted on surveymonkey.com included 45 items addressing teacher perceptions in relation to formative assessment and their responses to assessment results.

The 45 statements from the survey reduced to a total of 13 factor clusters. From the resultant pattern matrix (Moore, 2009) of the factor analysis, each of these 13 factors was examined, first in terms of its eigenvalue. Eigenvalues allow for an evaluation of how distinct a potential group (factor) of items might be (Moore, 2009).

Figure 4.1 shows the scree plot derived from this factor analysis. Although eigenvalues of 1 and above are typically acceptable indicators of distinct factors (Moore, 2009), the small size in this study (n=164), especially in relation to the number of items (45), led to the establishment of an especially high minimum eigenvalue of 2 as indicative of acceptable factor distinctiveness.

Figure 4.1. Scree Plot for Eigenvalues.



Five factors emerged as distinctive from the factor analysis:

- The teacher's disposition toward acting decisively on formative test results and involving students in formative assessment and resultant instructional planning.
- The teacher's use of professional development opportunities to improve teaching and student performance.
- CA's perceived role in instruction, student engagement, and student performance on summative state assessments.
- CA's impact on student performance.
- The teacher's use of anticipated state assessment coverage to determine which curricular topics to emphasize instructionally.

A pattern matrix (PM) was used to identify factor elements, and the structure matrix (SM) was used to interpret overall relationships among items within a factor. The PM contains coefficients that represent unique contributions of the items. The more items, the lower the pattern coefficients. In contrast, an SM, which represents the

correlations between the variables and the factors, is a representation of the variance in a measured variable, and is measured by a factor on both a unique and common contributions basis (Moore, 2009). To supplement the meaning of factors 1, 2, 3, and 4, factor analytic procedures yielded one or two additional relevant items; however, all relationship magnitudes (coefficients) are derived from the structure matrix. Below, each of these five factors is discussed in terms of its related item statements (see Table 4.8).

Table 4.8.

Summary Results of Factor Analysis

Factor Description	Eigenvalues	Survey Items*	Structure Matrix Coefficients*
1. The teacher's disposition toward acting decisively on formative test results and involving students in formative assessment and resultant instructional planning.	8.513	PM: 25 28 26 22 SM: 23 44	.661 .640 -.528 .519 .488 .369
2. The teacher's use of professional development opportunities to improve teaching and student performance.	2.942	PM: 5 4 20 35 SM: 17	.828 .813 .575 .559 .480
3. CA's perceived role in instruction, student engagement, and student performance on summative state assessments.	2.641	PM & SM: 29 34 18	 -.654 -.620 .514
4. CA's impact on student performance.	2.201	PM: 1 2 SM: 3	-.887 .726 .333
5. The teacher's use of anticipated state assessment coverage to determine what curricular topics to emphasize instructionally.	2.039	PM & SM: 32 31	 .650 .643

Factor 1 was "The teacher's disposition toward acting decisively on formative test results and involving students in formative assessment and resultant instructional planning." Given the eigenvalue of 8.51, data for this factor suggest most teachers

responding to the survey have positive dispositions towards formative assessment and are using the formative CA data to make changes in instruction when students perform poorly. Data also suggests these teachers are acting on the data obtained from the CA to change and improve instruction for poor performers. Examples of items in this factor were Q25, "I feel comfortable with how I utilize formative assessment results for my students"; Q28, "I provide feedback to my students as to what areas they performed poorly in on the state formative assessments"; Q22, "I include my students in the assessment process when formative scores are obtained"; and Q23, "I take time to reflect on what I can do to help students who perform poorly the CA."

Factor 2 was "the teacher's use of professional development opportunities to improve teaching and student performance," with an eigenvalue of 2.94. Data for this factor suggest that teachers are seeking professional development opportunities in the area of assessment, and are utilizing professional development to improve instruction for students who perform poorly on the CA. Examples of the items in this factor were Q4, "I feel like I am a more effective teacher because I attend professional development in the area of assessment, personal research, and individual and group instructional adjustment with the CA"; Q5, "My students are better prepared due to my personal preparation because I attend professional development in the area of CA"; and Q17, "I have had professional development training in reading and math in relation to utilizing formative assessment scores."

Factor 3 was "CA's perceived role in instruction, student engagement, and student performance on summative state assessments," with an eigenvalue of 2.64. The data for this factor suggest teachers do not have a strong perception that the CA is a tool to

increase student performance on summative assessment and student engagement in the classroom. Examples of items in this factor were Q18, "When I use the CA, my students don't perform any better on the spring summative assessments"; Q29, "My students who are eligible for special education services receive better instruction after the CA is given"; and Q34, "My students are more engaged when I change the way I teach after I get CA results."

Factor 4 was "CA's impact on student performance," with an eigenvalue of 2.20. Data for this factor suggest teachers do not have a strong perception that the formative assessment has an impact on future success in the spring summative assessments. Examples of items in this factor were Q1, "I use the Computerized Assessment (CA) and my students score well every year"; Q2, "I use the CA and my students perform poorly every year"; and Q3, "I have used the CA but only rarely."

Factor 5 was "the teacher's use of anticipated state assessment coverage to determine what curricular topics to emphasize instructionally," with an eigenvalue of 2.04. Data for this factor suggest teachers are using knowledge learned on the job from a mentor and/or from their building administrator to determine curricular focus for the summative assessment. Examples of items in this factor were Q31, "I find out what is going to be addressed in state summative assessments and teach only those topics specifically"; and Q32, "I am told by my building principal or district what specific topics from the state assessments to teach in the curriculum." Overall, factors 1, 2 and 5 provided the most pertinent information about the original question of this study, "Do intermediate elementary teachers who administer formative assessments use the data to adjust their teaching?"

Chapter 5

Conclusions and Recommendations

Findings suggest a response to the research question, "Do intermediate elementary-level teachers who administer formative assessments use the data to adjust their teaching?" Of the 164 teachers responding to the survey, ~97% agreed that they "use research based strategies to change instruction after formative assessments are given"; 98% of the respondents disagreed with the statement that they "do not change strategies after the CA is given." Overall, intermediate elementary-level teachers perceive that they are utilizing research-based strategies for remediation purposes when students perform poorly on the formative assessment.

In relation to the question of whether intermediate level teachers report they would change instructional methods and strategies as a result of students' poor performance on formative assessments, results indicated 98.8 of respondents do make changes in their teaching on the basis of these assessments. This finding may be attributable to NCLB-driven assessment policies, procedures, and routines that pervade the work environments of teachers. Teachers want and need to focus on student success and thus they try new methods and approaches to reach poor performers. In addition, respondents to the survey may also be reflecting a need to apply clinical problem-solving skills inherent in effective instruction, re-teaching content to students whose lack of mastery prevents them from performing successfully on formative assessments.

Another explanation possibly underlying the findings of this study is that teachers who responded to the survey may also be stating implicitly that they not only use the opportunity formative assessment procedures provide to apply their clinical skills, but

also to analyze curricula in terms of what students are and are not learning. As Quenemoen and Thurlow (2004) have noted, formative assessment procedures, such as those in place in the state in which this study was conducted, enable teachers with good curriculum knowledge, i.e. subject matter standards and indicators, to make reasonably precise determinations of what content a student knows and does not know.

Overall, the factor analysis provided valuable insights that shed light on the formative assessment process. First of all, this study showed that teachers perceive professional development to be an important tool for increasing student performance on summative assessments. Second, teachers perceive that their changes to instruction, after formative assessments, increase student performance and have an impact on student learning. However, teachers' have mixed perceptions about whether formative assessment increases student performance. Survey responses show that teachers do not perceive that using the formative assessment program increases student performance or student engagement in the classroom.

Finally, these findings could also reflect some degree of respondents' confidence in their ability to influence learning in their students. This possible relationship between teachers' self-report of formative assessment usage and their apparent belief that they can be instrumental in student curriculum mastery (i.e. efficacy) has been addressed in previous work in this area. MacMillan (2006) posited that teachers who embrace formative assessment are probably also likely to perceive themselves as knowledgeable and skillful enough to make successful instructional changes, i.e. they see themselves as having efficacy in promoting positive student learning outcomes.

Limitations of the Study

As with all research, there are limitations that need to be considered when trying to generalize interpretations of these analyses. The response rate of 164 teachers (3.1%) was very low in comparison with the total number of teachers in the midwestern state where this study took place. Therefore, no generalizations can be made about all teachers at the 3rd, 4th, 5th and 6th grade levels. The following is a discussion of some of additional limitations of the study.

First of all, the issue of the small sample size must be addressed. The number of teachers who participated in this study might not be representative of the entire population of teachers who use the formative assessment program. Of the 5,228 teachers who were invited participate in the study, only 164 completed the survey.

Additionally, a limitation of this study was that it looked across generations of teachers rather than differentiating by certain age ranges. Different generations of teachers may not have felt as though they could respond freely, or the survey might not have been specific to these generations.

This study was an observational study, not an experimental study, and so had no control components. Thus, it is difficult to determine exactly how age, experience, and geographic region might have affected survey results, given the countless other variables not measured or controlled. Therefore, any interpretation of these findings should recognize there may be additional important variables not included in this study.

Another limitation was the survey tool used in this study. The survey was self-generated and not a standardized survey. Additionally, because a very small number of teachers participated in the pilot study, survey statements underwent few revisions. It is,

therefore, possible that these statements were less valid than they may have been if a larger number of teachers had participated in the pilot study.

The manner in which grades are assigned to middle and elementary schools was a limitation of this study. This study focused on 3rd, 4th, 5th, and 6th grade teachers who used the formative tool. When teachers responded on the survey about the number of students in their respective classes, it was determined that some 6th grade teachers were actually teaching at the middle school level. Thus, what the researcher anticipated to be a typical 6th grade elementary school class of 18-20 students may have actually been a 6th grade middle school class of 50-100 students.

The limited diversity sampling of urban and suburban schools was also a limitation of this study. One reason for this was the inability of the researcher to survey teachers in two large school districts in this midwestern state. These two districts would have provided additional responses on surveymonkey.com. An improvement to this study would have been to follow the research protocol of these two school districts to gain access to their teachers. This would provide an additional cross section sampling of teachers for this study.

The unavailability of these school districts proved to be a limitation to the data collection of this study. After the survey link was emailed, teachers from these two districts were not allowed to participate because research protocols were not followed for these districts. The researcher was unable to follow protocols for these districts because the names of individual participants and individual schools were not known. Once this was made known to the researcher, approximately 40 responses had to be removed from the study.

Furthermore, an additional limitation was a lack of knowledge about professional development. Finding out what types of training and where this training was completed would have been beneficial to this study. Consequently, this information is a vital missing portion of this study.

Lastly, a limitation of this study is its focus on one midwestern state. Only a limited amount of information related to formative assessment utilization can be gathered from a single state. In the future, collecting data from multiple states would provide more results, and would be more indicative of how teachers utilize formative assessment data.

One of the goals of any study is to improve the methods and procedures of research. The following paragraphs will provide ways in which this study could have increased its validity and improved its quality and implementation.

One specific way this study could have been improved would have been to more proactively and creatively enlist districts to participate in the pilot study. Originally, three large districts were chosen for the pilot study. After numerous email and phone contacts with no response from one of those districts, it was determined the pilot should go forward. Other school districts were not substituted due to time constraints placed on the researcher. Another issue with the pilot study was one local school district allowed only a small number of schools to participate.

Another way in which this study could have been improved would have been to approach it purely as a descriptive study, instead of as a factor analytic study. The survey statements lent themselves to an assessment of how teachers feel about their methods and strategies. The inclusion of open-ended survey questions might have provided a richer, higher quality understanding of the teachers' dispositions about formative assessment. An

even better approach might have been to implement a training program for teachers in how to utilize formative assessment results, and then to determine quantitatively the efficacy of that training on student performance.

An important addition to this study would have been to determine which research-based strategies are used by teachers in conjunction with the formative tool. NCLB mandates the use of research-based strategies in best practice for teachers. This would have provided valuable information for the state's Department of Education as well as individual school districts.

Lastly, an improvement to this study would have been to increase focus on teachers who have students eligible for IDEA services. According to data collected in this study, teachers might be providing IDEA eligible students access to the general education curriculum, but might not be doing so differently from the way they teach the general education population. By increasing the focus on IDEA eligible students, this study could have described the research-based strategies teachers use to meet the needs of students in special education.

Recommendations

This study provided valuable information regarding how intermediate elementary-level teachers utilize state formative assessment results. The following are proposed recommendations for better understanding the use of formative results and the relationship of formative results to the summative assessment process, as a whole.

First of all, a longitudinal study of experienced and novice teachers over a fixed period of time could provide information about how teachers change their methodologies, or about when and how much training makes a difference in how they utilize formative

assessment results. This study would examine training programs at the building level. This suggestion leads logically to the next recommendation.

A future study would be to determine the influence, at the building level, of a training program focused on assessment score utilization. A quantitative study could determine the effect of a formal training program on how teachers utilize formative assessments (Khattri, Reeve, & Kane, 1998). This study could document the effect of a professional development program for teachers on the formative and summative assessment results of students. This program might look like a typical "in-house" professional development program, where master teachers share experiences and research in teaching, assessment, and learning, but would focus mainly on the interpretation and utilization of state assessment results, with an emphasis on learning new and innovative research-based strategies.

Another recommendation for further research would be to provide a training program for teachers who want to learn to utilize formative results and adjust their teaching methods and strategies for students who are IDEA eligible. Because students who are IDEA eligible are included in the summative assessment process, and their results are included with those of the general population, a formal training program to increase the knowledge and skills of general education teachers may increase the success of special education students.

Determining which professional development programs teachers have attended is also another area of future research. This study addressed only the notion that teachers attended professional development in the areas of state assessments and teaching strategies; it did not focus on the specific types of training teachers have attended.

Another recommendation for future study would be to survey teachers of grades 3, 4, and 5 only and not include grade 6. Because this study included 6th grade teachers, some of the class sizes were skewed. Using 3rd, 4th and 5th grades would allow for a consistent sampling size of students in each class.

Initially, one of the goals for this study was to determine specific strategies teachers use to remediate poor performers on the formative assessment. When pilot testing of the survey instrument was conducted, a portion of the survey was dedicated to allowing teachers to indicate specific strategies they used in the classroom. Unfortunately, only three percent of teachers in the pilot indicated specific strategies and methods they had used. In the future, a recommendation would be to have a larger emphasis on the name and derivation of specific strategies teachers use in their classrooms to remediate poor performers on the formative assessment.

Lastly, a recommendation for further research would be to determine the effects of the formative assessment tool on student performance on the state summative assessments. Since the formative tool is used to help provide teachers a snapshot of student performance at a given time, a future study might ask teachers about the effects of the formative assessment tool on summative results. This study did not address the effects of the formative assessment tool on student learning; it tried to determine only whether teachers utilize the formative results to remediate and improve instruction for poor performers.

Summary

The field of assessment, especially high stakes assessment, will be around for quite some time (Johnson, Johnson, Farenga, & Ness, 2008). With the reauthorization of

NCLB pending at the time of this study, its results provide timely information as to how 3rd, 4th, 5th and 6th grade teachers who utilize the formative assessment tool approach teaching in the areas of increasing summative results and in response to the increased accountability of NCLB. Hopefully, this information will inspire teachers to look at the way they "do business" in the classroom.

References

Ainsworth, L. B., & Viegut, D. J. (2006). *Common formative assessments: How to connect standards-based instruction and assessment.* Thousand Oaks, CA: Corwin Press.

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd Ed.). New York, NY: Jossey-Bass.

Appleton, K. (Ed.). (2006). *Elementary science teacher education: International perspectives on contemporary issues and practice.* Mahwah, NJ: Lawrence Erlbaum Associates.

Beare, H., & Boyd, W. L. (Eds.). (1993). *Restructuring schools: An international perspective on the movement to transform the control and performance of schools.* Bristol, PA: The Falmer Press.

Best, S. J., & Kreuger, B. S. (2004). *Internet data collection.* Thousand Oaks, CA: Sage Publications.

Boaler, J. (Ed.). (2000). *Multiple perspectives on mathematics teaching and learning.* Westport, CT: Ablex Publishing.

Brehm, J., & Gates, S. (2008). *Teaching, tasks, and trust: Functions of the public executive.* New York, NY: Russell Sage Foundation.

Bushweller, K. (1997). Teaching to the test. *The American School Board Journal.* 184(9), 20-25.

Butler, S. M., & McMunn, N. (2006). *A teacher's guide to classroom assessment: Understanding and using assessment to improve student learning.* San Francisco, CA: Jossey-Bass.

Cunningham, A. (2003). Supporting student-centered teaching and learning: Technology in Wake Forest university education programs. *Contemporary Issues in Technology and Teacher Education,*3(1).

Dell'Olio, J. M., & Donk, T. (2007). *Models of teaching: Connecting student learning with standards.* Thousand Oaks, CA: Sage Publications.

Firestone, W. A., Monfils, L. A., & Schorr, R. Y. (2004). *The ambiguity of teaching to the test: Standards, assessment, and educational reform.* Mahwah, NJ: Lawrence Erlbaum Associates.

Fiske, E.B., (1987, April 12). The teacher: Redesigning the American teacher. *The New York Times*, pp. 12, 18.

Gallager, C., & Worth, P. (2008). *Formative assessment policies, programs, and practices in the Southwest Region* (No. 041): U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment.* Philadelphia, PA: Routledge.

Gordon, E. W., & Bridglall, B. L. (2006). *Affirmative development: Cultivating academic ability.* Lanham, MD: Rowman & Littlefield.

Hall, K., & Burke, W. M. (2004). *Making formative assessment work: Effective practice in the primary classroom.* Berkshire, England: Open University Press.

Hashaway, R. (1998). *Assessment and evaluation of developmental learning: Qualitative individual assessment and individual models.* Westport, CT: Praeger.

Johnson, D. D., Johnson, B., Farenga, S. J., & Ness, D. (2008). *Stop high stakes testing: An appeal to America's conscience.* Lanham, MD: Rowman & Littlefield.

Justice, L. M., & Vukelich, C. (2008). *Achieving excellence in preschool literacy instruction.* New York, NY: The Guilford Press.

Kaplowitz, M., Hadlock, T. D., & Levine, R. (2004). A comparison of web and mail survey response rates. *Public Opinion Quarterly*, 68(1), 94–101.

Khattri, N., Reeve, A. L., & Kane, M. B. (1998). *Principles and practices of performance assessment.* Mahwah, NJ: Lawrence Erlbaum Associates.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: The University of Chicago Press.

Lancaster, G. (2005). *Research methods in management: A concise introduction to research in management and business consultancy.* Burlington, MA: Elsevier Butterworth-Heinemann.

Lagemann, E. C. (2000). *An elusive science: The troubling history of education research.* Chicago, IL: The University of Chicago Press.

Madaus, G., Russell, M. K., & Higgins, J. (2009). *The paradoxes of high stakes testing: How they affect students, their parents, teachers, principals, schools, and society.* Charlotte, NC: Information Age Publishing.

McLoughlin, C., & Taji, A. (Eds.). (2005). *Teaching in the sciences: Learner-centered approaches.* Binghamton, NY: Food Products Press.

McMillian, J. (2006). *Classroom assessment: Principles and Practice for effective standards-based instruction.* Upper Saddle River, NJ : Pearson.

Merrits, D., & Walter, R. (2006). *Starting point: Teaching entry level Geoscience* Lancaster, PA: Franklin and Marshall College.

Miller, A. H., Imrie, B. W., & Cox, K. (1998). *Student assessment in higher education: A handbook for assessing performance.* New York, NY: Routledge.

Miyazaki, I., & Schirokauer, C. (1981). *China's examination hell: The civil service examinations of imperial China.* New Haven, CT: Yale University Press.

Moore, D.S. (2009). *The basic practice of statistics* (5th ed.). New York, NY: W.H. Freeman and Company.

Olson, L. (2002, May 22). Up close and personal. *Education Week,* 21(37), 28-33.

Parsons, R.D., Hinson, L.H., & Sardo-Brown, D. (2001). *Educational psychology: A practitioner-researcher model of teaching.* Belmont, CA. Wadsworth-Thompson.

Popham, W. J. (2003). *Teach better, test better: The instructional role of assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Popham, W. J. (2010) *Classroom assessment: What teachers need to know* (6th ed.). Toronto, Ontario: Allyn & Bacon.

Quenemoen, R., &. Thurlow, M. L. (2004, July). *I say potato, you say potahto.* Paper presented at the annual meeting of the American Educational Research Association (AERA). National Center for Educational Outcomes, University of Minnesota. San Diego, CA.

Ravitich, D. (2002). A brief history of testing and accountability. Retrieved December 19, 2005, from the Hoover Institution Web site: http://www.hoover.org/publications/digest/4495866.html

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher.* 29(7), 4-14.

Stiggins, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan.* 81(3), 191-198.

Stiggins, R. J., & Conklin, N. F. (1992). *In Teachers' hands: Investigating the practices of classroom assessment.* Albany, NY: SUNY Press.

Sternberg, R., & Williams, W. (Eds.). (1998). *Intelligence, instruction, and assessment: Theory into practice.* Mahwah, NJ: Lawrence Erlbaum Associates.

Sunderman, G. L. (Ed.). (2008). *Holding NCLB accountable: Achieving accountability, equity, and school reform.* Thousand Oaks, CA: Corwin Press.

Thomas, R. M. (2005). *High-stakes testing: Coping with collateral damage.* Mahwah, NJ: Lawrence Erlbaum Associates.

Tienken, C., & Wilson, M. (2001). Using state standards and tests to improve instruction, practical assessment, research & evaluation, *Practical Assessment, Research & Evaluation,* 7(13). Retrieved January 4, 2010 from http://PAREonline.net/getvn.asp?v=7&n=13.

Tindal, G., & Haladyna, T. M. (Eds.). (2002). *Large-scale assessment programs for all students: Validity, technical adequacy, and implementation.* New York, NY: Lawrence Erlbaum Associates.

U.S. Department of Education. (2006). Raising achievement: Alternate assessments for students with disabilities. Retrieved January 5, 2010, from U.S. Department of Education Web site: http://www.ed.gov/policy/elsec/guid/raising/alt-assess-long.html.

Walsh, J., Kemerer, F. R., & Maniotis, L. (2005). *The educator's guide to Texas school law* (6th ed.). Austin, TX: University of Texas Press.

Watkins, C. E., & Campbell, V. (Eds.) (2000). *Testing and assessment in counseling practice* (2nd ed.). New York, NY: Lawrence Erlbaum Associates.

Weimer, M. G. (2002). *Learner-centered teaching: Five key changes to practice.* San Francisco, CA: Jossey-Bass.

Wormeli, R. (2006). *Fair isn't always equal: Assessing & grading in the differentiated classroom.* National Middle School Association. Portland, ME: Stenhouse Publishers.

Wiggins, G. (2002). *Toward genuine accountability: The case for a new state assessment system.* Retrieved March 15, 2006 from The George Lucas Educational Foundation Web site: http://www.glef.org

William, T. (2006) *Why do many students perform poorly in English composition and reading comprehension, and are there any relationships between instruction and student performance?* Unpublished master's thesis, University of Toronto, Toronto, Ontario, Canada.

Wilson, M. (2004). *Towards coherence between classroom assessment and accountability.* Chicago, IL: The University of Chicago Press.

Wilson, R., & Barenthal, M. (Eds.). (2006). *Systems for state science assessment.* Washington, DC: National Academies Press.

Yung, B. (2006). *Assessment reform in science: Fairness and fear.* Science and technology education library. Dordrecht, The Netherlands: Springer.

Appendix A Survey Items

1. I use the Computerized Assessment (CA) and my students score well every year.

2. I use the CA and my students perform poorly every year.

3. I have used the CA but only rarely.

4. I feel like I am a more effective teacher because I attend professional development in the area of assessment, personal research, and individual and group instructional adjustment with the CA.

5. My students are better prepared due to my personal preparation because I attend professional development in the area of CA.

6. I use research-based strategies to remediate problems after formative assessments are given.

7. I use the CA as it is intended to be used, not as a practice for the spring summative assessments.

8. My students are better prepared for the state assessments because of my teaching strategies and methods I incorporate after using the CA.

9. I reflect on my effectiveness as a teacher if I have students perform poorly on the CA.

10. I don't make instructional adjustments or changes when students perform poorly on the CA.

11. I re-teach the "big Ideas" when I have students perform poorly the CA.

12. I am encouraged by my principal to use and utilize the CA scores for my students.

13. My principal encourages me to adjust my teaching style, method, and materials when I have students who perform poorly on the CA.

14. I make changes in my teaching methodology when my students perform poorly on the CA.

15. This survey is helping me view the CA from a different "angle".

16. I use the CA as a practice run for the summative assessment in the spring.

17. I have had professional development training in reading and math in relation to utilizing formative assessment scores.

18. When I use the CA, my students don't perform any better on the spring summative assessments.

19. The way I adjust instruction for general education students looks different from the way I adjust instruction for students eligible for special education services.

20. I have attended professional development on how to change my teaching methods and learn additional strategies if I have poor performing students on the CA.

21. I don't make many changes in my teaching when students perform poorly on the CA.

22. I include my students in the assessment process when formative scores are obtained.

23. I take time to reflect on what I can do to help students who perform poorly the CA.

24. I include parents in the assessment process when formative scores are obtained.

25. I feel comfortable with how I utilize formative assessment results for my students.

26. Sometimes I feel like I have no idea what's going on when it comes to state formative and summative assessments.

27. I know I can engage students who perform poorly on the formative assessments.

28. I provide feedback to my students as to what areas they performed poorly in on the state formative assessments.

29. My students who are eligible for special education services receive better instruction after the CA is given.

30. I use a number of various teaching strategies in my class to prepare students for the CA.

31. I find out what is going to be addressed in state summative assessments and teach only those topics specifically.

32. I am told by my building principal or district what specific topics from the state assessments to teach in the curriculum.

33. I have viewed the state elementary standards and I have attended professional development that helps me teach my students more effectively in those standards.

34. My students are more engaged when I change the way I teach after I get CA results.

35. I have had professional development training in mathematics and utilization of formative assessment scores.

36. I have learned research-based teaching practices and I have integrated them into my classroom.

37. I don't feel the need to attend additional professional development to teach my students more effectively; I need more resources and support.

38. It makes no difference if I use the CA or not, my students will always perform the same way.

39. I don't need to change my teaching style; I am an effective teacher.

40. I regularly actively engage all of my students.

41. My students perform poorly on the CA and are rarely actively engaged in class.

42. I am not responsible for the success of students; it is the responsibility of the special education teacher and staff.

43. I feel that my students are NOT prepared for the CA.

44. When my students perform poorly on the CA, I know exactly what to do to remediate the problem.

45. After taking this survey, I feel like I may want to learn more teaching methods and strategies to meet the needs of poor performers on the CA.

Appendix B

Results of the Pilot Study

A reliability analysis was completed using SPSS (v.18, 2009) and it was determined that the survey had a Cronbach's alpha of .71, with .70 within the standardized items respectively. This data suggests that there is adequate reliability evident in the survey to continue this study (Moore, 2009).

Pearson correlations were also analyzed between individual statements (see Appendix C) and provide the raw correlation data for this pilot study. The definition of "higher correlations" for this pilot study are a correlation of .600 and above; thus, from these data one can conclude that the answers given were not made in error, nor did the participants misunderstand the statements. Seven pairs of statements correlated highly; thus, roughly 63 percent of the questions correlate positively within the survey. The scale analysis yielded higher positive correlations for the following:

1. In the area of teaching effectiveness and professional development, the correlations were .89 for statements 4 and 5: "I am more effective because I attend professional development"; and "My students are better prepared because of my professional development."
2. The correlation was .68 for statements 8 and 9: "My students are better prepared because of the strategies and methods I incorporate after the KCA is given"; and "I reflect on my effectiveness if I have students perform poorly on the KCA."

3. Another area was the teachers' belief that they actively engage students. The correlation was .60 for statements 1 and 39:. "I use the KCA and my students perform well every year"; and "I regularly actively engage all of my students."
4. A correlation of .75 was obtained for statements 6 and 27: "I use research based strategies to remediate problems after formative assessments are given"; and "I provide feedback to my students as to what areas they performed poorly in on the formative assessments."

Other high correlations between questions were as follows:

5. A correlation of .766 was found for statements 21 and 34: "I don't make any changes in my teaching when students perform poorly on the KCA"; and "My students are more engaged when I change the way I teach after the KCA results are provided."
6 The correlation was .624 for statements 29 and 5: "I use a number of various teaching strategies in my class to prepare students for the KCA"; and "My students are better prepared due to my personal preparation because I attend professional development in the areas of the KCA."
7 A correlation of .604 was found for statements 42 and 36: "I feel that my students are not prepared for the KCA"; and " I don't need to change my teaching style, I am an effective teacher."

Scale analysis of the survey provided higher negative correlations for the following comparisons:

1. A negative correlation of -.74 was found for statements 1 and 2: "I use the KCA every year and my students perform well"; and "I use the KCA and my students perform poorly every year."
2. A negative correlation of -.64 was found for statements 4 and 20: "I am more effective because I attend professional development in the area of assessment and instruction"; and "I don't make any changes in my teaching when students perform poorly on the KCA."
3. The negative correlation was -.64 for statements 8 and 17: "My students are better prepared because of the strategies I use after the KCA is given"; and "When I use the KCA my students don't perform any better on the spring summative assessments."
4. A negative correlation of -.76 was found for statements 7 and 15: " I use the KCA as it is intended to be used, not as practice for the summative assessments"; and "I use the KCA as a practice run for the spring summative assessments."
5. The negative correlation was -.60 for statements 17 and 39: "I use the KCA every year and my students don't perform any better on the summative assessments"; and "I regularly actively engage all of my students."
6. A negative correlation of -.70 was found for statements 20 and 9: " I don't make any changes in my teaching when students perform poorly on the KCA"; and "I reflect on my effectiveness as a teacher when students perform poorly on the KCA."

These sets of statements possess a highly negative correlate within the survey. This allows one to conclude that these sets of statements were answered in the opposite fashion. Thus, when the respondents answered the statements there were clear demarcations between the two sets of statements. When the pilot study was conducted, differentiation was made between the teachers and administrators. Pilot testing of the initial survey was completed in late 2007.

Appendix C

Correlations Matrix of Pearson Coefficients

Correlations Matrix

		Q1	Q2	Q3	Q4
Correlation	Q1	1.000	-.640**	-.271**	.015
	Q2	-.640**	1.000	.263**	.034
	Q3	-.271**	.263**	1.000	.024
	Q4	.015	.034	.024	1.000
	Q5	.080	.005	-.040	.709**
	Q6	.126	-.098	-.116	.200*
	Q7	.069	-.023	-.002	.256**
	Q8	.196*	-.165*	-.283**	.203*
	Q9	.245**	-.165*	-.309**	.190*
	Q10	-.126	.033	.109	-.202*
	Q11	-.083	.015	-.083	.076
	Q12	.163*	-.155*	-.193*	.124
	Q13	.063	-.155*	-.102	.142
	Q14	.076	-.111	-.142	.284**
	Q15	.142	-.060	.141	.133
	Q16	.126	-.179*	-.156*	-.019
	Q17	-.026	-.008	-.081	.292**
	Q18	-.237**	.299**	.230*	-.019
	Q19	-.114	.100	.050	-.100
	Q20	-.103	.184	.043	.440**
	Q21	-.031	.016	.183*	-.165*
	Q22	.067	-.147	-.177*	.033
	Q23	.169*	-.195*	-.096	.186*
	Q24	.214*	-.160	-.193*	.152
	Q25	.168*	-.186*	-.206*	.129
	Q26	-.268**	.260**	.280**	-.081
	Q27	.144	-.064	-.075	.127
	Q28	.239**	-.282**	-.152	.059
	Q29	.013	-.011	-.037	.064
	Q30	.210*	-.140	-.110	.149
	Q31	.029	.049	.006	-.110
	Q32	.089	-.094	-.067	-.087
	Q33	.101	-.048	-.120	.260**
	Q34	.115	-.036	-.054	.108
	Q35	-.017	.081	-.123	.334**
	Q36	.005	-.102	-.124	.227*
	Q37	-.043	-.068	.045	-.283**
	Q38	-.178*	.164*	.215*	-.101
	Q39	-.073	-.043	.091	-.211*
	Q40	.119	-.095	-.041	.028
	Q41	-.189*	.188*	.178*	.035
	Q42	-.013	.032	.100	-.059
	Q43	-.406**	.314**	.338**	-.081
	Q44	.218*	-.100	-.059	.052
	Q45	-.032	-.060	.052	.144

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

Correlations Matrix

	Q5	Q6	Q7	Q8	Q9	Q10
Q1	.080	.126	.069	.196*	.245*	-.126
Q2	.005	-.098	-.023	-.165*	-.165*	.033
Q3	-.040	-.116	-.002	-.283*	-.309*	.109
Q4	.709*	.200*	.256*	.203*	.190*	-.202*
Q5	1.000	.230*	.171*	.107	.194*	-.156*
Q6	.230*	1.000	.212*	.291*	.142	-.312*
Q7	.171*	.212*	1.000	.191*	.202*	-.276*
Q8	.107	.291*	.191*	1.000	.284*	-.338*
Q9	.194*	.142	.202*	.284*	1.000	-.337*
Q10	-.156*	-.312*	-.276*	-.338*	-.337*	1.000
Q11	.044	.211*	.045	.268*	.234*	-.354*
Q12	.104	.308*	.192*	.388*	.196*	-.331*
Q13	.142	.316*	.265*	.315*	.247*	-.233*
Q14	.274*	.460*	.196*	.372*	.326*	-.393*
Q15	.149	.023	.050	.174*	.024	-.174*
Q16	.035	.126	-.396*	.110	-.038	.019
Q17	.320*	.288*	.215*	.326*	.067	-.085
Q18	-.115	-.029	-.100	-.322*	-.239*	.370*
Q19	-.010	.025	-.150	-.031	-.056	.093
Q20	.412*	.189*	.198*	.116	.104	-.145
Q21	-.246*	-.386*	-.131	-.351*	-.303*	.439*
Q22	.126	.365*	.154*	.354*	.197*	-.294*
Q23	.128	.428*	.139	.416*	.197*	-.385*
Q24	.121	.204*	.136	.192*	.213*	-.301*
Q25	.069	.315*	.277*	.383*	.125	-.192*
Q26	-.134	-.345*	-.339*	-.356*	-.143	.334*
Q27	.101	.167*	.267*	.256*	.195*	-.130
Q28	.090	.369*	.141	.375*	.214*	-.320*
Q29	.214*	.061	.061	.123	.037	-.134
Q30	.140	.346*	.125	.349*	.120	-.241*
Q31	-.135	-.082	-.192*	.006	.005	.194*
Q32	-.021	-.096	-.280*	.038	.057	.156*
Q33	.321*	.367*	.211*	.263*	.117	-.282*
Q34	.176*	.170*	.198*	.257*	.259*	-.346*
Q35	.406*	.238*	.175*	.169	-.025	-.011
Q36	.239*	.291*	.126	.280*	.126	-.166*
Q37	-.254*	.049	-.105	-.030	-.116	-.043
Q38	-.147	-.131	-.204*	-.376*	-.401*	.411*
Q39	-.114	-.079	-.086	-.183*	-.210*	.258*
Q40	.043	.195*	.108	.143	.121	-.126
Q41	-.014	-.119	-.079	-.328*	.017	.174*
Q42	-.008	-.060	-.110	-.169*	-.045	.133
Q43	-.130	-.098	-.234*	-.253*	-.122	.221*
Q44	.159*	.267*	.089	.116	.033	-.067
Q45	.091	-.077	-.098	.079	.017	-.107

Correlations Matrix

	Q11	Q12	Q13	Q14	Q15	Q16
Q1	-.083	.163	.063	.076	.142	.126
Q2	.015	-.155	-.155	-.111	-.060	-.179
Q3	-.083	-.193	-.102	-.142	.141	-.156
Q4	.076	.124	.142	.284	.133	-.019
Q5	.044	.104	.142	.274	.149	.035
Q6	.211	.308	.316	.460	.023	.126
Q7	.045	.192	.265	.196	.050	-.396
Q8	.268	.388	.315	.372	.174	.110
Q9	.234	.196	.247	.326	.024	-.038
Q10	-.354	-.331	-.233	-.393	-.174	.019
Q11	1.000	.199	.205	.257	.004	.045
Q12	.199	1.000	.583	.346	.165	.019
Q13	.205	.583	1.000	.423	.112	-.040
Q14	.257	.346	.423	1.000	.132	.071
Q15	.004	.165	.112	.132	1.000	.118
Q16	.045	.019	-.040	.071	.118	1.000
Q17	.147	.279	.345	.306	.039	.114
Q18	-.134	-.278	-.224	-.263	-.209	-.106
Q19	.007	.013	.075	.044	-.148	.011
Q20	.181	.186	.297	.201	.042	-.129
Q21	-.289	-.361	-.345	-.526	-.105	-.170
Q22	.175	.176	.175	.262	.113	.179
Q23	.416	.351	.285	.386	.145	.157
Q24	.058	.212	.224	.207	.047	.061
Q25	.145	.335	.222	.324	.020	.018
Q26	-.114	-.294	-.143	-.261	-.079	-.022
Q27	.113	.196	.197	.163	.132	-.179
Q28	.279	.347	.204	.299	.025	.077
Q29	-.049	.128	.119	.056	.317	-.093
Q30	.299	.225	.301	.367	-.061	.223
Q31	-.057	-.107	-.003	-.060	.203	.143
Q32	-.007	-.008	.058	.022	.038	.231
Q33	.207	.326	.264	.398	.020	.099
Q34	-.018	.158	.248	.305	.289	-.076
Q35	.074	.299	.278	.173	.072	.056
Q36	.271	.230	.226	.255	.066	.034
Q37	-.125	-.005	.014	-.017	-.039	.177
Q38	-.144	-.173	-.210	-.271	-.192	-.064
Q39	-.132	-.126	-.149	-.179	-.085	.029
Q40	.143	.132	.115	.154	-.013	-.018
Q41	-.237	-.176	-.196	-.143	-.045	-.056
Q42	-.126	-.099	.028	-.022	-.046	.030
Q43	-.131	-.211	-.091	-.099	-.184	.028
Q44	-.053	.134	.049	.144	.125	.003
Q45	.180	-.018	.024	.044	.291	.093

Correlations Matrix

	Q17	Q18	Q19	Q20	Q21	Q22
Q1	-.026	-.237*	-.114	-.103	-.031	.067
Q2	-.008	.299*	.100	.184*	.016	-.147
Q3	-.081	.230*	.050	.043	.183*	-.177*
Q4	.292*	-.019	-.100	.440*	-.165*	.033
Q5	.320*	-.115	-.010	.412*	-.246*	.126
Q6	.288*	-.029	.025	.189*	-.386*	.365*
Q7	.215*	-.100	-.150	.198*	-.131	.154
Q8	.326*	-.322*	-.031	.116	-.351*	.354*
Q9	.067	-.239*	-.056	.104	-.303*	.197
Q10	-.085	.370*	.093	-.145	.439*	-.294*
Q11	.147	-.134	.007	.181	-.289*	.175
Q12	.279*	-.278*	.013	.186*	-.361*	.176
Q13	.345*	-.224*	.075	.297*	-.345*	.175
Q14	.306*	-.263*	.044	.201*	-.526*	.262*
Q15	.039	-.209*	-.148	.042	-.105	.113
Q16	.114	-.106	.011	-.129	-.170	.179
Q17	1.000	-.091	.024	.470*	-.130	.270*
Q18	-.091	1.000	.067	-.110	.322*	-.275*
Q19	.024	.067	1.000	-.097	-.011	.005
Q20	.470*	-.110	-.097	1.000	-.180*	.068
Q21	-.130	.322*	-.011	-.180*	1.000	-.359*
Q22	.270*	-.275*	.005	.068	-.359*	1.000
Q23	.210*	-.259*	-.068	.148	-.415*	.453*
Q24	.130	-.256*	.029	.141	-.184*	.223*
Q25	.311*	-.255*	-.179*	.140	-.284*	.370*
Q26	-.189*	.262*	.101	-.092	.294*	-.285*
Q27	.165*	-.166*	.001	.128	-.021	.169*
Q28	.165*	-.259*	.016	.121	-.331*	.395*
Q29	-.007	-.263*	-.106	.158*	-.091	.043
Q30	.272*	-.206*	.025	.207*	-.418*	.253*
Q31	.028	.078	-.077	-.119	.115	-.138
Q32	.058	-.119	.077	-.007	.073	.075
Q33	.384*	-.289*	-.061	.335*	-.317*	.284*
Q34	.042	-.258*	-.113	.134	-.189*	.106
Q35	.526*	-.163*	-.061	.421*	-.130	.115
Q36	.309*	-.223*	-.141	.296*	-.165*	.313*
Q37	-.070	-.012	.056	-.251*	.053	.156
Q38	-.172*	.452*	.184	-.097	.409*	-.275*
Q39	-.095	.132	.064	-.096	.106	.054
Q40	.188*	-.026	.041	.129	-.121	.189*
Q41	-.172*	.175*	.062	-.117	.138	-.249*
Q42	-.063	.117	.161*	-.087	.131	-.099
Q43	-.128	.259*	.283*	-.091	.130	-.172*
Q44	.220*	-.199*	-.077	.143	-.033	.252*
Q45	-.127	.004	-.022	-.048	-.101	-.060

Correlations Matrix

	Q23	Q24	Q25	Q26	Q27	Q28
Q1	.169*	.214*	.168*	-.268*	.144	.239*
Q2	-.195*	-.160*	-.186*	.260*	-.064	-.282*
Q3	-.096	-.193*	-.206*	.280*	-.075	-.152
Q4	.186*	.152	.129	-.081	.127	.059
Q5	.128	.121	.069	-.134	.101	.090
Q6	.428*	.204*	.315*	-.345*	.167*	.369*
Q7	.139	.136	.277*	-.339*	.267*	.141
Q8	.416*	.192*	.383*	-.356*	.256*	.375*
Q9	.197*	.213*	.125	-.143	.195*	.214*
Q10	-.385*	-.301*	-.192*	.334*	-.130	-.320*
Q11	.416*	.058	.145	-.114	.113	.279*
Q12	.351*	.212*	.335*	-.294*	.196*	.347*
Q13	.285*	.224*	.222*	-.143	.197*	.204*
Q14	.386*	.207*	.324*	-.261*	.163*	.299*
Q15	.145	.047	.020	-.079	.132	.025
Q16	.157*	.061	.018	-.022	-.179*	.077
Q17	.210*	.130	.311*	-.189*	.165*	.165*
Q18	-.259*	-.256*	-.255*	.262*	-.166*	-.259*
Q19	-.068	.029	-.179*	.101	.001	.016
Q20	.148	.141	.140	-.092	.128	.121
Q21	-.415*	-.184*	-.284*	.294*	-.021	-.331*
Q22	.453*	.223*	.370*	-.285*	.169	.395*
Q23	1.000	.243*	.363*	-.253*	.215*	.332*
Q24	.243*	1.000	.319*	-.089	.112	.305*
Q25	.363*	.319*	1.000	-.454*	.319*	.473*
Q26	-.253*	-.089	-.454*	1.000	-.193*	-.349*
Q27	.215*	.112	.319*	-.193*	1.000	.195*
Q28	.332*	.305*	.473*	-.349*	.195*	1.000
Q29	-.056	.040	.020	-.033	.101	.110
Q30	.434*	.194*	.220*	-.146	.221*	.269*
Q31	-.107	-.027	-.118	.178*	-.082	-.151
Q32	-.033	-.011	.072	.084	.150	.141
Q33	.303*	.154*	.257*	-.342*	.194*	.226*
Q34	.143	.195*	.104	-.179*	.144	.154
Q35	.152	.212*	.300*	-.181*	.203*	.171*
Q36	.364*	.092	.259*	-.201*	.354*	.255*
Q37	-.109	.035	.016	.012	-.210*	.038
Q38	-.296*	-.194*	-.189*	.358*	-.160*	-.191*
Q39	-.062	-.089	.018	.104	-.021	-.064
Q40	.253*	.083	.172*	.017	.325*	.230*
Q41	-.285*	-.220*	-.251*	.180*	-.159*	-.225*
Q42	-.222*	-.078	-.096	.142	-.139	-.006
Q43	-.169*	-.125	-.333*	.471*	-.109	-.211*
Q44	.156	.077	.342*	-.275*	.230*	.164*
Q45	.108	.004	-.166	.070	.028	-.056

Correlations Matrix

	Q29	Q30	Q31	Q32	Q33	Q34
Q1	.013	.210*	.029	.089	.101	.115
Q2	-.011	-.140	.049	-.094	-.048	-.036
Q3	-.037	-.110	.006	-.067	-.120	-.054
Q4	.064	.149	-.110	-.087	.260*	.108
Q5	.214*	.140	-.135	-.021	.321**	.176*
Q6	.061	.346**	-.082	-.096	.367**	.170*
Q7	.061	.125	-.192*	-.280**	.211*	.198*
Q8	.123	.349**	.006	.038	.263**	.257**
Q9	.037	.120	.005	.057	.117	.259**
Q10	-.134	-.241*	.194*	.156*	-.282**	-.346**
Q11	-.049	.299**	-.057	-.007	.207*	-.018
Q12	.128	.225*	-.107	-.008	.326**	.158*
Q13	.119	.301**	-.003	.058	.264**	.248**
Q14	.056	.367**	-.060	.022	.398**	.305**
Q15	.317**	-.061	.203*	.038	.020	.289**
Q16	-.093	.223*	.143	.231*	.099	-.076
Q17	-.007	.272**	.028	.058	.384**	.042
Q18	-.263**	-.206*	.078	-.119	-.289**	-.258**
Q19	-.106	.025	-.077	.077	-.061	-.113
Q20	.158*	.207*	-.119	-.007	.335**	.134
Q21	-.091	-.418**	.115	.073	-.317**	-.189*
Q22	.043	.253**	-.138	.075	.284**	.106
Q23	-.056	.434**	-.107	-.033	.303**	.143
Q24	.040	.194*	-.027	-.011	.154*	.195*
Q25	.020	.220*	-.118	.072	.257**	.104
Q26	-.033	-.146	.178*	.084	-.342**	-.179*
Q27	.101	.221*	-.082	.150	.194*	.144
Q28	.110	.269**	-.151	.141	.226*	.154
Q29	1.000	-.040	.078	.046	.016	.454**
Q30	-.040	1.000	.002	.048	.396**	.050
Q31	.078	.002	1.000	.389**	-.083	-.036
Q32	.046	.048	.389**	1.000	.003	-.054
Q33	.016	.396**	-.083	.003	1.000	.092
Q34	.454**	.050	-.036	-.054	.092	1.000
Q35	.152	.177*	-.096	.037	.402**	.101
Q36	.003	.303**	-.138	.038	.351**	.020
Q37	-.099	-.126	-.070	-.006	-.119	-.122
Q38	-.263**	-.324**	-.057	-.022	-.344**	-.328**
Q39	-.234*	-.035	-.080	-.004	-.189*	-.320**
Q40	-.125	.323**	-.192*	.048	.112	-.025
Q41	-.012	-.348**	.124	-.076	-.100	-.029
Q42	-.014	-.106	.093	.169*	-.156*	-.042
Q43	-.083	-.200*	.023	.031	-.186*	-.112
Q44	.091	.077	-.022	.111	.298**	.135
Q45	.044	.018	.155*	.020	.071	.211*

Correlations Matrix

	Q35	Q36	Q37	Q38	Q39	Q40
Q1	-.017	.005	-.043	-.178*	-.073	.119
Q2	.081	-.102	-.068	.164*	-.043	-.095
Q3	-.123	-.124	.045	.215*	.091	-.041
Q4	.334*	.227*	-.283*	-.101	-.211*	.028
Q5	.406*	.239*	-.254*	-.147	-.114	.043
Q6	.238*	.291*	.049	-.131	-.079	.195*
Q7	.175*	.126	-.105	-.204*	-.086	.108
Q8	.169*	.280*	-.030	-.376*	-.183*	.143
Q9	-.025	.126	-.116	-.401*	-.210*	.121
Q10	-.011	-.166*	-.043	.411*	.258*	-.126
Q11	.074	.271*	-.125	-.144	-.132	.143
Q12	.299*	.230*	-.005	-.173*	-.126	.132
Q13	.278*	.226*	.014	-.210*	-.149	.115
Q14	.173*	.255*	-.017	-.271*	-.179*	.154
Q15	.072	.066	-.039	-.192*	-.085	-.013
Q16	.056	.034	.177*	-.064	.029	-.018
Q17	.526*	.309*	-.070	-.172*	-.095	.188*
Q18	-.163*	-.223*	-.012	.452*	.132	-.026
Q19	-.061	-.141	.056	.184*	.064	.041
Q20	.421*	.296*	-.251*	-.097	-.096	.129
Q21	-.130	-.165*	.053	.409*	.106	-.121
Q22	.115	.313*	.156*	-.275*	.054	.189*
Q23	.152	.364*	-.109	-.296*	-.062	.253*
Q24	.212*	.092	.035	-.194*	-.089	.083
Q25	.300*	.259*	.016	-.189*	.018	.172*
Q26	-.181*	-.201*	.012	.358*	.104	.017
Q27	.203*	.354*	-.210*	-.160*	-.021	.325*
Q28	.171*	.255*	.038	-.191*	-.064	.230*
Q29	.152	.003	-.099	-.263*	-.234*	-.125
Q30	.177*	.303*	-.126	-.324*	-.035	.323*
Q31	-.096	-.138	-.070	-.057	-.080	-.192*
Q32	.037	.038	-.006	-.022	-.004	.048
Q33	.402*	.351*	-.119	-.344*	-.189*	.112
Q34	.101	.020	-.122	-.328*	-.320*	-.025
Q35	1.000	.323*	-.196*	-.065	-.128	.109
Q36	.323*	1.000	-.125	-.079	.016	.317*
Q37	-.196*	-.125	1.000	.166*	.335*	-.010
Q38	-.065	-.079	.166*	1.000	.349*	-.088
Q39	-.128	.016	.335*	.349*	1.000	.107
Q40	.109	.317*	-.010	-.088	.107	1.000
Q41	-.145	-.314*	.022	.189*	.024	-.346*
Q42	-.060	-.200*	.072	.195*	.077	-.230*
Q43	-.083	-.225*	.024	.250*	.117	-.088
Q44	.199*	.232*	-.036	-.077	.054	.138
Q45	-.135	.032	-.138	-.094	-.250*	-.053

Correlations Matrix

	Q41	Q42	Q43	Q44	Q45
Q1	-.189	-.013	-.406	.218	-.032
Q2	.188	.032	.314	-.100	-.060
Q3	.178	.100	.338	-.059	.052
Q4	.035	-.059	-.081	.052	.144
Q5	-.014	-.008	-.130	.159	.091
Q6	-.119	-.060	-.098	.267	-.077
Q7	-.079	-.110	-.234	.089	-.098
Q8	-.328	-.169	-.253	.116	.079
Q9	.017	-.045	-.122	.033	.017
Q10	.174	.133	.221	-.067	-.107
Q11	-.237	-.126	-.131	-.053	.180
Q12	-.176	-.099	-.211	.134	-.018
Q13	-.196	.028	-.091	.049	.024
Q14	-.143	-.022	-.099	.144	.044
Q15	-.045	-.046	-.184	.125	.291
Q16	-.056	.030	.028	.003	.093
Q17	-.172	-.063	-.128	.220	-.127
Q18	.175	.117	.259	-.199	.004
Q19	.062	.161	.283	-.077	-.022
Q20	-.117	-.087	-.091	.143	-.048
Q21	.138	.131	.130	-.033	-.101
Q22	-.249	-.099	-.172	.252	-.060
Q23	-.285	-.222	-.169	.156	.108
Q24	-.220	-.078	-.125	.077	.004
Q25	-.251	-.096	-.333	.342	-.166
Q26	.180	.142	.471	-.275	.070
Q27	-.159	-.139	-.109	.230	.028
Q28	-.225	-.006	-.211	.164	-.056
Q29	-.012	-.014	-.083	.091	.044
Q30	-.348	-.106	-.200	.077	.018
Q31	.124	.093	.023	-.022	.155
Q32	-.076	.169	.031	.111	.020
Q33	-.100	-.156	-.186	.298	.071
Q34	-.029	-.042	-.112	.135	.211
Q35	-.145	-.060	-.083	.199	-.135
Q36	-.314	-.200	-.225	.232	.032
Q37	.022	.072	.024	-.036	-.138
Q38	.189	.193	.250	-.077	-.094
Q39	.024	.077	.117	.054	-.250
Q40	-.346	-.230	-.088	.138	-.053
Q41	1.000	.249	.341	-.062	-.004
Q42	.249	1.000	.198	-.022	-.107
Q43	.341	.198	1.000	-.309	.103
Q44	-.062	-.022	-.309	1.000	-.186
Q45	-.004	-.107	.103	-.186	1.000