2011

The Impact of Using Formative Assessment Attributes in Daily Instruction on Student Affect

Joshua William Ruland

Loyola University Chicago

Recommended Citation

Ruland, Joshua William, "The Impact of Using Formative Assessment Attributes in Daily Instruction on Student Affect" (2011). Dissertations. 44.
https://ecommons.luc.edu/luc_diss/44
LOYOLA UNIVERSITY CHICAGO

THE IMPACT OF USING FORMATIVE ASSESSMENT ATTRIBUTES IN DAILY INSTRUCTION ON STUDENT AFFECT

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL OF EDUCATION
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF EDUCATION

PROGRAM IN CURRICULUM AND INSTRUCTION

BY

JOSHUA WILLIAM RULAND

CHICAGO, ILLINOIS

AUGUST 2011
ACKNOWLEDGMENTS

I would like to acknowledge the hard work and dedication of my committee chair, Dr. Brigid Schultz. Her guidance and support was a cornerstone in this project. I would also like to acknowledge the work of my committee, Dr. Anne Marie Ryan and Dr. Marla Israel. Their questions, suggestions, and feedback were greatly appreciated. I would like to acknowledge Dr. R.J. Breunlin for all the assistance he provide me, both in and out of the classroom, with designing and running the statistical analysis on the collected data. I would like to thank Dr. Russ White for his guidance in education and the dissertation process. I would also like to thank Valerie Collier for her editing and format work. Finally, I would like to thank all those who began questioning current educational practices and who started to think outside the box. As Dylan Wiliam has said, “it is not about being right or wrong, it is about learning.”
DEDICATION

I dedicated this work first and foremost to my loving, understanding, and supportive wife, Amber. She is my best friend and biggest supporter of my work. Without her, I would never have been able to finish. I also dedicate this work to my wonderful children. First, Brayden, my oldest who has suffered the longest, his encouragement has been life-saving. Second, Micah, who was born during my coursework, his energy and excitement helped push me forward. And finally, Ella, my daughter born during the completion of this journey, her unconditional love has meant the world to me.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ................................................................................................. iii

LIST OF TABLES ........................................................................................................... vii

ABSTRACT ..................................................................................................................... ix

CHAPTER I: INTRODUCTION ...................................................................................... 1
  Introduction ............................................................................................................... 1
  Statement of the Problem ....................................................................................... 5
  Background of the Study ....................................................................................... 7
  Research Questions ............................................................................................... 9
  Significance of the Study ...................................................................................... 10
  Definition of Terms ............................................................................................. 11
  Limitations of the Study ..................................................................................... 12

CHAPTER II: REVIEW OF RELATED LITERATURE ............................................... 13
  History of Grading and Assessments .................................................................. 13
  Current Perceptions of Grading and Assessments ............................................. 20
  Need for Change .................................................................................................. 32
  Formative Assessment as Possible Answer ....................................................... 35
  Research on Formative Assessment .................................................................. 38
  Research on the Five Attributes ....................................................................... 45
  Summation ............................................................................................................ 68
  Research Questions ............................................................................................ 70

CHAPTER III: METHODS .......................................................................................... 72
  Purpose of the Study ............................................................................................ 72
  Research Questions ............................................................................................ 73
  Sample .................................................................................................................. 74
  Instrumentation ................................................................................................... 80
  Intervention .......................................................................................................... 84
  Statistical Procedures ......................................................................................... 87
  Hypotheses .......................................................................................................... 90
  Ethical Considerations ....................................................................................... 91

CHAPTER IV: RESULTS OF THE STUDY ............................................................... 94
  Purpose of the Study ........................................................................................... 94
  Research Questions ........................................................................................... 94
  Results .................................................................................................................. 96

CHAPTER V: DISCUSSION OF FINDINGS .............................................................. 113
  Introduction ........................................................................................................ 113
  Purpose of the Study ......................................................................................... 113
LIST OF TABLES

Table 1: Experimental and Control Teacher Post Results for Value of Formative Assessment Attributes........................................................................................................ 98

Table 2: Experimental and Control Teacher Post Results for Frequency of Use of Formative Assessment Attributes ........................................................................... 99

Table 3: Experimental Teacher Pre and Post Results for Value of Formative Assessment Attributes......................................................................................... 100

Table 4: Experimental Teacher Pre and Post Results for Frequency of Use of Formative Assessment Attributes ....................................................................... 101

Table 5: Demographic Information of the Control Group.............................................. 101

Table 6: Demographic Information of the Experimental Group ................................. 102

Table 7: Comparison of Experimental and Control Group Post Results for Academic Efficacy .......................................................................................................... 103

Table 8: Comparison of Experimental Pre and Post Results for Academic Efficacy..... 104

Table 9: Comparison of Experimental and Control Post Results for Eagerness to Learn ................................................................................................................... 105

Table 10: Comparison of Experimental Pre and Post Results for Eagerness to Learn ................................................................................................................... 106

Table 11: Test of Correlation Between Academic Efficacy and Clear Learning Targets......................................................................................................................... 108

Table 12: Test of Correlation Between Academic Efficacy and Progress Monitoring Information ............................................................................................................. 109

Table 13: Test of Correlation Between Eagerness to Learn and Clear Learning Targets......................................................................................................................... 110
Table 14: Test of Correlation Between Eagerness to Learn and Progress Monitoring Information ......................................................................................................................... 111
ABSTRACT

The main focus of the research was to determine if the systematic use of formative assessment attributes had an impact on two affective variables—Academic Efficacy and Eagerness to Learn. Both the control and experimental students’ affect was measured using a pre and post test administration of the Student Affect Inventory created by Popham and Stiggins (2008) in conjunction with the CCSSO. The post administration results of the experimental and control group were compared to determine if the systematic use of formative assessment attributes had a statistically significant impact on student affect. In addition, the pre and post test data of the experimental group was compared to determine if the systematic use of formative assessment attributes had a statistically significant change in reported affect by students as measured by the inventory. Inventories by all student participants (experimental and control) were examined for association between Academic Efficacy, Eagerness to Learn, Clear Learning Targets, and Progress Monitoring Information.

The study utilized a quasi-experimental design along with a study of correlation. Student in the experimental group (N=12) received systematic formative assessment attributes in daily instruction for 18 weeks. The teachers in the experimental group received additional staff development and were to use formative assessment attributes systematically in their instruction. Students in the experimental group did not show a statistically significant change in inventory data. While the experimental group did
exhibit some change, the population was too small (N=12) to show a statistically significant change. The results of the test of correlation of all student inventories (N=337) did show a strong statistically significant correlation between the affect variables and the formative assessment attributes. Additional research to examine impact of formative assessment attributes on a larger experimental population would be valuable. Additional research on the association of various formative assessment attributes with affect variables would add to the body of research in this field.
CHAPTER I
INTRODUCTION

Introduction

State and national assessments have their place in a balanced assessment system: to provide large-scale data about the performance of a large number of students from kindergarten to graduation. That information can then be used to make a judgment about the success a school did or did not have. That type of data does not apply directly to the student level. It is summative in nature. This level of data is useful only in making large-scale decisions about state educational programs and federal funding and to inform the public about the state of public education as a whole, by region, state, district, or school. The most recent assessment data from the National Assessment of Educational Progress calls attention to the need to improve the state of public education.

With the additional pressure added to school districts brought on by the No Child Left Behind Act, assessment results are becoming increasingly more scrutinized by both those in education and those in the public. The public release of large-scale assessments results calls attention to the need to improve public education nationwide. According to the Nation’s Report Card (Statistics, 2009), the assessment results for 4th grade and 8th grade students were well below the expectation. Fourth grade students have not improved math scores for 3 years, and 8th grade students have marked only marginally better. Only five states increased math scores for both tested grade levels. Reading
scores increased only four points since the national assessment began 15 years ago—30 states had no change in reading performance as measured (Statistics, 2009). In addition, the reading performance of 12th grade students has declined since 1992. All student groups showed a decline in reading performance, except the 90th percentile. Even with the major focus of NCLB on closing achievement gaps between subgroups, the achievement gap between racial subgroups has continued to increase (Grigg, Donahue, & Dion, 2007). While the United States claims to be a model country for education, the rest of the world has been catching and surpassing the United States in assessment results.

The comparative data of international assessments from around the world does not reflect the struggle the United States has had in education. Although the United States did not show a marked improvement in reading, eight other countries did—which improved the number of countries outperforming the United States in 4th grade reading ability to 10. Several Asian countries, including Japan, Singapore, and Hong Kong, along with 23 other countries have continued to outperform the United States in mathematics (Provasnik, Gonzales, Miller, & National Center for Education, 2009). The poor performance and lack of growth shown by the United States in reading and math, while many other countries continued to improve, called attention to the need for change in educational practice. Although much has been done to collect data, little has been done with using that data to make positive changes to assist academic learning. Large-scale testing is not the answer to improving academic achievement.

Academic reporting of student ability has increased the confusion and decreased the student motivation to learn. Grade reporting in most school districts is not an
accurate description of a student’s academic ability or knowledge. Most school districts continue to calculate non-academic factors along with academic factors to determine a grade (Marzano, 2006). Most school districts include factors that discourage student learning: late-work deductions, zeros, homework.

The current tradition of grading encourages a student climate focused on point collecting and finding right answers, as opposed to authentic learning. All activities become graded and thus used as a summative measure. Students are not encouraged to continue to learn after the assessment because the grade has already been given. Experts such as Ken O’Connor (2002), Douglas Reeves (2009), and Robert Marzano (2006) have suggested making significant changes to grading policies to encourage a climate focused on learning, instead of point collecting (Marzano, 2006; O’Connor, 2002; Reeves, 2009). Reeves (2009) went so far as to claim that the single most important change to affect student achievement was to change the report card to a standards-based grading system and the subsequent instructional shift on learning required to support such a change. O’Connor (2002) mirrored this remark and included several changes to individual grading, such as identify the learning target; use appropriate, high quality assessments that match the target; avoid using non-academic factors in determining a grade; do not score everything; allow students the opportunity to learn and show mastery even after the assessment; and—the most important—involve the students in the assessment process.

Students have learned to play the game called “School.” Students quickly discover that school is not a place to learn and discover, but rather a place to find the right answers. As Howard Gardner (2004) has said, “… schools everywhere have
embraced ‘correct answer compromises’ instead of undertaking ‘risks of understanding’” (Gardner, 2004, p. 141). Rick Stiggins (2007) has pointed out that historically, assessments have been used to identify differences in students and then rank those students accordingly. The current role of assessment elicits negative responses in students who are struggling academically. The student who receives negative summative assessment data feels hopeless and panicked, thinks that he or she is a failure, and is embarrassed. These negative reactions to summative assessment data encourages the student to seek out learning experiences that are easy, avoid new concepts or critical thought, and stay away from challenges, and the results promote giving up when they are not immediately successful. Stiggins has offered the use of formative assessment attributes included in instruction and daily classroom experience as a method to address student affect concerning assessments.

Much research surrounding the use of formative assessment attributes has supported the claim that they will improve student academic performance. Stiggins (2007) has claimed that the reason that formative assessment has this impact is that its philosophy is rooted in changing the student’s affect concerning the use of assessments and in including the student as the number one user of that data. Leahy, Lyon, Thompson, and Wiliam (2005) have supported this claim, stating that education needs to change its function of collecting rights and wrongs and to encourage teachers to collect information to inform instructional decisions. Likewise, students should also be provided that information to make decisions about learning. Leahy et al. (2005) made the comparison that current education is like quality control in manufacturing: at the end of
teaching, identify those who did not learn. They suggested that education should rather be a quality *assurance* process: collecting information through formative assessments to determine what needs to happen so each student learns (Leahy et al., 2005).

**Statement of the Problem**

Despite the increase in summative assessments of learning, little has changed in how students perform on national and state assessments. Current classroom instruction has not led to large gains in learning as measured by these assessments. The research has suggested that the added pressure of large-scale accountability assessments has pressured teachers and schools to provide test preparation for students, not genuine critical learning. Research has shown that teachers who used formative assessments to provide specific and timely feedback to their students have had a greater impact on their students’ academic achievement. Students who were provided with such feedback became more attentive and involved in the learning process and began to see the assessment process as a tool to help foster growth. Meta-analyses and early studies have supported, with large amounts of evidence, that using formative assessment in the classroom had a large impact on student academic achievement—especially for those students who were perennial low achievers (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 1998b; Crooks, 1988; Herman, Choi, National Center for Research on Evaluation, & Student, 2008; Kirton, Hallam, Peffers, Robertson, & Stobart, 2007; Leahy et al., 2005). Students tended to learn more and achieve at a higher level when they responded to assessment results by knowing where they were on the path to success, where they were headed, and what they needed to do to reach that accomplishment. The increased use of formative
assessment attributes by the classroom teacher on a daily basis has shown a positive impact on students’ affect toward the learning and assessing process.

The current practice of high-stakes testing given once a year to students to determine the performance that a school or district is making towards adequate yearly progress does little to improve the learning of students. These state-mandated assessments do nothing to provide information to those who need data to inform instructional decisions. The results are often reported 6 months later when those students have finished the school year and are now learning at another grade. The users of that information are preparing to teach students that were probably assessed the previous year on a different level of standards. For assessment data to impact student learning, it must provide immediate, specific feedback to the decision makers (Black & Wiliam, 1998b). The most important users of assessment data are the student and the teacher. This type of large-scale testing does not provide either of those individuals with the necessary information in a manner that will have a useful impact on student learning. For assessments to have an impact on student learning, the results need to provide immediate, specific feedback to the learner and the teacher to inform instructional decisions by the teacher and learning decisions by the student (Stiggins & Chappuis, 2005). Assessments that inform the teacher and student as to where the class or individual student is on the learning continuum and where the class or student need to progress have a positive impact on student learning when they are used during the process of learning—not 6 months later, when learning has stopped (Reeves, 2007). The most important response to the information gained from assessments is the psychological response of the student.
Large gains in student learning have occurred when the student received assessment information that helped identify where he or she was and where he or she needed to go on the learning continuum and that student responded to the information with the decision to keep trying (Stiggins & Chappuis, 2008). Ultimately it is the student’s reaction and choice to learn.

**Background of the Study**

Students have control over their learning. Despite a fantastic teacher, if a student chooses not to learn, then that student will not learn until he or she is persuaded to choose to learn. To collectively improve student performance, the individual student must become invested and involved in his or her own learning. Therefore, it has become increasingly more important to examine the factors that promote increases in student achievement. Although many studies have been conducted about the impact that formative assessment has had on student achievement, very little has been studied about the impact that formative assessment has had on student affect: eagerness to learn and academic efficacy. This researcher has postulated that this impact on student affect is the reason that formative assessment has such positive results in student achievement. This researcher claims that the impact that formative assessment has on student achievement is the result of the impact those attributes have on student affect. That change in student affect impacts the student’s desire and attention to authentic learning which improves student achievement. This study adds to the body of knowledge on formative assessment and the impact that formative assessment has on student affect.
With the added pressure for school districts to make adequate yearly progress as defined by the federal government, educational curriculum has shifted away from authentic learning to more skill, drill, and test prep. School districts are struggling to keep up with the level of adequate yearly progress (AYP), set by the federal government through NCLB. Because the AYP assessments are primarily for reading and math, most schools have placed a greater emphasis on curriculum in these areas. The reaction by school districts has been to cover what will be assessed. Unfortunately, this idea of coverage did not necessarily mean that students will have learned what was covered. The emphasis was on knowledge and fact finding not critical thought, investigation, and discovery.

To prepare students for the “big test” schools began to restructure their assessments to mirror the state assessment. The result was that students were exposed to even more emphasis on finding the right answer instead of learning by discovery and critical thought. More emphasis was placed on points and finding the right answer than on authentic learning. Larry Ainsworth (2007) described the amount of large-scale testing as a deluge placed on teachers and students. The stress over large-scale assessments pushed teachers to establish a pace of instruction that was often too much and too fast—resulting in coverage of materials, but a lack of student understanding (Reeves, 2007). Many experts have supported the use of assessments, but described these assessments to be effective only when they are short—cycle, quick assessments (Marzano, 2006; Popham, 2008; Reeves, 2007; Stiggins, 2004).
Research Questions

The purpose of this study was to examine the impact that systematic use of formative assessment attributes in classroom instruction had on student affect and to examine the association of formative assessment attributes in the classroom with student affect. The study sought to answer the following research questions:

1. Is there a statistically significant impact on students’ Academic Efficacy when the systematic use of formative assessment is used in daily classroom instruction?

2. Is there a statistically significant impact on students’ Eagerness to Learn when the systematic use of formative assessment is used in daily classroom instruction?

3. Is there a statistically significant association between students’ (control and experimental) perception of Clear Learning Target and Academic Efficacy?

4. Is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Academic Efficacy?

5. Is there a statistically significant association between students’ (control and experimental) perception of Clear Learning Target and Eagerness to Learn?

6. Is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Eagerness to Learn?
Significance of the Study

The current methods used in public schools contribute to students becoming more concerned with the behaviors associated with school, rather than the experience of learning. Current practices have taught students to be concerned with the amount of points an assignment is worth, completing assignments on time, and avoiding any unnecessary effort. In other words, students have become more concerned with point collecting than actually learning. They have become motivated by completing what is necessary to accumulate enough points to acquire the grade that may satisfy their need. Worse yet, struggling students are often not motivated by grades and points, but receive only poor grades—often reflective of a behavior such as not turning in the work as opposed to actually struggling with the assignment. According to much of the literature, students have a negative affect towards learning and assessing. They do not believe that assessments have any reflection on their actual academic ability. When teachers implemented formative assessment attributes in daily classroom practice, it had a positive impact on changing students’ affect towards learning and the assessing process. Students began to understand the nature of what was to be learned, associated assessment results as information that could be used to help improve their learning toward that target, and became more concerned with learning and less concerned with collecting points.

Many research studies supported that the use of formative assessments and those attributes have had a positive impact on student achievement. Although these research studies have included the student as the most important user of information and concluded that the student’s emotional response to assessments was important to that
increase in academic achievement, few studies have examined the direct connection with the use of formative assessment attributes and the resulting impact on student affect. Most of the research on formative assessment has showed the impact it had on student achievement, but little research has examined why formative assessment has had this impact on student achievement.

**Definition of Terms**

**Formative Assessment** is defined as: “…a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievement of intended instructional outcomes” (S. McManus, 2008, p. 3).

**Summative Assessment** is defined as any assessment that shows how the student performed against others or how many learning goals he or she has mastered at the end of learning. The information is used to assign judgment about a student’s performance.

**Formative Assessment Attributes** are the critical features used in the teaching process to make the formative assessment process effective.

**Progress Monitoring Information** is “… evidence-based feedback that is linked to the intended instructional outcomes and criteria for success” (S. McManus, 2008, p. 4). Descriptive feedback should be about specific qualities of student learning to continue the learning process through discussion or suggestions for improvement. It should be based on learning goals, progressions, or learning targets and not comparisons to other students and should help the student identify where he or she is on the learning continuum (S. McManus, 2008).
**Learning Targets** are the descriptions of skills, concepts, or knowledge that students are expected to learn.

**Academic Efficacy** is defined as “…the student’s perceived ability to succeed and the student’s sense of control over her or his academic well-being” (Stiggins, 2008).

**Eagerness to Learn** is defined as the student’s desire to learn.

**Limitations of the Study**

This study was limited by its sample size and narrow demographics. The number of participants limited the generalizability of the research findings to this population. The location of the study limited the use of an experimental design. The school setting allowed for convenience grouping based on student groupings through the school guidance office. Students had been placed into middle school teams through a computer program. The study thus used a quasi-experimental method. Because of the nature of the study, the intervention was applied through the current teachers. The teachers received training in formative assessment attributes and how to systematically use them in a classroom on a daily basis. Because the training was conducted by someone outside of the school, the amount and fidelity in which these teachers used the attributes were subject to self-reporting. The study lasted approximately 18 weeks; a longer study with a larger population would provide findings that are applicable to more populations.
CHAPTER II
REVIEW OF RELATED LITERATURE

History of Grading and Assessments

As Howard Gardner (2004) has stated, “…those who operate schools need to determine whether students under their charge are learning anything” (Gardner, 2004, p. 132). As the burdens on the school have increased, pressure to develop an efficient way to measure performance and student progress increased, resulting in the objective-format assessment also commonly referred to as the “test.” In modern society, the test (or assessment) is the method of determining who will receive the rewards of society (Gardner, 2004). Historically, assessments and grading have been used to measure the success of students or schools compared with other students or schools to place them in rank order. This process produces winners and losers. The mission of schools has changed from rank ordering of students to making sure that all students achieve at or above a minimum standard of achievement. This mission has been highlighted with the signing of the No Child Left Behind Act in 2002, which has placed a heightened sense of accountability on helping all students succeed in place of merely sorting, classifying, and placing them in rank order (Stiggins, 2007). Although the mission of education has changed the use of assessment has remained the same.

The first attempts to scientifically measure the social impact of education occurred in England during the Enlightenment. In the late 17th and 18th centuries,
scholars studied the order of nature through scientific methods. The way the world was and the way people that thought and reacted could finally be empirically studied through the fields of psychology and educational psychology. The American educational psychology was rooted in the European method: “…to come to understand how nature works and through such knowledge to control its operations” (Eisner, 2002, p. 196). The core beliefs of this movement were that nature was orderly and possessed a pattern that could be recognized, rational procedures could be used to discover natural regularities, theoretical ideas about those regularities could be constructed, and truth could be found. The last belief was an admiration for quantification. Measurement, rationality, theoretical explanation, predication, and control were the goals of this movement (Eisner, 2002).

In the latter half of the 19th century, the belief in these scientific methods led to the belief that educational practice could be guided by science and could be grounded in the understanding of how humans thought and learned. The first schools (developed in the 1850s) sought the rationale, brought by the new social science, to provide these institutions with an intellectual respect and legitimacy. This influence was mainly carried on through the test-and-measurement movement (Eisner, 2002). One of the results of this focus on test and measurement was the efficiency movement, started by Frederick Taylor as “scientific management,” which used the stopwatch to measure human performance to determine efficiency (Eisner, 2002). This movement, applied to education, led to the belief that success was measured by the speed at which children were exposed to material. The efficiency movement in education appeared as a conveyor belt of
education on which large numbers of students would undertake the same tasks presented in the same manner with presumably the same result. The unexpected result was a monotonous educational experience. In the early part of the 20th century, a factory inspector, Helen Todd, interviewed 500 factory child laborers. Of the 500, 412 preferred to work in the squalor of the factory than return to the monotony of school (Kliebard, 1986).

Before 1850, American schools were mostly one-room school-houses. Although assessments existed before this time, the outcome of an assessment was rarely a grade. Generally, the assessment was one of performance or oral defense, in which the outcome was either “pass” or “fail.” It was not until the mid-19th century that schools began to incorporate grades as a form of measuring student achievement. After the passing of compulsory-attendance laws, student enrollment and the number of schools increased dramatically. Although elementary schools continued to use description about student progress, high schools began to use mathematical scales—the percentage. The main reason for the use of this grading system was to aid teachers in tracking students. Students were organized by age and ability. When schools began to develop levels of high school, students in those same levels could also be compared with each other. Students would then be grouped with similar students with the premise of more efficient mass education (Hargis, 2003).

Although grades in high school served no function other than to organize students by academic ability, it did provide information for colleges. As competition for college
enrollment increased, colleges sought ways to screen applicants. High school grades were used again to sort and classify students (Hargis, 2003).

The validity of the 100-point scale was called into question as early as 1912 when Starch and Elliot published the results of a study that questioned the use of the percentage scale: Drastically different scores were assigned by numerous teachers on the same English paper (Hargis, 2003). This was the most prominent form of grading scale used then in American schools. The percentage scale was also combined with a letter grade of A, B, C, D, or F, which arbitrarily fell at different points on the 100-point percentage scale, which differed by school (Hargis, 2003).

In the 1920s, schools began grouping students by ability by using results from standardized intelligence tests (Hargis, 2003). Standardized achievement tests began at approximately the same time; however, they increased in popularity much faster. The use of the standardized achievement test was to identify the average performance for a given subject area. Schools encouraged their teachers to educate students to the median score. The results created even more ability grouping (Cureton, 1971).

In response to the need to assess large numbers of individuals quickly (beginning in the early 20th century), the multiple-choice format was created and used in assessments. Robert Yerkes’ work on the Alpha and Beta tests was commissioned by the Army to identify levels of literacy of enlisted men in World War I. It also helped to sort and classify men into or out of officers training (Eisner, 2002). The Army was in need of a way to ease the scoring of the large amount of soldiers who would take the tests. The Alpha tests relied heavily on the multiple-choice format in which correct answers were
given a point. The results were used to quickly and efficiently place soldiers in specific jobs based on their performance (Marzano, 2000). This view of assessment would be carried on to education and this stated purpose of assessment would remain in effect until the mid-20th century. Assessment would be used to classify students and schools by how they performed when compared with other students or districts. This view embraced assessments as a way to classify and order students. Eventually, assessments would be used for accountability—both for the student and for the school.

The belief that assessment should be used for accountability for schools’ impacts on school improvement began in the 1930s with the implementation and use of the “College Entrance Examination Boards,” which administered the Standardized Admissions Test (SATs). Although these tests were used to determine college admission, they were quickly adapted to be used for American education accountability. When the schools became supported by public money, the public also demanded to determine whether the money was spent wisely (Cureton, 1971). If the average SAT scores went up, then schools were doing better; if they went down, then schools were performing worse. This quickly established the driving force behind school improvement: raise SAT scores (Stiggins, 1999). The average SAT score for districts was one of the first methods to rank order states and their educational systems—a practice that still occurs today with a variety of tests (Stiggins & Chappuis, 2005; Stiggins, 2002). Not only has this practice been used to rank states, it can now be used to rank order districts and schools.

Norm-referenced tests, or tests designed to compare a student’s performance with that of his or her peers, were commercially developed in the 1950s and 1960s. Districts
began to initiate standardized testing to establish local accountability (Stiggins & Chappuis, 2005; Stiggins, 1999). Decisions about school improvement steps were made in attempts to raise these scores. It became evident during the educational reform efforts of the 1960s that standardized achievement tests could not accurately assess the outcomes of the new curricula. This reform movement saw the end of the sorting and rank ordering goals of education and shifted to catching up with the Soviets’ achievement in space. The movement called for students to think like scientists. The push was for students to understand the concepts, not just memorization of fact. The development of new curricula emphasized the need to evaluate the impact that it had on student learning. During the 1960s, Michael Scriven’s terms of “formative and summative evaluation” became popular as methods of examining student learning (Eisner, 2002).

Educational evaluation was focused on the quality of curriculum, the character of those activities, and the ease in which teachers could gain access to curriculum materials, not just on the measurement of student achievement. This shifted the view of student evaluation as a tool in decision making, not just in knowledge seeking. This practice was expected to enhance instruction (Eisner, 2002). Even with this new view of education and learning, policy remained in place to continue annual testing to measure the effectiveness of schools.

Statewide testing began in the 1970s with three statewide testing programs; by the end of the decade, more than 40 were being used. The 1970s and 1980s saw implementation of a national assessment program. The 1990s gave rise to the international assessments (Stiggins & Chappuis, 2005; Stiggins, 1999). With the creation
of the No Child Left Behind Act, states have felt even more pressure to perform well on statewide standardized tests. The federal government has increased the penalties for not making adequate yearly progress on these statewide tests. NCLB requires that every pupil in the United States in grades 3--8 and 11 be tested in mathematics and reading. The widely held belief has been that checking the status of achievement and reporting those results to the public will motivate and increase the speed of school improvement. Such tests and testing does provide educators with opportunities to reflect on what is and is not being achieved, but there is little evidence that these tests have had a positive impact on student achievement. The result has forced test giving as a matter of compliance with policy, rather than as a method for school improvement (R. Stiggins, 2004). As a result of this compliance, “…schools everywhere have embraced ‘correct-answer compromises’ instead of undertaking ‘risks for understanding’” (Gardner, 2004, p. 141).

In this country’s infancy, students were guided through learning and provided with a description of their ability and their growth. Education was not organized by age, but rather by ability, achievement, and need for growth (Hargis, 2003). Upon the development of levels of education and the advent of accountability, schools began finding ways to document and organize the achievement of students. Assessing students went from collecting information about ability to ranking students by achievement and performance (Eisner, 2002). The education system began collecting rights and wrongs and focused less on helping those who struggled to improve their learning. As the system changed to one that classified and ordered using grades and assessments, the participants
in this system also changed their response to grades and assessments. The summative nature of grades and assessments does little to improve the learning performance of today’s students.

**Current Perceptions of Grading and Assessments**

Research has been conducted examining the various stakeholders’ perceptions of academic reporting, grading and assessment practices. The following review examined the current practice of grades and assessment practices—most often used as a measurement of whether a student has learned or not—and the resulting impact those practices have on students. The majority of this research reflected the difference in perceptions of grades and assessments. The vast majority of grading and assessment practices were concerned with presenting information after learning was to have occurred and did not serve to continue the learning process. As described in the following review of the literature, grades and assessment practices vary tremendously. Marzano’s (2006) *Classroom Assessment and Grading that Work*, lists four classroom assessment best practices:

- Feedback from classroom assessments should give students a clear picture of their progress on learning goals and how they might improve.
- Feedback on classroom assessments should encourage students to improve.
- Classroom assessment should be formative in nature
- Formative classroom assessment should be frequent.

Unfortunately, most often the reviewed current practices do not resemble Marzano’s prescription and cannot fulfill their above-stated purpose because of the lack of
Thomas Guskey (2005) conducted a study of 314 educators from three states about teacher and administrator perception of the validity of a variety of evidence of student learning and achievement. Both groups agreed on the value of student portfolios, teacher-created assessments, compositions, and writing assessments. Teachers rated observations, homework completion and quality, and student behaviors and attitudes as most valid in determining student achievement. On the other hand, administrators believed district assessments, state assessments, and nationally normed assessments to be more valid predictors. It became apparent that teachers valued their classroom assessments as more valid than third-party assessments (Guskey, 2005). The student receives more frequent teacher assessments than larger district, state, or nationally normed assessments. The frequency determines that teacher-created assessments have the potential to impact student learning. Paul Black and Dylan Wiliam (1998b) stated from their meta-analysis that the largest impact on student achievement is frequent, immediate, short cycle feedback to students in the classroom. With the evidence that supports the impact that valid, frequent, and local assessments can have on student achievement, it is increasingly important that grades and assessments provide accurate feedback that can be interpreted correctly by all stakeholders. A review of the literature showed a lack of such formative assessment practices in current school districts. Grading practices lacked accuracy by including multiple factors such as behaviors and were not interpreted by the different stakeholders in the same way.
In her study presented at the Annual Meeting of the American Educational Research Association, Kathryn Davinroy (1994) presented her findings on the different perspectives on grades between teachers and 3rd grade students. According to Davinroy, teachers viewed grades and assignments as valid measures of student academic performance related to the concepts taught, while the students in the study viewed their resulting grades as a reflection of the neatness of their work. Davinroy posited that if students are unable to recognize that they are assessed on their understanding of mathematics, then they will be unable to value understanding (Davinroy et al., 1994).

In similar studies, Thomas Guskey (2002) and Janet Carlton (1992) investigated the differences in the perceptions of grades and grading practices of three stakeholder groups: teachers, students, and parents. In Guskey’s study he reported the perception of the purpose of grading. Although all stakeholders expressed similar ranking on a given list of words and phrases, teacher responses were much more varied than parents or students. Teachers also valued grades as incentive much lower than both parents and students. Parents, students, and teachers all ranked the importance of grades as a measure of communicating information about student achievement and behavior to students as more important as the years in school increased. As the grade levels increased, teachers valued grades as incentive increasingly lower, while parents and students ranked them as an increasingly more important incentive. In examining the rankings of performance elements—the elements used in determining grades—students and teachers ranked homework completion as an important factor in determining grades. Likewise, teachers
viewed homework quality (accuracy and correctness) as equally important, while students valued homework quality as significantly less important (Guskey, 2002).

Although Guskey’s (2002) study illustrated the difference in value that grades hold for the different stakeholders, Carlton’s (1992) study highlighted the differences in what the primary participants think grades represent. Carlton’s study illustrated the difference in interpretation of grades by the various stakeholders. Parents of high school students responded that they knew how to interpret grades only as a reflection of academic ability. However, Carlton reported on the factors that teachers include in 6-week grades. Eighty-three percent of teachers reported including effort in assigning grades to students. Carlton continued by illustrating the connection of behavior with grades and the misconception that parents have about final grades. Parents in this study believe that letter grades are reflective of only academic ability and that a student who received a “C” was an average academic student. The reality is that it was quite possible that in this scenario a student was excelling in academic understanding, but was receiving an average grade due to behavior such as turning in assignments late. The parents involved in this study did not approve of this grading practice. Parents reported that grades should represent academic achievement and that if other factors are included (such as behavior) then interpretation of the grade is impossible. The teacher responses further illustrated the connection of behavior and the disconnection of academic achievement to grades: the teachers responded that most students who apply effort do not fail. Student participants in this study acknowledged the impact of effort on grades. In describing the difficulty in interpreting grades, student participants pointed out that each teacher grades
assignments differently and factors effort into struggling students’ grades more than into those of high-achieving students. Carlton found that there was no discrepancy between students and teachers about the expectation to achieve a certain grade, but there was a discrepancy between teachers and parents. Both parents and students expressed disagreement with the method that teachers used to assign grades.

Although Carlton’s (1992) study identified the difference in how parents interpreted grades and teachers determined grades, Susan Austin and Richard McCann’s (1992) study evaluated the grading policies and procedures of 144 school districts in a northeast state. The study was an investigation into the validity of grades and the potential conflict that grading policies and procedures may have on state educational reform efforts. In the study, parents were identified at each of the organizational levels as the primary audience for grade reporting, followed by students as second and teachers as third (Austin, McCann, & Research for Better Schools, 1992). According to Carlton (1992), parents interpreted grades differently than both students and teachers. If the primary audience of grades is parents and the primary role is to inform parents about their children’s academic ability then school districts are not achieving this goal with any consistency. If the role of grade reporting is to accurately inform parents and students about students’ academic ability, then behavior such as turning in assignments late or not at all cannot be factored into a grade. Grades have been reported to be the key to informing parents and students about students’ academic performance but often they have more than academic implications. As Ken O’Connor (2002) has advocated, grades and
grade reporting should be an accurate reflection of academic ability and should not factor in such things as behavior.

An additional difference in perception occurred within the same group of stakeholders. The academic status of a student affected his or her perception of grades. The high-achieving student attributed high grades to personal ability, while the low-achieving student attributed grades to how well the teacher liked a student (Carlton, 1992). Therefore, grades were not a valid or reliable form of disseminating information about academic ability to students or parents. Students of varying academic ability interpreted grades in different ways, which impacted students’ behavior and perceptions in different ways.

Black, Swan, & Wiliam (2006) have examined students’ beliefs about learning from a larger study on The Learning How to Learn project. The findings presented suggested that students did not see the connection between their involvement in learning and particular school practices. Students were unable to articulate any connection about their learning to what occurs in the school setting. They attributed getting good grades to behavior such as avoiding conflict, attendance, and turning in assignments on time (Black, Swann, & Wiliam, 2006). High school students from VanderHeide’s (1994) qualitative study on student perceptions of test grades displayed a frustration with their belief that knowledge was not accurately assessed through classroom assessments. All participants displayed a disparity in how accurately they viewed grades to represent personal assessment. Her study articulated the belief of students that a student may know the answer or concept but is unable to present that information in the exact way the
teacher required for the test (VanderHeide, 1994). In these studies grades were not an accurate reflection of student academic ability or understanding or an effective device to communicate to the student or parent information helpful in making academic decisions.

Findings from VanderHeide’s (1994) qualitative study on student perceptions of assessment grades showed that grades had a profound relationship on how successful or unsuccessful the students in this study felt. The findings also presented the importance that these students placed on behavior to achieving high grades, which thus took away from their understanding that grades were a reflection of academic learning. VanderHeide stated “Consequently, competition to obtain good grades overshadowed the learning experience” (VanderHeide, 1994, p. iv). Students were more concerned with their ability to gain a high grade than to learn new information or concepts. The students involved in this study remarked that teacher-prepared tests emphasized recall and allowed little opportunity for insight. The students interviewed also said that the tests were lacking in validity and did not assess an individual’s intelligence. The students stated that tests did not measure what they knew. All interviews showed that students perceived grades as a way to rank and classify students; however, all participants agreed that the meaning of grades was more than a reflection of academic ability.

Austin and McCann (1992) supported this belief: They reported that the five factors used to determine grades were student performance, class participation, attendance, attitude, and discipline. Of these criteria, student performance was cited as the most used in the studied districts, but all other factors were listed as impacting grades in all levels of the organizations. Seventy-five out of 90 districts participating in the
study provided information that showed that they asked their teachers to apply multiple criteria to determine grades. In addition, the responding districts also showed that staff received little to no in-depth professional development on grading and grade reporting and none of the districts reported providing staff development to help teachers grade with consistency. Austin and McCann (1992) speculated that districts were attempting to create grading policies that addressed all stakeholders’ views of the purpose of grades: reflection of achievement against standards, describing the effort and progress students are making, and student performance as compared with that of other students (Austin, et al., 1992). In attempting to address the various perceptions of the meaning of grades, school districts were reducing any validity in the accuracy of grades as a reflection of student academic ability.

Susan Brookhart (1993) conducted a study of various grade-level teachers’ beliefs and values in grades and grading. In her findings, she presented the conclusion that participating teachers viewed grades as something to be earned as compensation for work done. Although achievement was part of the construct it was not the whole. Her findings also suggested that grades were used as a large part of classroom management as the reward for the desired behavior.

Teachers were willing to change grades and procedures about grades depending on the individual student. Teachers in multiple studies (Brookhart, 1993; Carlton, 1992; VanderHeide, 1994) reported manipulating the grade of students who academically were failing, but exhibited positive behaviors in class. For example, a teacher who observed students who achieved a failing grade but appeared to “work hard” would most likely
inflated the grade to a passing mark. On the other hand, teachers also reported that they were more likely to give a true academic grade achieved by higher students. Teachers have a dual role as advocates and judges of students, and because of this dual role, concern over student consequences will trump grading interpretability, and teachers will mix these roles depending on students of differing abilities (Brookhart, 1993).

According to Davinroy (1994), “If reform is to make a difference, it must gain access to the perceptions of those it seeks to educate” (Davinroy et al., 1994, p. 3). According to the review of literature, if grades are to be used to make decisions about learning by any of the stakeholders then grades need to be perceived to have the same meaning. According to Marzano (2006), grades were so imprecise that they were virtually meaningless to all who interpreted them (Marzano, 2006). Because of this discrepancy in grades and grading the results from these differing perceptions impacted the students in vastly different ways from those intended.

As a negative result of grades used as a motivational factor, students have shown an increase in motivation to achieve a higher grade but not necessarily to achieve a higher level of learning. Feedback in the form of grades and assessment of whether a student has learned the subject matter or not (summative assessment) were motivating students to attain better grades but not necessarily to learn more or at a better quality. Carlton (1992) found that in student conversations they wanted the grading system that assigned them the highest grade. Teachers also agreed that the inconsistency of grades produced a negative consequence for students.
The confusion, inconsistency, and misconception connected with grading and assessments appeared to have an unintentional consequence on students’ affect. It can be widely agreed that the optimal result of grading and assessment practices should be to improve the academic performance of students, but in the following review of the literature, the response reported showed a more widely negative response to grading and assessment practices.

Gavin Brown and Gerrit Hirschfeld (2007) made claims from their study that supported self-regulation theory—the belief that an individual has control over his or her own learning through the decisions that he or she makes. In their study of New Zealand students, they showed evidence that mathematics scores increased when students viewed assessment as a way to hold students accountable for learning and as beneficial. In contrast, students who viewed assessment as interfering with learning or as being ignored by the teacher showed lower achievement in mathematics. Daniel Ragland (2008) presented the importance of student perception and its effect on achievement in Hispanic students. He presented in his major findings that academic self-perception—how the individual views his or her academic ability—was the only attitude variable that had a significant effect on proficiency level. Ragland’s findings also posited that reported higher motivation and academic self-perception levels had a positive correlation on GPA level. He also presented that perception was the strongest predictor of proficiency level. Based on his findings, Ragland recommended that all stakeholders should focus on improving students’ positive academic self-perceptions and students self-motivation.
VanderHeide (1994) displayed the emotional impact that grades had on students’ self-efficacy. In interviews, when students discussed high grades, they displayed enthusiasm in their tone of voice and demeanor; when recalling poor grades they tended to avoid eye contact, their shoulders drooped, and their voices dropped or became quiet. The students who enjoyed competition thrived in this climate, but for those who were forced to compete, grades created feelings of discouragement and disillusionment. Fear of grades motivated only those students who were concerned about success as measured by grades. In other words fear of grades motivated only those motivated by grades. Generally they were not the struggling students (Chappuis & Stiggins, 2002). Success or positive feedback through grades resulted in a sense of fulfillment, enjoyment, and an increase in motivation. The repeated exposure to poor grades created a negative view that studying and effort would have no impact on the pattern of failure and it reinforced a negative behavior (VanderHeide, 1994). One student from VanderHeide’s (1994) study commented “that it was better to seem not to care than to appear stupid by doing poorly” (VanderHeide, 1994, p. 74). Even though students consistently displayed a belief that teacher-created tests were poorly designed and inaccurate, they placed blame for poor grades on personal work ethic. The students in this study expressed the opinion that grades should be used as a means of feedback to students to promote further learning by them (VanderHeide, 1994). Feedback about where students were on the learning continuum has shown to have a positive effect on all students and specific immediate feedback had a profound effect on the struggling students (Black & Wiliam, 1998b; Stiggins, 2004).
The use of grades has been shown to impact students’ motivation and focus of motivation. High-achieving students are motivated through the fear used with grades while struggling students experience the opposite effect. Increasingly more detrimental is the shift in the student’s focus in his or her motivation. Students become increasingly more dedicated to executing behaviors that achieve a desired grade. In most cases, those behaviors are not connected or are coincidentally connected to learning. For example, students become more concerned with finding answers they believe the teacher is looking for than creating knowledge. Students also show an increase in the perceived correlation between completion of assignments on time and neatly done with higher grades rather than the quality or understanding of the concepts as displayed through the assignment.

As VanderHeide (1994) has pointed out in her study through a comment by one of her participants, “I don’t think I really worry about what I’ve learned in the past three years, except I worry about what my marks are” (p. 90). Student desire to determine what the teacher wants as an answer, as a means to achieve a better score, has surpassed the personal learning experience.

As described in the review of literature, a lack of common understanding exists in the interpretation of grades (Austin et al., 1992; Carlton, 1992; Davinroy et al., 1994; Guskey, 2002, 2005). There is also a lack of commonality in what factors are included in determining a student’s grade (Austin et al., 1992; Brookhart, 1993; Carlton, 1992; Guskey, 2002). The variance in determining grades and interpreting grades decreases the usefulness of grades (Marzano, 2006). This disconnection results in a variety of negative consequences for all stakeholders (VanderHeide, 1994). The research supports that the
largest impact on student academic improvement occurs with frequent and accurate feedback (Black & Wiliam, 1998b; Davinroy et al., 1994; S. M. McManus, 2008). The research also supports that grades, either summative or as letter grades on assignments are neither accurate measures of academic ability nor do they provide necessary feedback to any of the stakeholders (Brookhart, 1993; Carlton, 1992). The purpose of education has shifted from classifying students in rank order to providing an education for all to meet or exceed the minimum standard; however, our assessment and grading practices have not made the same shift. Brookhart (1994) posited in her review of teachers’ grading practices and theories, “Grading theory and practice will be better connected once the role of classroom assessment and grading practices in student achievement motivation and classroom management is understood” (p. 279).

Need for Change

With the creation and reauthorization of No Child Left Behind, school districts, research firms, psychometricians, and professional development providers have been seeking innovations to improve the teaching and subsequent learning of students in public education to meet the high standards set for adequate yearly progress. According to the Nation’s Report Card (2009), fourth grade students have not increased math scores on the National Assessment of Educational Progress (NAEP) since 2007, and 8th grade students have marked a marginal increase of only 2 points. Only five states showed gains in both 4th and 8th grades. Most student demographic group gaps remained (Statistics, 2009). Fourth grade students showed only a 4-point increase in reading as measured by the NAEP since the first assessment 15 years ago. Eighth grade students showed only
marginal gains of three points since 1992. Only four states showed gains in both grade levels for reading, and 30 states showed no significant change in either grade level (Lee, 2007).

Although the performance of the nation’s 4th and 8th grade students has not increased significantly in the last 15 years, the academic performance of 12th grade students was more alarming. Overall reading performance, as measured by the NAEP, declined in comparison with the assessment results from 1992. Reading performance showed declines for all students except the top performers (90th percentile). The percentage of students performing at or above the basic level decreased from 80 to 73 percent. The percentage of students performing at or above the proficient level decreased from 40 to 35 percent. The gap between reading performance of White and Black students increased, while all other gaps remained unchanged since 1992. Less than one-quarter of 12th grade students tested in mathematics performed at or above the proficient level (Grigg, Donahue, & Dion, 2007).

According to the international assessments of 2001 and 2006, the United States did not show an increase in reading performance for 4th grade students; however, eight other countries improved bringing the number of countries outperforming the United States to 10. Although 4th grade U.S. students improved in mathematics, the countries of Hong Kong, Japan, and Singapore consistently outperformed the United States in 4th grade and 8th grade. Twenty-three countries outperformed U.S. 15-year-olds in mathematics (Provasnik et al., 2009).
Unfortunately, as the United States remains relatively unchanged in academic performance, we continue to misuse classroom and local assessments as a tool for collecting “rights” and “wrongs” instead of as an instructional tool. According to Howard Gardner (2004), “…formal testing has moved too far in the direction of assessing knowledge of questionable importance in ways that show little transportability…quite different forms of assessment need to be implemented if we are to document student understanding” (p. 134). Stiggins (2004) adds to Gardner’s sentiment that “the belief in the power of standardized testing has blinded public officials and school leaders to completely different application of assessment—day-to-day classroom assessment—that has been shown to trigger remarkable gains in student achievement” (p. 23).

Currently, the most common use of assessments is as a summative measure. Summative assessment is a measurement at the end of student learning to determine how much or many of the intended goals the student has learned or how many students have learned those goals. It is an assessment after learning has stopped. Formative assessment, on the other hand, has been defined as “…a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievement of intended instructional outcome” (S. McManus, 2008, p. 3). Although summative is an assessment of learning, formative is an assessment to help students continue to learn. The use of formative assessment has shown profound gains in student achievement, and the most profound are in the achievement of struggling students. With the increasing pressure on school districts and states to show increases in student academic performance, educators need to shift their
schema to one that uses methods that increase and encourage student learning not merely measure it.

**Formative Assessment as Possible Answer**

Formative assessment is not a new term and can be defined in many ways. Black and Wiliam (1998b) defined assessment as “…all those activities undertaken by teachers—and by their students in assessing themselves—that provide information to be used as feedback to modify teaching and learning activities” (p. 140). This definition does not limit itself to formal tests, quizzes, or homework. Assessment is a collection of evidence about student learning through a variety of ways such as portfolios, journals, dialogue, questioning, interviewing, work samples, formal testing, and projects. They defined formative assessment as “such assessment…when the evidence is actually used to adapt the teaching to meet student needs” (p. 140). The key difference between summative and formative assessment is what is done with the information. Summative uses the information to show how the student performed against others or how many learning goals he or she has mastered at the end of learning. Formative assessment uses the information collected to determine where the gap of learning is for the student and then is used to determine how to close the gap.

Stiggins and Chappius (2006) explained assessment for learning as a formative assessment philosophy that involves the student in their assessments by giving the students clear classroom-level targets based on state or local standards. Those targets are then transformed into dependable and accurate assessments. The vision of the successful outcome is shared and understood by the students through models of success and quality
work and or the use of descriptive rubrics. The teacher generates feedback, either written or verbal, that describes where the student is on the learning continuum of that target and provides specific communication to the student on how to narrow that gap.

This study used the definition provided by the Council of Chief State School Officers (CCSSO) in which formative assessment was defined as: “…a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievement of intended instructional outcomes” (S. McManus, 2008, p. 3). The CCSSO also included the five critical attributes of formative assessment: learning progressions, learning goals, descriptive feedback, self- and peer assessment, and collaboration (S. McManus, 2008).

The first wave of research on formative assessment focused on the collection of information by the teacher, school, or district to make systemic changes in curriculum or instruction for the previous year. The focus was on the teacher as user of the information collected. The first major research findings on this level of formative assessment were presented from the meta-analysis by Terry Crooks’ (1988), which articulated the effect that formative assessment can have on instruction and thus academic achievement.

Since Crooks (1988), the study of formative assessment and the student-centered version often referred to as “assessment for learning” was brought to the education community again 10 years later in a second wave of interest. This second wave of interest and research on formative assessment not only examined the teacher as a user of formative assessment but also focused on the student as the primary and more important user of the collected information.
The paramount findings from this second wave on formative assessment were attributed to Black and Wiliam (1998a), who published the results of an extensive meta-analysis on assessment and classroom learning. They presented evidence, from numerous studies, that supported the use of frequent feedback to students about their learning and that such feedback can aid in large learning gains for the individual student and class. They also examined the role of student self-assessment and peer assessment alongside formative assessment strategies used by teachers. Their meta-analysis comprised a collection of 681 publications, and 250 of the original collection were selected. The results of this selection were published in summary in a later article, “Inside the Black Box: Raising Standards Through Classroom Assessment” (Black & Wiliam, 1998b).

Although the terms “formative” and “summative” assessments or evaluation have been around since the 1950s with Benjamin Bloom and the 1960s with Michael Scriven, these two studies reignited attention to the impact that formative assessment strategies can have on teacher instruction and student learning.

Although Crooks (1988) and Black and Wiliam (1998a) have presented compelling arguments for the use of formative assessment in public education, it is an area that has not been studied extensively since the publishing of Black and Wiliam’s findings. Subsequently, the use of formative assessment and student inclusion as a decision maker and a user of assessment information, are rare in America’s public educational systems and the research of its use is likewise as rare (Herman et al., 2008). However, in the cases and studies that have since been reported, the impact of formative assessment attributes has mirrored the effects originally published by both Crooks (1988)
and Black and Wiliam (1998a). Although the majority of these studies have focused on the impact on test scores or student achievement, little has been studied on the effects such a shift in assessment would have on students’ eagerness to learn and academic efficacy.

**Research on Formative Assessment**

The following section shows evidence of the impact that using the attributes of formative assessment can have on students and their academic achievement.

Terry Crooks (1988) in a meta-analysis of studies on classroom evaluation practices has summarized results from 14 specific fields of research to clarify the impact between classroom evaluation practices and student outcomes. Crooks’ review synthesized research as related to the impact of classroom evaluation on students. Crooks defined classroom evaluation as “…evaluation based on activities that students undertake as an integral part of the educational programs in which they are enrolled. These activities may involved time spent both inside and outside the classroom. This definition includes tasks such as formal teacher-made tests, curriculum-embedded tests (including adjunct questions and other exercises intended to be an integral part of learning materials), oral questions asked of students, and a wide variety of other performance activities (cognitive and psychomotor)” (p. 467).

Crooks (1988) has summarized his findings about the importance of classroom evaluation as it affects students. Based on his evaluation of research, he found that classroom evaluation guides students’ judgment of what is important to learn, affects their motivation to learn, forms their self-perception of competence, helps them make
decisions about what and how much to study, consolidates learning, and impacts the development of their learning strategies and skills. Crooks posited that classroom evaluation “…appears to be one of the most potent forces influencing education” (p. 467).

Crooks’ (1988) evaluation of research uncovered that the practice of classroom evaluation relies heavily on recall of isolated bits of information, but research has repeatedly shown that such fragments or details are readily forgotten without a context or broader framework. More concerning is the focus of such evaluation on knowledge base information when, according to Crooks’ examination, accumulation of knowledge is less important than learning skills and habits. The research that Crooks examined displays a disparity between the importance placed on higher-order thinking and transference of learning and the evaluation of such thinking.

Based on their extensive meta-analysis, Black and Wiliam (1998b) have supported that innovations that strengthen the use and practice of formative assessment produce learning gains. They have cited such substantial and profound learning gains in studies in which the participants range from 5-year-olds to undergraduates and range over several school subjects and countries. The results of such studies reported a typical effect size between .04 and .07. Black and Wiliam pointed out that such effect sizes are larger than most effect sizes reported for educational interventions. They continued by clarifying the impact such an effect size would have on the United States. An effect size of .07 would change the status of the United States from the middle of 41 countries in mathematics to one of the top five.
The most dramatic of the findings reported by Black and Wiliam (1998b) was the impact that these strategies had on struggling students—those with learning disabilities and low-achieving students. The results showed that frequent and specific feedback yielded substantial gains in both groups of students, with the greatest gains for low-achieving and learning-disabled students. Although formative assessment has been shown to have a large positive impact on all students it yields substantial impact on low achievers by concentrating on specific problems they are having difficulty with, providing them a clear understanding of where they are in their learning and providing a clear understanding of what needs correcting and how to correct it (Black & Wiliam, 1998b).

Black and Wiliam (1998b) articulated the nature and extent that formative assessment should be used in the field of education. Their studies showed that the primary user of assessment information to promote and improve learning is the student; however, the student has responded to the current educational system by focusing on “rewards,” also known as “grades” or “class ranking.” The student is encouraged by the collection of more grades or points. Students are avoiding authentic learning for fear of poor grades or less points and pursue finding answers instead of generating answers (Black & Wiliam, 1998b). It is necessary to refocus students on learning and away from point-collecting or reward-seeking behavior.

Black et al. (2004) followed this examination of research with a research study of 19 secondary school teachers and their students in the United Kingdom. The study conducted used the suggestions from Black and Wiliam’s (1998b) meta-analysis, applied
them to a school setting, and measured the results on student achievement. The main interventions of this study were questioning, feedback through grading, peer and self-assessment, and the formative use of summative tests. The result of this experiment was an average effect size around 0.3 standard deviations in a variety of externally administered standardized achievement tests (Black et al., 2004).

Wiliam, Lee, Harrison, and Black (2004) have conducted research examining the impact that formative assessment practices of 24 teachers had on student achievement in schools in the United Kingdom. The intervention was several full-day and half-day workshops provided to the teachers about formative assessment practices. The teachers were then observed throughout the course of the year, and their curriculum and lesson plans were also examined to determine the extent that formative assessment strategies were used in instructional planning. The quantitative results of achievement scores for students taking the local standardized assessment used by the school and the graduation exit exam known as the “national school-leaving examination” (GCSE) from previous or tandem classes showed a statistically significant increase in the average score of students in the various courses. The results showed an impact on achievement scores on external assessments or assessments created by an outside agency such as national standardized tests (Wiliam, Lee, Harrison, & Black, 2004). The authors concluded by explaining the impact such an increase would have on a large-scale inclusion: If used in a full school setting these results would raise a school in the 25th percentile to the upper half (Wiliam et al., 2004).
The results from the August 2008 CRESST (Center for Research on Evaluation, Standards, and Student Testing) Report 740 supported the use of formative assessment and the effects that its use has on student achievement. The study employed a model of formative assessment that used the components of specified goals for student learning (targets of learning), frequent formative assessments aligned with goals or targets, and instructional decisions made from formative data. The study focused on one area of the model: the quality of teachers’ interpretation of assessment results and how the accuracy of teachers’ judgment would affect student performance. Teachers in this study were asked to establish learning goals or targets, assess students on goals or targets, give goal- or target-based instruction, and use assessment data to make changes in instruction. The study examined the accuracy of teachers’ prediction of student achievement and the relationship to middle school learning. Analysis of results showed a consistent, positive relationship between teacher accuracy and middle school student learning (Herman, et al., 2008).

This study involved seven experienced middle school science teachers from districts across California in the implementation of a unit on buoyancy from the Foundational Approaches in Science Teaching (FAST) curriculum. The unit used formative assessments embedded in daily instruction. The teachers received intensive sustained training and support to use formative assessment strategies. The study examined teaching logs, pretest and post-test data, and teacher judgment data compared with formative assessment data. The data showed a strong correlation between the accuracy of a teachers’ perception of what percentage of her class was on target with the
expected level of understanding and the overall increase in student achievement. This study suggested that teachers who collect formative data and use such data to inform their instructional decisions have a larger impact on student achievement (Herman et al., 2008).

A study of the impact that formative assessment had on some Scottish primary and middle schools provided evidence that the use of formative assessments has a positive impact on increase in student responsibility for their learning and improved motivation, confidence, and classroom achievement. Kirton et al. (2007) studied the impact of Project One of the Assessment Is for Learning Development Programme in 16 Scottish primary schools and two middle schools. Teachers were given strong professional development in formative assessment practices through workshops, learning communities, and support from Scottish national learning and education agencies. Teachers were given the opportunity to choose which formative assessment practices they would use and monitor. The study sought to discover the extent that this project was perceived to have on classroom practice; improved student learning, motivation and behavior; change in teachers’ beliefs, attitudes, and understandings of assessment; school climate; and parental interest and involvement in their child’s education. The study collected data through self-evaluation, examining action plans, teacher journals, case study reports, field visits to ensure validity of documented evidence, interviews of staff and students, and classroom observations. The collected generated results indicated that the project was perceived by the participants to have had a positive impact on students, teachers, and pedagogy, but little impact on involving parents. Final evaluations declared...
that all 33 schools perceived the project to be successful. All of the collected data suggested that 14 schools appeared to have embraced the strategies, 14 seemed to have made adequate progress, and 5 seemed to have gained less (Kirton et al., 2007).

Research conducted by Lisa Smith (2008) reported that frequent formative assessments can predict achievement on measures of Adequate Yearly Progress indicators in math as measured by standardized criterion-referenced competency tests in the Gainesville School District in Georgia. Smith conducted research in one school district of 2,900 middle school student scores over three years. In examining the data, the scores of posttest formative assessments given on a quarterly basis were shown to accurately predict increases in the state’s AYP measurement test. For every one unit increase in quarterly score on the post-formative assessment, a positive gain in student achievement could be predicted. The research model correctly predicted participant 84.87 percent of the outcomes (Smith, 2008).

Although the majority of research has been conducted on the most stressed areas of needed improvement—math and science—according to a study by Christian Colby-Kelly and Carolyn E. Turner (2007), research has also shown the impact that including formative assessment attributes has had on the second language classroom for pre-university students. Colby-Kelly and Turner reported on the results they collected from nine teacher and 42 student participants. The 42 students all reported a variety of original languages other than English and were enrolled in pre-university classes in England. The research questions were summarized in this way: What are the teacher and student perception, the nature, and the evidence that formative evidence benefits learning in a
second language classroom setting? Colby-Kelly and Turner’s research findings suggested that teacher-student feedback with a motivational component appeared to be effective in motivating some English language learners to focus on learning.

According to interviews and questionnaires, teachers demonstrated that they were in strong favor of using formative assessment practices in their classrooms. The surveyed teachers also agreed that student involvement in assessment was positive and that self-evaluation and feedback fostered learning; however the teachers were not in agreement on whether students believed that assessments contributed to learning. All but one teacher agreed that assessments and teacher comments did impact student learning (Colby-Kelly & Turner, 2007).

**Research on the Five Attributes**

This study used the definition provided by the Council of Chief State School Officers (CCSSO) in which formative assessment was defined as “…a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievement of intended instructional outcomes” (S. McManus, 2008, p. 3). The CCSSO also included the five attributes critical of formative assessment: learning progressions, learning goals, descriptive feedback, self- and peer assessment, and collaboration (S. McManus, 2008). The following section describes each of the five attributes and presents the relevant research related to each attribute.
Learning Progressions

According to the CCSSO (2008), “Learning progressions should clearly articulate the sub-goals of the ultimate learning goal” and “[l]earning goals and criteria for success should be clearly identified and communicated to students” (S. McManus, 2008, p. 4). Kirton et al. (2007), Herman et al. (2008), Chappuis and Stiggins (2002), and Black and Wiliam (1998b), claimed that students who had clear targets of learning have become more committed and more effective learners.

Popham (2008) described learning progressions as “…a sequenced set of subskills and bodies of enabling knowledge that, it is believed, students must master en route to mastering a more remote curricular aim.” It is the design and sequence of the sub goals or “building blocks” of learning leading to the larger learning goal or target of learning. Targets of learning can be large in scope such as a target curricular aim—usually a higher-level skill or concept, or they can be used as subskills—levels of understanding and ability that are necessary to learn or perform the target curricular aim. Popham described that learning progressions are important in the design of what a student must master to continue moving forward or toward the larger curricular aim or higher-order skill or concept. It is also important to design learning progressions to include formative assessments, which, if used in a learning progression, will help determine where on the learning continuum toward that target curricular aim and will help the teacher provide specific feedback to guide the student on the course of attainment. Popham described four levels of formative assessment:

- Level 1: Teachers’ Instructional Adjustments
• Level 2: Students’ Learning Tactic Adjustments
• Level 3: Classroom Climate Shift
• Level 4: Schoolwide Implementation (p. 49)

Level 1 includes the information a teacher collects to make instructional decisions (this form is the most common). Level 2 views the student as a data decision maker. Popham is clear that level 2 requires the presentation of clear learning goals or targets of learning to the student in a manner that the student can understand and view as attainable. Stiggins (2004) described these targets as “student friendly” and stated that students need formative feedback about where they are in progress toward that goal. According to Popham (2008), including level 1 (Teachers’ Instructional Adjustments) and level 2 (Students’ Learning Tactic Adjustments) in classroom instruction will have a positive impact on level 3 (Classroom Climate Shift) shifting it to one of an atmosphere of collaboration.

According to Margaret Heritage (2008) in a paper prepared for the Formative Assessment for Teachers and Students (FAST) state collaboration, learning progressions are an essential component of effective formative assessment. She described learning as a trajectory of development. Heritage defined learning progressions similar to Popham’s (2008) definition and acknowledged that the empirical research was lacking. Heritage (2008) concluded that those involved in constructing learning progressions would need to draw on existing research on learning, curriculum specialists, and experienced teachers to create clear paths of learning. Once these constructs had been developed, they could be empirically studied.
As a national effort, 51 states and territories agreed to work together to design and adopt a set of common core standards. According to the National Governors Association and the Council of Chief State School Officers, the initiative was a way to articulate common learning progressions without the current variability that exists from state to state. According to their website, the common standards will help clarify the expectations and goals to involve parents and students. It should help parents understand exactly what students need to know and be able to do. More importantly, it will help students to become self-directed by knowing the expectations of successful learning ("Common Core State Standards Initiative," 2010).

In a study conducted by Lenz, Adams, Bulgren, Pouliot, and Laraux (2007), they found that use of learning progressions in the form of curriculum maps significantly increased learning, as shown through summative assessments, of students with learning disabilities. In the study 30 students, who qualified as Learning Disabled, enrolled in general education classes from two different high schools were chosen as participants. These students were administered the curriculum map intervention. The intervention consisted of a visual map depicting the structure of the content of the lesson; during the lesson, the instructor displayed and referred back to the map; finally, the closure of the lesson consisted of a review using questions and full review of the critical points on the map. The intervention resulted in a mean score of 63% compared with the control group’s mean score of 34%--a difference of almost 30% in mean scores on a summative assessment of the unit taught (Lenz et al., 2007).
Donna Shanks (2002) conducted a study to determine the effects of using curriculum maps to communicate learning progressions with students had on learning in second through sixth grade students. The study was conducted over a 2-year period—year 1 without curriculum maps and year 2 with them—using grades two through six in a rural school in Tennessee. The sample size was approximately 100 per grade level. The study found statistical significance in the mean Terra Nova scale scores for each grade-level student’s pre- and post-intervention results. The scores of the previous year were compared with those of the following year—when curriculum mapping was used. The use of curriculum maps as a means to communicate learning progressions to second through sixth grade students had a positive impact on summative learning as measured by Terra Nova scale scores (Shanks, 2002).

*Learning Targets*

Learning targets are descriptions of skills, concepts, or knowledge that students are expected to learn. According to McManus (2008), learning targets must be identified and communicated to students. The learning target should be aligned with a skill, concept, or body of knowledge essential to reaching the next level of understanding. The criteria for which learning will be assessed must also be communicated to the students for them to know if whether they are progressing toward the successful mastery of the communicated target of learning (S. McManus, 2008).

Jansen, Bartell, and Berk (2009) reported in the *Elementary School Journal* that learning targets are the major component in building a knowledge base for teacher education. The article stated that the two key features to support knowledge-building are
(1) targeted learning goals—specific enough to suggest interventions to learners to achieve and indicate the evidence of success, and (2) learning targets should be shared and communicated to the students and understood by all participants.

A research project conducted by Campos and O’Hern (2007) examined the use of learning targets and its effect on student empowerment. The project targeted one classroom of first and one classroom of fifth grade students. The researchers conducted an analysis of probable cause that revealed that students were unclear about what they needed to know, students were not taking responsibility for their learning, students were not receiving adequate descriptive feedback on their work, and they were not encouraged to reflect on learning. The intervention included introducing key concepts at the beginning of the lesson, making reference to the learning targets throughout the lesson, and using class and individual charts and graphs to track progress toward mastery of identified key concepts and targets of learning. The teachers would use feedback to provide individualized focused instruction to struggling students in small group settings. By collecting data through student survey and classroom discussions, the researchers reported an increase in student awareness of learning targets and goals. The examination of student portfolios supported that students as young as first graders were able to indicate skills that they had mastered and skills that they needed to continue to learn (Campos & O’Hern, 2007). The results of this study support the necessity of engaging students in learning through the use of identifying clear targets of learning and implementing self-assessment strategies to monitor individual progression toward mastery.
A research project conducted by Althoff, Linde, Mason, Nagel, and O'Reilly. (2007) supported that the use of daily learning targets had a positive effect on high school students in a Midwest metropolitan area. The study examined the effect that the use of posting daily targets of learning standards in student-friendly language would have on 150 high school students with five different teachers of English, social science, and world language. Student and teacher surveys were used to collect data. The intervention comprised posting daily targets and implementing a bi-monthly student comprehension checklist to self-assess mastery of posted targets. The post intervention collection of data occurred after nine weeks of intervention. The data showed an increase in achievement and awareness of learning targets. As reported by the student comprehension checklist, students reported a 28% increase in awareness of learning targets and level of comfort in facing assessments. Class achievement records displayed an increase of 13% in class averages, and the survey results showed an increase in 9% of students reporting a conscious knowledge of daily expectations (Althoff et al., 2007).

Simon and Taylor (2009) hypothesized that student and instructor misconceptions and the related negative learning consequences with such differences in perception of explanations and instruction could be addressed by the use of explicit learning targets for lectures in a post-secondary setting. Simon and Taylor studied the effect that the use of learning targets had on individual students in three different college courses. The studies sought evidence to ascertain whether students found the targets as useful and how the student used the learning targets. The results of the study indicated that nearly all students indicated that the learning targets were valuable. Eighty-five percent of all
comments were positive. The most frequent comment on the use of learning targets was that they helped students know what they needed to know to be successful for the course. Students used descriptors such as: “focus,” “guided me,” “kept me on track,” and “outlined the course,” to describe the use of learning targets. Students also reported that learning targets helped them “…get more out of the lecture.” The results of the study suggested that learning targets gave students a clear direction and expectation of learning, helped them self-monitor their progress, and caused the classroom instruction to become more efficient (Simon & Taylor, 2009).

Crooks (1988), Stiggins (2008), Wiliam (2004) among others stated that if students clearly understand the goals and measures of their learning then the student will be able to make choices concerning to the tasks that will display their learning. By allowing students choice it stimulates the intrinsic motivation and provides relevant challenges specific to individual needs. Overall, these authors make a case for the emphasis of evaluations to be on the skills and knowledge perceived to be most important.

**Descriptive Feedback**

According to the Council of Chief State School Officers, “Students should be provided with evidence-based feedback that is linked to the intended instructional outcomes and criteria for success” (S. McManus, 2008, p. 4). Descriptive feedback should be about specific qualities of student learning to continue the learning process through discussion or suggestions for improvement. It should be based on learning goals, progressions, or learning targets and not comparisons with other students and should help
the students identify where they are on the learning continuum (S. McManus, 2008). Anne Davies (2008) asserted that evaluative feedback has a negative impact on student learning by decreasing student motivation. She agreed that the most effective feedback is descriptive and specific and is designed to improve learning.

Through his research, Crooks (1988) stated that too much emphasis has been placed on the role of grading as a method of evaluation and too little on its role of assisting students to learn. He continued by claiming that there is little need or justification for the use of normative grading especially with its production of undesirable results as supported by Marzano (2006), Reeves (2007), Guskey (2005) and O’Connor (2002) among others. Crooks claimed that his evaluation of the research suggested that evaluation activity would be more beneficial if its sole purpose was to provide meaningful feedback to students.

Upon evaluating the research, Crooks (1988) suggested several methods that research had shown to make classroom evaluation more effective. First, feedback is most effective when it provides information and focus for the student on his or her progress toward mastery of an educational task and thus enhances self-efficacy, encourages effort, and limits attention to comparative practices (Black, McCormick, James, & Pedder, 2006; Black & Wiliam, 1998a, 1998b; Crooks, 1988; Stiggins & Chappuis, 2005). Second, feedback should occur while it is still relevant and soon after or during the completion of the task. The student should then be able to show mastery or improved learning after the feedback is given (Black & Wiliam, 1998b; Crooks, 1988; Leahy et al., 2005). Third, the feedback should be specific to the student and the student’s need in a
method easily understood by the student. Students should also be given clear learning
standards that are high but attainable and receive appropriate feedback directed at
attaining these standards (Althoff et al., 2007; Chappuis & Stiggins, 2002; Crooks, 1988;
Stiggins, 2007).

Students should be given frequent opportunities to practice and use skills learned
and receive appropriate feedback on their performance and progress. This evaluation
does not require formal evaluations or formal feedback.

According to a study conducted by Nelson and Schunn (2009), not only is
feedback important but also the type of feedback can have varying levels of impact on
student performance. This study examined five different types of feedback features used
to respond to writing—summarization, specificity, explanation, scope, and affective
language. The study used 1,073 feedback segments given over the course of a college-
level class with participants ranging in age from 18 to 21. The study coded the feedback
given on student-generated writing. The study then collected data about the student’s
response to that feedback and the implementation of that feedback on the final product.
Nelson and Schunn summated their findings of the type of feedback that students rated as
most helpful and had the most implementation in the final draft: feedback should be
given with a summary of students’ performance, should be specific, and should not only
indicate the problem but also should generate possible solutions.

For students to continue to self-assess their location on the learning continuum, it
is equally vital for the teacher to develop methods of feedback throughout the course of a
lesson. Students need to be able to express understanding and receive specific feedback
from variety of sources—self, peer, and teacher (Black & Wiliam, 1998b; Leahy et al., 2005). According to the examination of studies, Black and Wiliam (1998b) found that the majority of teacher feedback from the selected studies was described as inhibiting future learning. The teacher was often seen as redirecting the student to provide the appropriate and expected response, and such a response rarely was reflective of critical thought. Such question-and-answer sessions reported shared a common problem of teacher wait time. The question was posed and the teacher then selected a student to answer within seconds of asking the question. The examined research suggested that such questions were of lower-level thought—such as factual information—and did not require time to think through an answer or were followed by the correct answer provided by the teacher, which prevented the engagement of students through learned behavior: The student knew the answer was coming and therefore waited for the teacher’s expected answer instead of developing his or her own. This teacher behavior was shown to inhibit the thought processes of the students and conditioned him or her to seek out the answer that the teacher wants instead of developing the student’s own understanding through trial and error and genuine learning. Black and Wiliam summated that effective dialogue used as feedback should be focused to promote exploration of understanding and should allow all students the opportunity to express ideas.

As presented in the research findings from their study Black et al. (2004) showed gains of 0.3 standard deviations in effect size through addressing this issue of questioning. In the intervention teachers provided additional wait time by several seconds to give students time to think. Teachers also used random selection techniques
to encourage all students to formulate an answer. All answers were encouraged to promote “thoughtful improvement.” The purpose of asking questions was to raise issues in which the teacher needed information or to promote student thinking. The results of this intervention showed an increase in more active participants and a refocus for students on learning and less on spotting the correct answer. The teacher also shifted his or her perspective from presenter of content to guide for exploration and development of ideas and critical thought (Black et al., 2004; Leahy et al., 2005).

In the study by Black et al. (2004) summative tests were also used to provide formative feedback. These tests, combined with the three other formative assessment techniques, contributed to the increase in 0.3 standard deviations of the effect size. The results of the study suggested that to improve student achievement, students need to be engaged in reflection, and to review the work they have done, and to plan revisions to continue the learning process beyond the test. Students should practice creating questions and providing examples of mastery answers. Students should be provided criteria to help them understand how their work should be improved and then given the opportunity to continue the learning and display their mastery at another point for example, reworking exam questions and then taking an alternate assessment (Black et al., 2004).

The role of feedback is to increase student achievement and student desire to learn. Tan, Biswas, and Schwartz (2006) compared the effects on student achievement of two different forms of feedback: corrective-cognitive and guided-metacognitive. This study consisted of fifth grade students working with a computer-based feedback system.
One group received corrective feedback, while the other received guided feedback. The results showed that immediate corrective feedback had an impact on the student’s achieving immediate goals and that guided feedback better prepared the student for future learning goals. The students who received corrective feedback performed better on the immediate assessment of the taught concept, but were unable to carry over that knowledge to the next lesson. The students who received guided feedback did not perform as well on the immediate assessment of the lesson, but were able to apply self-monitoring and assessing and thus performed better on subsequent assessments of new lessons. The results of this study suggest that guided metacognitive feedback helps students learn skills and application of knowledge and essentially encourages learning, while corrective feedback impacts the students’ performance on the lower-level knowledge base. Both have a positive impact on student achievement but address different levels of learning and application of that learning (Tan, Biswas & Schwartz, 2006).

Black and Wiliam (1998a, 1998b), William et al. (2004), Stiggins (1999), and Stiggins and Chappuis (2008) have shown research evidence to support the positive effects that accurate, specific, and timely feedback about student learning had on student achievement and motivation. Assignment grades, assessment grades, and course progress reports all were opportunities to provide formative feedback to students; however, as previously examined, stakeholders perceived grades differently. Teachers incorporated many factors beyond academics into grades such as timely, complete, behavior, and neatness (Austin et al., 1992; Carlton, 1992; Davinroy et al., 1994; Guskey, 2002;
VanderHeide, 1994). The examined research studies by Black and Wiliam (1998b) and others concluded that assignment grades, assessments grades, and progress reports were not being used in a formative method in the majority of the selected environments. Conversely, the examined research has shown that student learning improved when students received specific feedback on strengths and weaknesses with methods of improvement. Studies that showed the greatest impact did not associate a grade to such feedback. Examined research studies have established that the use of numerical scores or grades had a negative effect and encouraged students to ignore comments as feedback when given (Black et al., 2004).

Black et al. (2004) concluded that the use of grades on homework did not provide students with feedback directed at improving and therefore did not serve the purpose of enhancing learning. In their study, teachers improved feedback through homework in several ways. Questioning and tasks should encourage students to develop and show understanding. Comments should identify strengths and areas of weaknesses and give guidance on how to improve. Students should be given opportunities to respond to comments and continue the learning process. Effective feedback will promote student thinking to take place. The evidence presented in this study demonstrated that implementation of such feedback methods changed the perspective of the student to view homework and the assessment of such work as a step in the learning process and not as the summative or end of learning (Black et al., 2004).

According to research conducted by Lipnevich (2007) on college students, the assessment feedback that promoted the most increase in student achievement was
This study examined the impact of differing forms of feedback on student affect and performance in an attempt to determine the optimal forms of assessment feedback. The study consisted of a randomized experiment design on college students writing an essay. Students received feedback from the course instructor, no feedback, or a computer-based program. Each type of feedback was also crossed with grades and no grades; and praise and no praise. The results reported that the most effective feedback was specific and descriptive, with no grades. Student performance was depressed when grades were given along with the descriptive feedback (Lipnevich, 2007).

Self- and Peer-Assessment

According to the CCSSO (S. McManus, 2008), “Both self- and peer-assessments are important for providing students an opportunity to think meta-cognitively about their learning” (p. 5). The CCSSO defined self-assessment as a process in which, “…students reflect on and monitor their learning using clearly explicated criteria for success” and peer-assessment as a process in which “…students analyze each others’ work using guidelines or rubrics and provide descriptive feedback that supports continued improvement” (p. 5). Well-designed learning progressions and clear learning targets are necessary to provide specific feedback and for students to engage in self-assessing and peer assessing.

According to the analysis of research by Black and Wiliam (1998b) as supported in position by Rick Stiggins (1999, 2002, 2006, 2007), many successful formative assessment innovations have involved the use of self-assessment and peer assessment, the
ability for students to accurately determine where they or their peers are on the learning continuum toward mastery of the skill or concept, and the research has shown success in impacting learning from age 5 upward. According to Black and Wiliam (1998b), students lacked clear and understood targets of learning. Students were only able to self-assess or peer-assess if they understood the learning targets and what success or mastery of those targets looked like (Althoff et al., 2007; Black & Wiliam, 1998b; Stiggins, 1999; Stiggins & Chappuis, 2008). Black and Wiliam (1998b) stated that the ability to self-assess is “…an essential component of formative assessment” (p. 143). The student must be able to recognize where they are supposed to be headed, where they currently are, and how to narrow the gap. Students must be trained to self-assess to understand the main purpose of their learning and to refocus on learning and improving learning (Black & Wiliam, 1998b; Chappuis & Stiggins, 2002).

The results from the study conducted by Black et al. (2004) reinforced the need for students to be able to understand the target of learning and to self-assess their progress or mastery of that goal or target. Likewise, peer assessment was shown to be valuable in the gains of effect size through this study. The peer assessment was reported, through this study, to be particularly helpful: The language of the feedback was familiar to students because it appeared in the natural language of their peers. It was also reported to have a positive impact on the students providing the assessing because students learned by taking on the role of teacher and thus applied that skill to their own performance through self-assessing. Another reported positive result of peer assessment was placing the work in the hands of the students which allowed the teacher to observe and provide
specific feedback to struggling students and also to reflect on the lesson and adapt based on observation and feedback provided from the class to create helpful interventions for instruction to address such feedback. The dramatic results from this study came from creating strategies based on the following criteria for effective assessments: The criteria for assessing any task or assignment should be clear to the student. The student should know what mastery or success looks like. Students should be taught how to peer-assess. This type of assessment not only provided additional feedback to the peer being assessed but also, maybe more important, provided the necessary practice and skills for students to experience self-assessment. Students should be encouraged to focus on the clear learning targets and should be taught to self-assess their progress and mastery of that target, the result of which provides a base for students to move toward self-directed and independent learning (Black et al., 2004).

According to a study by Butler (1990) of 80 5-, 7-, and 10-year-old students, students who were involved in mastery learning using a standards-based approach practiced better reliability and were more realistic about their performance when self-assessing. Students involved in a competitive normative approached were more likely to overestimate their ability at a younger age and the older the children, the less likely they would overestimate their performance. The students in the mastery condition showed no age difference in their ability to self-assess. Butler’s results implied that the young children in this study were capable and interested in self-assessment and did so with accuracy. The research implied that schools should foster mastery of learning goals and self-evaluations, instead of norm-based competition.
Sadler and Good (2006) conducted a study of the effect of self- and peer grading on four middle school science classes. They compared teacher-assigned grades with those awarded by students either to themselves or to other students. Students used a scoring rubric and were trained in assessing. The study reported a high correlation between the students’ self-assessed scores and their teachers’ scores. In addition to this correlation, students who practiced self-assessment showed an increase in student learning far greater than the control group when a second test was administered after the initial assessment. Students at all levels gained from performing self-assessment, with the largest gains from those at the lower and middle levels of achievement (Sadler & Good, 2006).

Ellen Callahan (2007) found that peer assessment had a positive effect on the attitudes of high school students toward science courses. The results supported that peer assessment was reported by the participating students as providing meaningful feedback and promoting collaboration and improved understanding. The participating students responded that these identified changes directly led to an improved attitude toward science instruction. Callahan used the action research form of inquiry to articulate the attitude of high school students and the subsequent change of attitude. The study was performed over six weeks using 115 11th and 12th grade high school science students in one high school enrolled in a common course, Human Anatomy. The researcher implemented peer assessment techniques, defined as “…evaluation procedures through which students provide feedback for improvement as they assess their peers’ performance and work” (p. 2), into the course and collected data through observation, questionnaires,
and interviews. The peer assessment checklists created by the researcher were used as a guide to aid the use and focus of peer assessment. The worksheet helped the peer assessor to present objective data, such as correct and incorrect, and also guiding questions to promote discussion and specific feedback. Callahan summarized her findings by reporting that the participants agreed that the use of peer assessment helped create an atmosphere of collaboration by allowing the recipients to remain free of summative assessments while receiving specific feedback. The results of the questionnaire data presented that 94% of participants reported feeling comfortable receiving help from their peers during peer assessment activities. Ninety-three percent responded that they felt the activities helped them learn and understand, and 84% reported that the peer assessment provided them with an identification of areas in which they needed to improve. Eighty-six percent of participants also reported an increase in interest level for the content when peer assessment was utilized, and 95% of the participants reported an increase in responsibility for their peers’ learning. The student interviews supported the above reports. All interviews made reference to giving immediate feedback as a benefit of the peer assessment activities (Callahan, 2007).

Laurynn Evans’ (2009) research using a quasi-experimental design in 9th grade English classrooms reported a statistically significant impact on achievement of students who used reflective assessments. In this study, reflective assessment was defined as a meta-cognitive function designed to help students assess where they are on the learning continuum through the use of self-monitoring and evaluation and therefore is considered a self-assessment activity. Evans conducted research using nine 9th grade English
classrooms in a large suburban high school. The classrooms were randomly assigned to one of three treatment conditions: control group with no changes, co-planned curriculum, or co-planned curriculum with reflective assessments. The co-planned classes were identical except for the use of reflective assessment. The experimental classes used reflective assessment activities in which students would either write personal learning statements or verbalize their thinking of what they have learned in the course of the lesson. The teacher would collect this information and then provide feedback through either written or verbal communication (Evans, 2009).

Data was collected and analyzed using posttest results from the summative assessment for the unit of study. The posttest was a collaboratively designed instrument by the researcher and teacher, and comprised multiple-choice and short-answer questions. Outside experts also examined the instrument for validity, and it tested positively for reliability. The results of the statistical analysis supported that students in this study who were involved in reflective assessments had a higher academic gain than those in the control or comparison group (Evans, 2009).

Terri Faitel (2007) also reported that using self-assessment practices can increase teacher understanding of student thought during high stakes standardized testing. Faitel presented a description of the impact that student self-assessment followed by teacher examination had on an eighth grade teacher preparing students for the Michigan Educational Assessment Program (MEAP). Participating eighth grade students were given practice MEAP tests, which were graded and returned to the students. The students then were asked to perform a self-assessment strategy in which they were to examine
each problem. They were to describe what test-taking and/or problem-solving strategy they used for the correctly answered questions and categorize their mistakes into two categories: “silly mistakes” and “genuine mistakes.” The teacher emphasized eliminating “silly mistakes” and had students focus future learning on the areas of “genuine mistakes.” Faitel reported that through the course of three years, a discrepancy appeared between the multiple-choice answers chosen by students and their written demonstration of knowledge exhibited in the self-assessment practice. In other words, students often chose correct answers for incorrect reasons. In contrast, many students with incorrect selections often described correct reasoning and understanding of the tested concepts. Faitel reported a decrease in this disconnect with the increase use of the test, return, self-assess, retest cycle.

**Collaboration**

The final attribute identified by the CCSSO, Collaboration, is the result of implementing formative assessment attributes into the classroom instruction and assessment of students. The CCSO put it this way: “A classroom culture in which teachers and students are partners in learning should be established” (S. McManus, 2008, p. 5). Popham (2008) described this as level 3 of formative assessment, whereas level 1 is the teacher’s use of formative assessment information and level 2 is the student’s use of formative assessment and his or her ability to make decisions about learning. It is the combination of these two levels that, according to Popham, create a classroom climate shift to collaboration and shared interest in learning.
In similar studies, students’ perceptions of assessment and grading impact the milieu of the classroom. S. M. McManus (2008) presented that when participating high school students had a clear understanding of their learning targets and what successful mastery of those targets looked like, both students and teachers reported that their classrooms shifted to a more open and collaborative setting. McManus also presented data to support that students became more involved in their own learning and took more ownership of their learning. This resulted in justifying answers and analyzing solutions. Likewise, teachers reported that students developed more positive attitudes and displayed this attitude through respectful behavior, more homework completion, and more positive comments about assignments. McManus presented that students’ self-efficacy increased, as shown through the students’ increased engagement in the learning process.

Callahan’s (2007) study on peer assessment of 11th and 12th grade science students presented responses that students who were involved in peer assessment and received specific teacher feedback reported a positive, collaborative classroom atmosphere. Students reported that the use of peer assessment helped them feel responsible for their own learning and also their peers’ learning.

Research completed by Croley (2003) reported that factors such as classroom atmosphere and interaction and/or perception of the teacher contributed to and/or reduced anxiety in middle school female students during math assessments. The study consisted of 25 female seventh grade students who performed poorly on math assessments because of anxiety as determined by the MARS-A survey. The study consisted of a series of semi-structured interviews with the participants. Of the reported factors, classroom
atmosphere ranked as one of the highest factors in both contributing to and reducing anxiety. The teacher and his or her interaction with students was another highly rated factor. Participants generally responded that a teacher who cared about student learning and helped students learn and did not just assess or judge performance had a large impact on relieving anxiety. Likewise, classmates who helped participants to understand concepts and answer questions were also a factor in minimizing anxiety. Croley also reported that students believed that a more standards-based approach was helpful in minimizing anxiety. Students reported wanting to have less topics and skills introduced and instead have more practice and better performance with key concepts before introducing additional topics and skills.

Sara Ahern (2009) reported the results of her study on transparency in assessment through web-based communications: Providing parents and students greater access to assessment data and feedback increased students’ desire to learn. Ahern studied the use of a web-based grade book, PowerSchool, and the subsequent impact the use of this tool had on the teachers’ perception of parent, student, and teacher collaboration. The study reported that teachers perceived that their students were more interested in their learning and increased their desire to continue to learn. Although the study reported that teachers did not witness an increase in parent inquiry, they did witness an increase in student inquiry and use of the grade book tool. Teachers reported that students exhibited greater motivation, effort, and ownership of learning and grades. The use of the online grade book increased the collaborative nature of the classroom setting, and this study reported a perceived increase in student desire, ownership, and motivation to learn (Ahern, 2009).
**Summation**

The historical research shows that assessments have been widely used to measure students against each other or a set of standards. The use of assessments to measure schools and district performance has added even more, high-stakes pressure on performing well on tests. The result is a system that places students and schools in rank order. Even as our focus has shifted from placing students in order of achievement to leaving no child behind, we have not yet widely addressed the change in assessments (Gardner, 2004; Stiggins, 2007). Our education system still uses assessments to measure performance after learning has occurred rather than as a tool for student learning and continual school improvement. The widely accepted use of assessments is as a measurement of learning rather than as a tool to promote learning. Summative assessments—assessments after learning has occurred—far outnumber formative assessments—assessing where students are on the learning continuum to inform decision making while learning is occurring (Hargis, 2003).

According to the review of literature, formative assessment attributes can have great impact on student achievement. Formative assessment is not a “what,” but rather a “how.” It is the process used to collect information and to use that information to make decisions by the educators and the individual student (Black & Wiliam, 1998b; S. McManus, 2008). The CCSSO identifies the most powerful attributes used in this process as the use of a thoughtfully designed learning progression, communicating learning targets to students, providing descriptive feedback, encouraging self- and peer assessment, and harboring a collaborative climate (S. McManus, 2008).
To engage the student in the assessment process and thus encourage the student to focus on learning rather than point collecting, educators need to shift student culture by developing learning progressions that focus on learning targets that can be communicated to students, providing descriptive feedback to those students on their performance toward that target, and encouraging self-and peer-assessment (Leahy et al., 2005; S. McManus, 2008; S. M. McManus, 2008).

Educators need to emphasize and use methods that describe student performance and make suggestions on what that student needs to do to narrow the gap between where he or she is and the learning target. This is referred to as “descriptive feedback.” The research also suggests that the majority of feedback a student receives is evaluative or summative in nature (Leahy et al., 2005). In other words, the student receives feedback after learning has occurred. Most often this feedback appears as grades or scores. In this form, that feedback is not received as feedback to encourage learning, but rather as a judgment or measurement of learning (Carlton, 1992; Davinroy et al., 1994; Marzano, 2006). Current research implies that grades and marks are not effective means of descriptive feedback. Many non-academic factors are included and are interpreted by the stakeholders differently.

Although much research exists on how formative assessment can impact student academic performance as measured by standardized achievement tests, little research is available on how the use of formative assessment can impact student affect and create a climate of collaboration. To prepare students for a future of continual learning it is increasingly more important to study the impact that schools can have on addressing the
student perception of learning and assessing as measured through changes in student academic efficacy and eagerness to learn. If we are truly going to make an impact on student learning, it is important to study the impact on student academic efficacy and what encourages students to be eager to learn. The goal of education is to encourage students to continue to learn even after they have left the place called “school.” With the need to examine ways to help students achieve, it is increasingly more important to examine the impact that instructional and assessment methods have on students’ eagerness to learn and their academic efficacy. The review of literature shows a lack in such research that attempts to answer these questions.

**Research Questions**

The purpose of this study was to examine that impact systematic use of formative assessment attributes in classroom instruction had on student affect and to examine the association of formative assessment attributes in the classroom with student affect. The study sought to answer the following research questions:

1. Is there a statistically significant impact on students’ Academic Efficacy when the systematic use of formative assessment is used in daily classroom instruction?

2. Is there a statistically significant impact on students’ Eagerness to Learn when the systematic use of formative assessment is used in daily classroom instruction?

3. Is there a statistically significant association between students’ (control and experimental) perception of Clear Learning Target and Academic Efficacy?
4. Is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Academic Efficacy?

5. Is there a statistically significant association between students’ (control and experimental) perception of Clear Learning Target and Eagerness to Learn?

6. Is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Eagerness to Learn?
CHAPTER III

METHODS

Purpose of the Study

The main purpose of this study was to determine if the systematic use of formative assessment attributes in daily classroom instruction and assessment had an impact on students’ Academic Efficacy and Eagerness to Learn. The second purpose of the study was to determine if there was an association between student perception of Clear Learning Targets and Progress Monitoring Information with students’ Eagerness to Learn and Academic Efficacy in a middle class, medium sized, suburban middle school.

The main focus of the research was to determine if the systematic use of formative assessment attributes had an impact on two affective variables—Academic Efficacy and Eagerness to Learn. Both the control and experimental students’ affect was measured using a pre and post test administration of the Student Affect Inventory created by Popham and Stiggins (2008) in conjunction with the CCSSO (see Appendix A). The post administration results of the experimental and control group were compared to determine if the systematic use of formative assessment attributes had a statistically significant impact on student affect. In addition, the pre and post test data of the experimental group was compared to determine if the systematic use of formative assessment attributes had a statistically significant change in reported affect by students as measured by the inventory.
The secondary focus of this research was to examine if there was an association with the formative assessment attributes of Clear Learning Targets and Progress Monitoring Information with Academic Efficacy and Eagerness to Learn. The study sought to determine if there was an association between student responses about perception of formative assessment attributes—Clear Learning Targets and Progress Monitoring Information—with their responses about student affect—Academic Efficacy and Eagerness to Learn. Student responses from all student, both control and experimental, were examined to determine if there was an association between student perception of Clear Learning Targets and Academic Efficacy and Eagerness to Learn. In addition, student responses were also examined to determine any association between student perception of Progress Monitoring Information and Academic Efficacy and Eagerness to Learn. All control and experimental student participants were administered the Student Affect Inventory created by Rick Stiggins and James Popham (see Appendix A). The inventory collected student responses of their perception to four areas—receiving clear learning targets, receiving progress monitoring information, academic efficacy, and eagerness to learn. The responses were then examined to determine if there was an association with student responses between the four areas.

**Research Questions**

The purpose of this study was to examine the impact that systematic use of formative assessment attributes in classroom instruction had on student affect and to examine the association of formative assessment attributes in the classroom with student affect. The study sought to answer the following research questions:
1. Is there a statistically significant impact on students’ Academic Efficacy when the systematic use of formative assessment is used in daily classroom instruction?

2. Is there a statistically significant impact on students’ Eagerness to Learn when the systematic use of formative assessment is used in daily classroom instruction?

3. Is there a statistically significant association between students’ (control and experimental) perception of Clear Learning Target and Academic Efficacy?

4. Is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Academic Efficacy?

5. Is there a statistically significant association between students’ (control and experimental) perception of Clear Learning Target and Eagerness to Learn?

6. Is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Eagerness to Learn?

**Sample**

The sample used for this study consisted of students enrolled in sixth through eighth grade during the 2010-2011 academic year in a middle school 30 miles south of Chicago.
Demographics of the School District

The school consisted of 562 students in grades 6 to 8—55% white, 22% black, 11% Hispanic, less than 10% Asian, and 11% Multiracial. This middle school had 43% of its students on free and reduced lunch and considered low-income and had a limited English proficient rate of 8%, and a mobility rate of 28%, while maintaining an attendance rate of 94%. Class sizes on average for grades 6 to 8 ranged from 17.4 to 19.7. The school maintained a per pupil instructional expenditure of $5,602, which was under the state average by $200. This middle school had made AYP for the 2009 school year in all areas and subgroups. At the time of the study, 82% of all students had met or exceeded standards in reading and 82% had in math and had met the AYP expectations since 2006 in all areas except for the subgroup of students with disabilities—this school had made AYP for this subgroup in the 2009-2010 school year. This school received students from three different elementary schools all within the same school district. The school operated in a modified version of the middle school method. Grade level teams of teachers were assigned a group of students. Each team consisted of a content level teacher for math, science, social studies, and language arts. The middle school method builds time in the teaching schedule, called team prep, for the team to meet and discuss student concerns or instructional delivery on a regular basis. The modification to this method was that this team of teachers followed their group of students on to the next grade level each year until they exited middle school.
Participants

This study utilized a quasi-experimental design. The researcher was granted permission to access students and teachers in the study by allowing the groupings to remain intact. Students had been assigned classes and teacher teams through a computer application. Teachers were assigned to teams by the administration through a combination of need and group dynamics. In this design, the researcher utilized pre-existing student and teacher groups. This design allowed for random assignment of intact groups (Gay, 2003). In this setting, one team, teachers and their students, were selected to be the experimental group while the remaining teams were assigned as the control group. The control and experimental groups were equivalent groups with no known major differences.

The researcher maintained a letter of authorization from the superintendent of schools authorizing the research study. The original letter was submitted to IRB, and a copy of the letter was kept in a secure file cabinet in the researcher’s office.

Control and experimental teacher consent was collected through signed consent letters (see Appendix E) administered on the first day of summer professional development by the researcher. Signed consent letters were kept in a secure file cabinet in the researcher’s office. All teachers in the school were recruited for the study. A presentation about the study was made to the staff. Those teachers who chose to participate were assigned randomly as control or experimental.

Consent letters (see Appendix F) were sent home with both control and experimental students the first day of the fall semester. Letters were given by the
students’ homeroom teachers to their students during homeroom. The homeroom
teachers read the uniform description of the study to all students. Students had five days
to turn in parent consent letters. Students returned letters to the homeroom teacher. The
homeroom teacher collected letters in an envelope and returned the envelope to the
building principal daily. The researcher collected the envelopes at the end of each day
during the five days. All students that turned in parent consent letters were asked for
assent using the same procedure on the fifth day. The homeroom teacher read the
uniform instructions to the students. Students who agreed signed the assent letter (see
Appendix G) and returned it to the homeroom teacher. The homeroom teacher collected
letters in an envelope and returned it to the building principal. The researcher collected
the letters from the building principal at the end of the day; letters of consent and assent
were stored in a secure file cabinet in the researcher’s office.

Identification of participating school district, teachers, and students was kept
secure. While no harm would result to the participants if their participation were known,
all information was coded and did not contain personal identifiable information. All such
information was kept in a secure file cabinet in the researcher’s office for the duration of
the study and the corresponding time required until such a time that it may be destroyed.

*Experimental Group*

One team of teachers was chosen to act as the experimental group and received
additional professional development in the attributes of formative assessment and their
classroom use. The experimental group teachers consisted of four male teachers and two
female teachers. The teachers varied in years of experience of three to seven years of
educational teaching experience. The median years of experience was 4 years and the average was 4.5. The experimental group’s students ranged from grades 6 to 8. The total number of possible student participants in the experimental group was 86. The total number of student participants in the experimental group was 12. The study consisted of four sixth grade students, seven seventh grade students, and one eighth grade student. Of the 12 participants, eight were female and four were male.

Participating teachers in the experimental group received sessions of professional development. These teachers received professional development about the attributes of formative assessment. It was expected that these participating teachers would systematically include the attributes in daily classroom instruction. Teachers were to increase the use of articulating learning targets, providing feedback, promoting self- and peer-assessment, and promoting a collaborative learning environment. The Formative Assessment Questionnaire for Teachers (see Appendix B) was used to collect information about teacher value and frequency of use of formative assessment attributes at the beginning and end of the study.

There was no risk for students in the experimental group. Teachers continued to instruct using the school and district curriculum and the state learning standards. Students were given the same access to the skills and concepts as the control group. Students in the experimental group may have received more information about their progress. Students in the experimental group were taught by teachers who had additional professional development in areas of articulating learning targets, providing feedback, promoting self- and peer-assessment, and promoting a collaborative learning
environment. The result of this professional development would be that the teacher included instructional methods utilizing these formative assessment attributes in class. This methods included: an increase in identifying the learning target to students, increase in rubrics, decrease in summative assessment, increase in formative assessments, increase in written or verbal feedback, increase in opportunities for students to show understanding, increase in encouragement or activities for self- and peer-assessment and increase in including the students in collecting and evaluating data about their own performance or understanding—reflection, meta-cognition, self-monitoring, or self-tracking performance through charts and graphs. These students were asked to complete the Student Affect Inventory at the beginning and end of the study.

Control Group

The control group consisted of 18 female teachers and two male teachers. The median years of experience were seven years, and the average years of experience for the control group were 7.2. The total number of student participants in the control group was 149. The study consisted of 59 sixth grade students, 39 seventh grade students, and 51 eighth grade students.

Participating teachers that were assigned as the control group continued to instruct their students as they normally did. These teachers continued to follow the school and district’s curriculum guides and state standards for learning. They continued to assess students as they have done in the past. These teachers did not receive the additional professional development on formative assessment attributes. They continued to teach as they have been instructed to teach by the school and district. At the beginning and end of
the study period (18 weeks) these teachers were asked to complete the Formative Assessment Questionnaire (see Appendix B).

Participating students in the control groups participated in class as normal. No changes other than the normal instructional changes that occur in a regular classroom happened. These students were asked to complete the Student Affect Inventory (see Appendix A) at the beginning and end of the study.

**Instrumentation**

*Student Affect Inventory*

In order to examine the impact the systematic use of formative assessment attributes had on student Eagerness to Learn and Academic Efficacy and the association of student perception of Clear Learning Targets and Progress Monitoring Information with Academic Efficacy and Eagerness to Learn, participating control and experimental group students were administered the Student Affect Inventory (see Appendix A) created by Stiggins and Popham (2008) in conjunction with the Council of Chief State School Officers at the beginning and end of the study. This study utilized the Student Affect Inventory instrument to collect student responses to perception in four areas—Clear Learning Targets, Receiving Progress Monitoring Information, Academic Efficacy, and Eagerness to Learn. The inventory was administered to participating experimental and control group students in grades sixth through eighth. The inventory was distributed during homeroom at the same time for each grade level. The inventory consisted of eight questions, two for each area, and took approximately 10 minutes or less to complete. The
students remained anonymous; only demographic information of gender and grade level was collected. The data was coded and entered into SPSS.

The researcher received permission to use the student inventory from Dr. Stiggins (see Appendix C). The Student Affect Inventory was created as a tool to collect information about student affect—academic efficacy and eagerness to learn—in order to make decisions about teaching. Richard Stiggins was the founder and executive director of ETS Assessment Training Institute and author of many books and articles on student assessment including *Classroom Assessment for Student Learning*. Dr. Stiggins served on the research and development staffs of ACT and the Northwest Regional Educational Laboratory as a psychometrician. James Popham was the Emeritus Professor in the UCLA Graduate School of Education and Informational Studies. Dr. Popham established the IOX Assessment Associates as a research and development group creating statewide achievement tests. He was the former president of the American Educational Research Association. Dr. Popham has authored numerous research articles and books on student assessment including *Transformative Assessment*. The student inventory’s construct validity was verified by a panel of curriculum and education specialists. The panel consisted of seven k-12 curriculum specialists with a median 23 years of educational experience and an average of 20.5 years of experience. The panel consisted of a school psychologist, three curriculum directors, one curriculum consultant, one assistant superintendent for curriculum, and one high school principal.

The participating teachers administered the Student Affect Inventory (see Appendix A) created by Stiggins and Popham (2008) in conjunction with the Council of
Chief State School Officers at the beginning of the school year in September to all experimental and control student participants after consent and assent was collected. The inventory took students approximately 10 minutes to complete. The teacher administered the inventory on a specific date during the second week of student attendance by following the script and uniform instructions; the inventory was administered during homeroom class time. Students not taking part in the study completed silent reading or independent study during that time. Students completed the inventory anonymously, and the students placed the completed inventories into an envelope. The teacher administering the inventory collected the sealed envelope and delivered it to the building principal. The envelope remained in a secure file cabinet, unopened. The researcher collected sealed inventories at the end of that day and examined at a site removed from the school. Completed inventories were stored in a secure file cabinet in the researcher’s office. The same inventory and procedure was repeated after approximately 18 weeks. The results of the inventory were analyzed for statistically significant change in student responses concerning Academic Efficacy and Eagerness to Learn. The results were also examined to determine if any association between Clear Learning Targets and/or Progress Monitoring Information existed with Academic Efficacy or Eagerness to Learn.

*Formative Assessment Questionnaire for Teachers*

The study also administered the Formative Assessment Questionnaire to teachers in the control and experimental group (see Appendix B). The questionnaire was used to determine if the teachers in the experimental group valued and systematically used
formative assessment attributes in daily instruction significantly more than the control group.

In order to examine the difference in value and systematic use of formative assessment attributes between the experimental and control group teachers, participating control and experimental teachers were given the Formative Assessment Questionnaire for Teachers (see Appendix B) to complete at the beginning and end of the 18 weeks. The questionnaire was created by the Qualification and Curriculum Authority (QCA). The QCA granted written permission to the researcher to use the questionnaire for research (see Appendix D). This organization was in charge of the development and delivery of education and instruction for the United Kingdom. The questionnaire contained items requiring a likert scale (an indication of agreement from “strongly agrees” to “strongly disagrees” and frequency of use from “most lessons” to “never”) and short response questions (Neesom, 2000). The instrument was used by the Learning How To Learn Project in London. The Assessment Reform Group commissioned numerous studies in which the questionnaire was used to collect data. The Learning How to Learn group established reliability through the use of three levels (classroom, school, and district). James, Black, McCormick, Pedder, and Wiliam (2006) used multiple regression analysis to establish validity. Analysis of the qualitative data utilized constant comparative analysis to establish validity.

Data was collected at the beginning and end of the time frame from the participating control and experimental teachers rating their value and frequency of formative assessment in the classroom. The Formative Assessment Questionnaire for
Teachers (see Appendix B) created by the Qualification and Curriculum Authority (2000) was administered through an online service—Opinio. Opinio is a secure database with security and masked IP addresses from the researcher. The questionnaire took approximately 20 to 30 minutes to complete. The teachers were informed of the online questionnaire time window one week ahead of time. The participants were invited to complete the online questionnaire by an email invitation the day the window opened. The window was open for three days. The questionnaire was accessible only once. Participants were not able to complete multiple questionnaires. The researcher was the only person who had access to the results of the questionnaire online.

Intervention

The study was conducted over an 18 week period and consisted of an experimental and control group. The researcher utilized one middle school team as an experimental group and the remaining grade level teams as the control. The experimental team teachers received training on formative assessment through a series of ongoing professional development meetings and coaching by the researcher (see Appendix H). The training consisted of an overview of the philosophy and more specific training on the five attributes of formative assessment, as identified by Sarah McManus (2008) in conjunction with the Council of Chief State School Officers, and the first two levels as identified by Popham (2008).

Popham’s Levels of Formative Assessment

Level 1: Teachers’ Instructional Adjustments

Level 2: Students’ Learning Tactic Adjustments
Level 3: Classroom Climate Shift

Level 4: School-wide Implementation (Popham, 2008, p. 49)

Five Attributes of Formative Assessment

1. Learning Progressions
2. Learning Goals or Targets
3. Descriptive Feedback
4. Self- and Peer-Assessments
5. Collaboration

The first meeting occurred prior to the beginning of the school year. The experimental group teachers were presented with an overview of formative assessment and its attributes. Teachers were coached on how to incorporate these attributes into classroom instruction. Subsequent meetings were held by the team on a weekly basis and the team met with the researcher as a coach every 2 to 3 weeks. Each session was used to increase the level of understanding and implementation of formative assessment attributes. Since the coaching method was utilized, each session began with the participants describing their use of formative assessment, success, and difficulties or concerns. The training session were driven by the participants’ identified needs in order to improve their fidelity with formative assessment.

Successful systematic use of formative assessment in the classroom included an increase of identification of the learning target to the student, an increase in short cycle assessments that generated progress monitoring information as feedback in place of a grade or score, an increase in progress monitoring information provided to the student
about where that student is on the learning continuum such as rubrics, an increase in opportunities for students to show mastery of a skill or concept, an increase in the use of peer- and self-assessment. Teachers were to systematically increase the amount of instructional decisions they make based on assessment data. The classroom that fully utilized formative assessment involved the student as a user of assessment information. In other words, the student received progress monitoring information as feedback about performance with the intent to make decisions about learning and study behaviors to continue learning. Students were guided by short formative assessments on their path to mastery of a skill or concept. Grades, points, or scores were collected after the student was expected to have learned. If a score did not reflect that a student had learned at the appropriate level, then that student had opportunities to continue to learn and show learning again at another time.

Training Sessions Descriptions

Session 1 was a full day in-service over the attributes and related research concerning formative assessment. It described formative assessment as a process and used the description of Formative Assessment Attributes as defined by McManus (2008). This session also addressed classroom policies and procedures that encourage learning and assessment as a tool for learning.

Session 2 was a two hour session focused on the use of learning targets, articulating targets to students, and creating assessments aligned to targets. This session also focused on strategies to collect information to make decisions about instruction.
Session 3 was a two hour session to continue training how to identify learning targets to students and provide descriptive feedback through progress monitoring on students’ performance on the learning continuum towards those targets.

Session 4 was a two hour session focused on involving the students in assessment, peer assessment, and self assessment. This session also focused on strategies to teach students how to receive feedback and make decisions about learning and studying strategies.

Session 5 was a two hour session focused on a collaborative classroom climate. This session focused on ways to encourage students to be engaged in the decision making process for learning and instruction and to minimize an atmosphere of fear.

The control group did not receive this training and was not included in any of the experimental group’s team planning meetings.

**Statistical Procedures**

The Student Affect Inventory data was collected, coded, and analyzed for statistical significance and associations using SPSS 17.0. Student and teacher responses were kept by the researcher in a secure file in a location separate from the research site. No personal identification information was collected. Student and teacher responses were coded to identify participant as either in the control group or experimental group.

In order to determine the impact the systematic use of formative assessment attributes used in classroom instruction had on Eagerness to Learn and Academic Efficacy, all student data was examined for statistical significance using SPSS 17 statistical analysis software. In order to collect data about student Eagerness to Learn and
Academic Efficacy, experimental and control students were administered the Student Affect Inventory before the study (pre-test) and following the end of the study time frame (post test), approximately 18 weeks later. The inventory collected student responses on a likert scale in four areas—Academic Efficacy, Eagerness to Learn, Clear Learning Targets, and Progress Monitoring Information. The Student Affect Inventory was analyzed for statistically significant difference in student responses in each of the four areas between the two groups—control and experimental.

The student post test inventory was examined using inferential statistics to determine if the difference between control and experimental group was statistically significant. A two-tailed, non-parametric, Mann-Whitney U was utilized to compare median differences of the post test inventory results. The responses were converted to data by coding responses on the likert scale from 1-5. The level of significance, or P-value, was set at < .05.

In order to examine the impact the systematic use of formative assessment attributes had on the experimental group’s Academic Efficacy and Eagerness to Learn, the results of the experimental group’s pre and post administration of the Student Affect Inventory was analyzed using the Mann-Whitney U to determine if there was a statistically significant difference between the pre-test and post test inventory. The level of significance, or P-value, was set at < .05.

In order to determine if there was an association between Clear Learning Targets, Progress Monitoring Information, Academic Efficacy, and Eagerness to Learn, all participating control and experimental students completed the Student Affect Inventory.
The collected data was collected, coded, and entered into SPSS 17. The responses to the four different areas of the student affect inventory was examined within the student affect inventory using the Spearman rho correlation coefficient to determine any association for each of the four areas of responses. For example, was there a correlation between student Eagerness to Learn and Clear Learning Targets? The student inventory was examined for statistically significant correlation between response for Eagerness to Learn and Clear Learning Targets and Progress Monitoring Information. The student inventory was also examined for statistically significant correlation between responses for Academic Efficacy and Clear Learning Targets and Progress Monitoring Information.

Teacher response data for the Formative Assessment Questionnaire for Teachers was examined using descriptive statistics and tested for significant differences between the control and experimental group. The teacher questionnaire collected responses from experimental and control teachers about value and frequency of use of four areas of formative assessment attributes—involving pupils in their learning, modeling quality, giving feedback, and self-assessment at the beginning and end of the study. All participating teachers, control and experimental, completed the questionnaire. The data was coded as ordinal data by the responses 1-5.

In order to determine if the teachers in the experimental group valued or reported using formative attributes in daily instructional significantly more than the control group, the post test administration for the control and experimental group of the Formative Assessment Questionnaire for Teachers response data was analyzed using a two-tailed, non-parametric, Mann-Whitney U test to determine any difference in the reported value.
and use of the formative assessment attributes between the experimental and control group.

In order to examine the change in value or use of formative assessment attributes in daily instruction, the results of the experimental group’s pre and post administration of the Formative Assessment Questionnaire for Teachers was analyzed using the Mann-Whitney U to determine if there was a statistically significant difference between the pre and post test inventory. The level of significance, or P-value, was set at < .05.

**Hypotheses**

H1: There will be a statistically significant difference between experimental and control groups in student responses on the post administration of the Student Affect Inventory for Academic Efficacy.

H2: There will be a statistically significant difference between the experimental group’s pre and post administration results on the Student Affect Inventory for Academic Efficacy.

H3: There will be a statistically significant difference between experimental and control groups in student responses on the Student Affect Inventory for Eagerness to Learn.

H4: There will be a statistically significant difference between the experimental group’s pre and post administration results on the Student Affect Inventory for Eagerness to Learn.

H5: There will be a statistically significant association between student (control and experimental) perception of Clear Learning Targets and Academic Efficacy.
H6: There will be a statistically significant association between student (control and experimental) perception of Clear Learning Targets and Eagerness to Learn.

H7: There will be a statistically significant association between student (control and experimental) perception of Progress Monitoring Information and Academic Efficacy.

H8: There will be a statistically significant association between student (control and experimental) perception of Progress Monitoring Information and Eagerness to Learn.

**Ethical Considerations**

*Voluntary Participation*

This study utilized voluntary participation from both teachers and student populations. The selected district and school volunteered to be a research site due to its congruence with the research topic of formative assessment.

*Informed Consent*

Since children were participants in the study, both parental consent (see Appendix F) and child assent (see Appendix G) were obtained. Parents received a letter, seeking volunteers and consent, sent home with their child explaining the study and providing information about their rights as parents and their child’s rights to participate and withdraw at any time with no penalty imposed by the researcher or school district or school. Signed letters of parental consent were collected and stored in a secure file cabinet in the researcher’s office removed from the site. Students who had parental consent were asked to participate through a letter of assent containing the same pertinent
information in the parental consent letter, but written at an age appropriate level. The assent letters were distributed during the school day to students by their homeroom teacher. Signed letters of assent were collected by the school and stored in the researcher’s office in a locked file cabinet.

All teachers were asked to participate through a presentation by the researcher during the summer professional development session held by the school. All teachers were invited to participate in the study. The researcher collected letters of consent (see Appendix E) from all wishing to participate. Convenience grouping was used to determine the participants who received the formative assessment professional development.

No Harm to Participants

Students were not exposed to any harm beyond that of everyday life. Teachers were not exposed to any harm beyond that of everyday life. Students and teachers in the control group proceeded as normal with no changes in instruction or learning other than the normal fluctuations and changes associated with the regular instruction methods employed in the school. Teachers in the experimental group received additional professional development in formative assessment congruent with the direction of staff development for the school. Formative assessment was a recognized instructional method. Students in the experimental group were taught by teachers who had additional professional development in formative assessment. All participating control and experimental students were asked to complete the Student Affect Inventory twice during the study. All participating control and experimental teachers were asked to complete
the Formative Assessment Questionnaire for Teachers at the beginning and conclusion of the 18 week study.

Anonymity and Confidentiality

No personal identifying information was collected from students or teacher participants. Inventories and questionnaires were coded with gender and as experimental or control group. All collected data was kept in a secure file cabinet in the researcher’s office at a removed site. Teacher questionnaire data was collected through a Loyola approved web-based survey collection program that met the requirements for security and anonymity as defined by Loyola University Chicago.
CHAPTER IV
RESULTS OF THE STUDY

Purpose of the Study

The purpose of this study was to examine the impact that systematic use of formative assessment attributes in classroom instruction had on student affect and to examine the association of formative assessment attributes in the classroom with student affect at a selected middle school located south of Chicago. The study addressed the following questions:

Research Questions

1. Is there a statistically significant impact on students’ Academic Efficacy when the systematic use of formative assessment is used in daily classroom instruction?

2. Is there a statistically significant impact on students’ Eagerness to Learn when the systematic use of formative assessment is used in daily classroom instruction?

3. Is there a statistically significant association between students’ (control and experimental) perception of Clear Learning Target and Academic Efficacy?

4. Is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Academic Efficacy?
5. Is there a statistically significant association between students’ (control and experimental) perception of Clear Learning Target and Eagerness to Learn?

6. Is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Eagerness to Learn?

Research Question One Null Hypothesis

$H_0$ There will be no statistically significant difference between experimental and control groups in student responses on the post administration of the Student Affect Inventory for Academic Efficacy.

$H_0$ There will be no statistically significant difference between the experimental group’s pre and post administration results on the Student Affect Inventory for Academic Efficacy.

Research Question Two Null Hypothesis

$H_0$ There will be no statistically significant difference between experimental and control groups in student responses on the post administration of the Student Affect Inventory for Eagerness to Learn.

$H_0$ There will be no statistically significant difference between the experimental group’s pre and post administration results on the Student Affect Inventory for Eagerness to Learn.

Research Question Three Null Hypothesis

$H_0$ There will be no statistically significant association between the students’ perception of Clear Learning Targets and Academic Efficacy.
Research Question Four Null Hypothesis

H₀ There will be no statistically significant association between the students’ perception of Progress Monitoring Information and Academic Efficacy.

Research Question Five Null Hypothesis

H₀ There will be no statistically significant association between the students’ perception of Clear Learning Target and Eagerness to Learn.

Research Question Six Null Hypothesis

H₀ There will be no statistically significant association between the students’ perception of Progress Monitoring Information and Eagerness to Learn.

Results

Tests and Data Collection Methods for Research Questions One and Two

To examine the impact the systematic use of formative assessment attributes by classroom teachers had on student Academic Efficacy and Eagerness to Learn, the Student Affect Inventory was administered to the experimental student group and control student group before the treatment and after 18 weeks of receiving the treatment. The results were analyzed for any statistically significant changes of the experimental group’s results from pre to post administration of the survey. The post administration results of the experimental and control group were also examined to determine any statistically different results. The results of the pre and post were coded and entered into SPSS 17.0. The data was examined for a statistically significant difference by performing a Mann Whitney U on the Experimental and the Control group. The significance level was set at .05.
**Experimental and Control Group Differences**

To determine if the systematic use of formative assessment attributes had a significant impact on student affect, the study employed a quasi-experimental design. Utilizing convenience grouping, one middle school team (students from sixth to eighth grade) was selected as the experimental group while the remaining teams were assigned as the control group. The teachers of the experimental group received ongoing professional development on formative assessment attributes to be used in their classroom teaching.

The study administered the Formative Assessment Questionnaire for Teachers to experimental and control group teachers to determine if the teachers in the experimental group valued and used formative assessment attributes more frequently in daily instruction more than the control group.

Table 1 lists the results of the Mann-Whitney U nonparametric test of experimental and control group responses to the value of the formative assessment attributes of: Involving Pupils in their Learning, Modeling Quality, Giving Feedback, and Self Assessment. The results indicated: a significance level of .786 for Involving Pupils in their Learning, a significance level of .509 for Modeling Quality, a significance level of .947 for Giving Feedback, and a significance level of .738 for Self Assessment.

The data, as illustrated in Table 1, indicated that there was no statistically significant difference (p>.05) in value of any of the four formative assessment attributes between the experimental teacher group and the control teacher group post study administration results of the Formative Assessment Questionnaire for Teachers.
Table 1

*Experimental and Control Teacher Post Results for Value of Formative Assessment Attributes*

<table>
<thead>
<tr>
<th></th>
<th>Involving Pupils Value</th>
<th>Model Quality Value</th>
<th>Giving Feedback Value</th>
<th>Self Assessment Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>44.000</td>
<td>38.000</td>
<td>47.000</td>
<td>43.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>54.000</td>
<td>48.000</td>
<td>347.000</td>
<td>343.000</td>
</tr>
<tr>
<td>Z</td>
<td>-.271</td>
<td>-.661</td>
<td>-.066</td>
<td>-.335</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.786</td>
<td>.509</td>
<td>.947</td>
<td>.738</td>
</tr>
<tr>
<td>Exact Sig. [2*(1-tailed Sig.)]</td>
<td>.825a</td>
<td>.547a</td>
<td>.975a</td>
<td>.776a</td>
</tr>
</tbody>
</table>

(p>.05)

Table 2 lists the results of the Mann-Whitney U nonparametric test of experimental and control group responses to the frequency of use of the formative assessment attributes of: Involving Pupils in their Learning, Modeling Quality, Giving Feedback and Self Assessment. The results indicated: a significance level of .369 for Involving Pupils in their Learning, a significance level of .217 for Modeling Quality, a significance level of .918 for Giving Feedback, and a significance level of .286 for Self Assessment.

The data, as illustrated in Table 2, indicated that there was no statistically significant difference (p>.05) in the frequency of use of any of the four formative assessment attributes between experimental and control groups post study administration results of the Formative Assessment Questionnaire for Teachers.
Table 2

Experimental and Control Teacher Post Results for Frequency of Use of Formative Assessment Attributes

<table>
<thead>
<tr>
<th></th>
<th>Involving Pupil Frequency</th>
<th>Model Quality Frequency</th>
<th>Giving Feedback Frequency</th>
<th>Self Assessment Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>33.000</td>
<td>28.000</td>
<td>44.500</td>
<td>30.500</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>43.000</td>
<td>38.000</td>
<td>54.500</td>
<td>40.500</td>
</tr>
<tr>
<td>Z</td>
<td>-.898</td>
<td>-1.234</td>
<td>-.103</td>
<td>-1.067</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.369</td>
<td>.217</td>
<td>.918</td>
<td>.286</td>
</tr>
<tr>
<td>Exact Sig. [2*(1-tailed Sig.)]</td>
<td>.490^a</td>
<td>.243^a</td>
<td>.921^a</td>
<td>.303^a</td>
</tr>
</tbody>
</table>

(p>.05)

Table 3 lists the results of the Mann-Whitney U nonparametric test of experimental group pre and post study responses to the value of the formative assessment attributes of: Involving Pupils in their Learning, Modeling Quality, Giving Feedback, and Self Assessment. The results indicated: a significance level of .503 for Involving Pupils in their Learning, a significance level of .775 for Modeling Quality, a significance level of .849 for Giving Feedback, and a significance level of .702 for Self Assessment.

The data, as illustrated in Table 3, indicated that there was no statistically significant difference (p>.05) in the value of use of any of the four formative assessment attributes between the experimental teacher group’s pre and post study results of the Formative Assessment Questionnaire for Teachers.
Table 3

*Experimental Teacher Pre and Post Results for Value of Formative Assessment Attributes*

<table>
<thead>
<tr>
<th></th>
<th>Involving Pupils Value</th>
<th>Model Quality Value</th>
<th>Giving Feedback Value</th>
<th>Self Assessment Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>10.500</td>
<td>12.500</td>
<td>13.000</td>
<td>12.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>20.500</td>
<td>22.500</td>
<td>41.000</td>
<td>40.000</td>
</tr>
<tr>
<td>Z</td>
<td>-.669</td>
<td>-.285</td>
<td>-.190</td>
<td>-.383</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.503</td>
<td>.775</td>
<td>.849</td>
<td>.702</td>
</tr>
<tr>
<td>Exact Sig. [2*(1-tailed Sig.)]</td>
<td>.527*</td>
<td>.788*</td>
<td>.927*</td>
<td>.788*</td>
</tr>
</tbody>
</table>

(p>.05)

Table 4 lists the results of the Mann-Whitney U nonparametric test of experimental group pre and post study responses to the frequency of use of the formative assessment attributes of: Involving Pupils in their Learning, Modeling Quality, Giving Feedback, and Self Assessment. The results indicated: a significance level of .340 for Involving Pupils in their Learning, a significance level of .343 for Modeling Quality, a significance level of .775 for Giving Feedback, and a significance level of .214 for Self Assessment.

The data, as illustrated in Table 4, indicated that there was no statistically significant difference (p>.05) in the frequency of use of any of the four formative assessment attributes between the experimental teacher group’s pre and post study results of the Formative Assessment Questionnaire for Teachers.
Table 4

*Experimental Teacher Pre and Post Results for Frequency of Use of Formative Assessment Attributes*

<table>
<thead>
<tr>
<th></th>
<th>Involving Pupil Frequency</th>
<th>Model Quality Frequency</th>
<th>Giving Feedback Frequency</th>
<th>Self Assessment Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>9.000</td>
<td>9.000</td>
<td>12.500</td>
<td>7.500</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>19.000</td>
<td>19.000</td>
<td>40.500</td>
<td>17.500</td>
</tr>
<tr>
<td>Z</td>
<td>-.954</td>
<td>-.949</td>
<td>-.285</td>
<td>-1.243</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.340</td>
<td>.343</td>
<td>.775</td>
<td>.214</td>
</tr>
<tr>
<td>Exact Sig. [2*(1-tailed Sig.)]</td>
<td>.412^a</td>
<td>.412^a</td>
<td>.788^a</td>
<td>.230^a</td>
</tr>
</tbody>
</table>

(p>.05)

Table 5 lists the number of participants by grade and gender for the control group.

The total number of participants in the control group was 149. The study consisted of 59 sixth grade students, 39 seventh grade students, and 51 eighth grade students. Of the 149 participants, 83 were female, 58 were male, and 8 did not identify either sex.

Table 5

*Demographic Information of the Control Group*

<table>
<thead>
<tr>
<th>Count</th>
<th>Grade level</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Gender</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Female</td>
<td>35</td>
<td>20</td>
</tr>
<tr>
<td>Male</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>59</td>
<td>39</td>
</tr>
</tbody>
</table>
Table 6 lists the number of participants by grade and gender for the experimental group. The total number of participants in the experimental group was 12. The study consisted of four sixth grade students, seven seventh grade students, and one eighth grade student. Of the 12 participants, eight were female and four were male.

Table 6

_demographic information of the experimental group_

<table>
<thead>
<tr>
<th>Count</th>
<th>Grade level</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

_research question one_

In order to reject the null hypothesis that there will be no statistically significant difference between experimental and control groups in student responses on the post administration of the Student Affect Inventory for Academic Efficacy, post survey results were analyzed using the Mann-Whitney U nonparametric test.

Table 7 lists the results of the Mann-Whitney U nonparametric test. The results indicated a significance level of .136.

The data, as illustrated in Table 7, indicated that there was no statistically significant difference between the experimental and the control group on the post administration results for Academic Efficacy. The hypothesis that there will be a
statistically significant difference between experimental and control groups in student responses on the post administration of the Student Affect Inventory for Academic Efficacy was not supported by the data (p>.05). A significance value at this level failed to reject the null hypothesis.

Table 7

*Comparison of Experimental and Control Group Post Results for Academic Efficacy*

<table>
<thead>
<tr>
<th></th>
<th>Academic Efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>667.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>745.000</td>
</tr>
<tr>
<td>Z</td>
<td>-1.491</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.136</td>
</tr>
</tbody>
</table>

In order to reject the null hypothesis that there will be no statistically significant difference of the between the experimental group’s pre and post administration results on the Student Affect Inventory for Academic Efficacy, pre and post survey results from the experimental group were analyzed using the Mann-Whitney U nonparametric test.

Table 8 lists the results of the Mann-Whitney U nonparametric test. The results indicated a significance level of .924.

The data, as illustrated in Table 8, indicated that there was no statistically significant difference between the experimental group’s pre and post administration results for Academic Efficacy. The hypothesis that there will be a statistically significant difference between the experimental group’s pre and post administration results on the
Student Affect Inventory for Academic Efficacy was not supported by the data (p>.05).

A significance value at this level failed to reject the null hypothesis.

Table 8

*Comparison of Experimental Pre and Post Results for Academic Efficacy*

<table>
<thead>
<tr>
<th></th>
<th>Academic Efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>94.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>172.000</td>
</tr>
<tr>
<td>Z</td>
<td>-.096</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.924</td>
</tr>
<tr>
<td>Exact Sig. [2*(1-tailed Sig.)]</td>
<td>.945^a</td>
</tr>
</tbody>
</table>

*(p>.05)*

*Research Question Two*

In order to reject the null hypothesis that there will be no statistically significant difference between experimental and control groups in student responses on the Student Affect Inventory for Eagerness to Learn, post survey results were analyzed using the Mann-Whitney U nonparametric test.

Table 9 lists the results of the Mann-Whitney U nonparametric test. The results indicated a significance level of .815.

The data, as illustrated in Table 9, indicated that there was no statistically significant difference between the experimental and the control group on the post administration results for Eagerness to Learn. The hypothesis that there will be a statistically significant difference between experimental and control groups in student
responses on the Student Affect Inventory for Eagerness to Learn was not supported by the data (p>.05). A significance value at this level failed to reject the null hypothesis.

Table 9

*Comparison of Experimental and Control Post Results for Eagerness to Learn*

<table>
<thead>
<tr>
<th></th>
<th>Eagerness to Learn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>858.000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>936.000</td>
</tr>
<tr>
<td>Z</td>
<td>-.234</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.815</td>
</tr>
</tbody>
</table>

(p>.05)

In order to reject the null hypothesis that there will be no statistically significant difference between the experimental group’s pre and post administration results on the Student Affect Inventory for Eagerness to Learn, pre and post survey results from the experimental group were analyzed using the Mann-Whitney U nonparametric test. Table 10 lists the results of the Mann-Whitney U nonparametric test. The results indicated a significance level of .312.

The data, as illustrated in Table 10, indicated that there was no statistically significant difference between the experimental group’s pre and post administration results for Eagerness to Learn. The hypothesis that there will be a statistically significant difference between the experimental group’s pre and post administration results on the Student Affect Inventory for Eagerness to Learn was not supported by the data (p>.05). A significance value at this level failed to reject the null hypothesis.
Table 10

Comparison of Experimental Pre and Post Results for Eagerness to Learn

<table>
<thead>
<tr>
<th></th>
<th>Eagerness To Learn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>74.500</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>210.500</td>
</tr>
<tr>
<td>Z</td>
<td>-1.012</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>.312</td>
</tr>
<tr>
<td>Exact Sig. [2*(1-tailed Sig.)]</td>
<td>.324</td>
</tr>
</tbody>
</table>

(p>.05)

Type II Error

In order to determine the impact the increased use of formative assessment attributes had on student affect, experimental and control group student pre and post survey results were analyzed. The control group did not have any statistically significant changes in responses in any of the four categories—Eagerness to Learn, Academic Efficacy, Clear Learning Targets, and Progress Monitoring Information. While the experimental group did have a change from pre to post survey, the results were not statistically significant in the same four categories. The collected data was unable to reject the null hypothesis.

The experimental group did show a change in the category of Eagerness to Learn. The mean change was 13.16 to 16.29. The researcher attributed the failure to reject the null hypothesis to a Type II error. The sample size for the experimental group was too
small \( (n=16) \) and suffered atrophy by the end of the study \( (n=12) \) to register a statistically significant change.

Tests and Data Collection Methods for Research Questions Three through Six

In order to examine for associations between formative assessment attributes and student affect, student inventory data was collected using the Student Affect Inventory created by James Popham and Rick Stiggins. The inventory collected both control and experiment group student responses about their perceptions in four areas—Clear Learning Targets, Progress Monitoring Information, Academic Efficacy, and Eagerness to Learn. The study collected 337 total student responses during the course of the study. The results of the inventories were coded and entered into SPSS 17.0. The data was examined for association by performing a Spearman Rho test of correlation on each of the four categories. The significance level was set at .05.

Research Question Three

To explore the potential association between student perception of Clear Learning Targets and Academic Efficacy, the correlation between students’ response about Clear Learning Targets and Academic Efficacy was calculated using a Spearman’s Rho test of correlation.

Table 11 lists the number of student responses collected for Academic Efficacy as 337 and Clear Learning Targets as 336. The results of the test of correlation indicated there was a statistically strong correlation between Clear Learning Targets and Academic Efficacy. Table 11 lists the correlation coefficient as .443 with a significance level of .000.
The data presented in Table 11 showed a strong statistically significant correlation between Clear Learning Targets and Academic Efficacy. The data supported the hypothesis that there will be a statistically significant association between student perception of Clear Learning Targets and student Academic Efficacy. A significance level of .000 was able to reject the null hypothesis.

Table 11

Test of Correlation Between Academic Efficacy and Clear Learning Targets

<table>
<thead>
<tr>
<th></th>
<th>Academic Efficacy</th>
<th>Clear Learning Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman's rho</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic Efficacy</td>
<td>Correlation Coefficient</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>337</td>
</tr>
<tr>
<td>Clear Learning Targets</td>
<td>Correlation Coefficient</td>
<td>.443**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>336</td>
</tr>
</tbody>
</table>

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

Research Question Four

To explore the association between student perception of Progress Monitoring Information and Academic Efficacy, the correlation between students’ response about Progress Monitoring Information and Academic Efficacy was calculated using a Spearman’s Rho test of correlation.

Table 12 lists the number of student responses collected for Academic Efficacy as 337 and Progress Monitoring Information as 335. The results of the test of correlation
indicated there was a statistically strong correlation between Progress Monitoring Information and Academic Efficacy. Table 2 lists the correlation coefficient as .461 with a significance level of .000.

The data presented in Table 12 showed a strong statistically significant correlation between Progress Monitoring Information and Academic Efficacy. The data supported the hypothesis that there will be a statistically significant association between student perception of Progress Monitoring Information and Student Academic Efficacy. A significance level of .000 was able to reject the null hypothesis.

Table 12

*Test of Correlation Between Academic Efficacy and Progress Monitoring Information*

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>Academic Efficacy</th>
<th>Progress Monitoring Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>337</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Progress Monitoring Information</th>
<th>Correlation Coefficient</th>
<th>1.000</th>
<th>.461**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>335</td>
<td>335</td>
<td></td>
</tr>
</tbody>
</table>

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

*Research Question Five*

To explore the association between student perception of Clear Learning Targets and Eagerness to Learn, the correlation between students’ response about Clear Learning Targets and Eagerness to Learn was calculated.
Table 13 lists the number of student responses collected for Eagerness to Learn as 337 and Clear Learning Targets as 336. The results of the test of correlation indicated there was a statistically strong correlation between Clear Learning Targets and Eagerness to Learn. Table 13 lists the correlation coefficient as .409 with a significance level of .000.

The data presented in Table 13 showed a strong statistically significant correlation between Clear Learning Targets and Eagerness to Learn. The data supported the hypothesis that there will be a statistically significant association between student perception of Clear Learning Targets and Eagerness to Learn. A significance level of .000 was able to reject the null hypothesis.

Table 13

*Test of Correlation Between Eagerness to Learn and Clear Learning Targets*

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>Eagerness To Learn</th>
<th>Clear Learning Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>337</td>
</tr>
<tr>
<td>Clear Learning Targets</td>
<td>Correlation Coefficient</td>
<td>.409**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>336</td>
</tr>
</tbody>
</table>

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

To explore the potential association between student perception of Progress Monitoring Information and Eagerness to Learn, the correlation between students’ response about Progress Monitoring Information and Eagerness to Learn was calculated.

Table 14 lists the number of student responses collected for Eagerness to Learn as 337 and Progress Monitoring Information as 335. The results of the test of correlation indicate there was a statistically strong correlation between Progress Monitoring Information and Eagerness to Learn. Table 14 lists the correlation coefficient as .388 with a significance level of .000.

Table 14

Test of Correlation Between Eagerness to Learn and Progress Monitoring Information

<table>
<thead>
<tr>
<th>Spearman's rho</th>
<th>Eagerness To Learn</th>
<th>Progress Monitoring Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>337</td>
</tr>
<tr>
<td>Progress Monitoring Information</td>
<td>Correlation Coefficient</td>
<td>.388**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>335</td>
</tr>
</tbody>
</table>

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

The data presented in Table 14 showed a strong statistically significant correlation between Progress Monitoring Information and Eagerness to Learn. The data supported the hypothesis that there will be a statistically significant association between student
perception of Progress Monitoring Information and Eagerness to Learn. A significance level of .000 was able to reject the null hypothesis.
CHAPTER V
DISCUSSION OF FINDINGS

Introduction

Much research has been conducted examining the impact of formative assessment on students’ academic performance and improvement. Meta-analysis and early studies support, with large amounts of evidence, that using formative assessment in the classroom will have a large impact on student academic achievement, especially those students that are perennial low achievers (Black et al., 2004; Black & Wiliam, 1998b; Crooks, 1988; Herman et al., 2008; Kirton et al., 2007; Leahy et al., 2005). Formative assessment attributes are the critical features used in the teaching process to make the formative assessment process effective. Formative assessment is a process that can look differently in various classrooms. The philosophy is consistent and the attributes are the descriptions of the features of that philosophy. While past research has focused and supported formative assessment as a way to improve academic achievement, examining why formative assessment has that effect will further help educators implement innovations and instructional methods more effectively.

Purpose of the Study

The purpose of this study was to examine the impact that systematic use of formative assessment attributes in classroom instruction had on student affect and to examine the association of formative assessment attributes in the classroom with student
affect at a selected middle school located south of Chicago. The results from this study add to the body of knowledge for the public education reform that this country and this nation’s children desperately need. The presented data supports the belief that the increased perception of formative assessment attributes by students has a positive on student affect.

Findings

Research Question One

There were two null hypotheses established for the first research question: “is there a statistically significant impact on students’ Academic Efficacy when the systematic use of formative assessment is used in daily classroom instruction?” The result of the Mann-Whitney U nonparametric test was unable to show that the systematic use of formative assessment attributes by the teacher has any statistically significant impact on student Academic Efficacy.

The first null hypothesis, “there will be no statistically significant difference between experimental and control groups in student responses on the post administration of the Student Affect Inventory for Academic Efficacy,” is not rejected at the .05 level of significance. The data resulting from the Mann-Whitney U nonparametric test indicates that there is no statistically significant difference in Academic Efficacy between control and experimental groups.

The second null hypothesis, “there will be no statistically significant difference between the experimental group’s pre and post administration results on the Student Affect Inventory for Academic Efficacy,” is not rejected at the .05 level of significance.
The data resulting from the Mann-Whitney U nonparametric test indicates there is no statistically significant difference in Academic Efficacy of the experimental group’s pre and post inventory results.

*Research Question Two*

There were two null hypotheses established for the second research question: “is there a statistically significant impact on students’ Eagerness to Learn when the systematic use of formative assessment is used in daily classroom instruction?” The result of the Mann-Whitney U nonparametric test is unable to show that the systematic use of formative assessment attributes by the teacher has a statistically significant impact on student Eagerness to Learn.

The first null hypothesis, “there will be no statistically significant difference between Experimental and Control groups in student responses on the Student Affect Inventory for Eagerness to Learn,” is not rejected at the .05 level of significance. The data resulting from the Mann-Whitney U nonparametric test indicates that there is no statistically significant difference in Eagerness to Learn between control and experimental groups.

The second null hypothesis, “there will be no statistically significant difference between the experimental group’s pre and post administration results on the Student Affect Inventory for Eagerness to Learn,” is not rejected at the .05 level of significance. The data resulting from the Mann-Whitney U nonparametric test indicates there is no statistically significant difference in Eagerness to Learn of the experimental group’s pre and post inventory results.
The collected data is unable to show a significant difference. The study utilized a quasi-experimental design. Control and experimental groups were created by convenience grouping. Already intact student groups were chosen randomly to become the experimental group. Both groups were administered the Student Affect Inventory at the beginning and end of the study. The experimental group’s results were examined to determine any change in responses from pre to post administration of the survey. The post results of the control and experimental were examined for statistical difference.

While the experimental group did show a change from pre to post survey, the results are not statistically significant for Academic Efficacy or Eagerness to Learn. The experimental group did show a change in the category of Eagerness to Learn. The mean change was 13.16 to 16.29. The collected data is unable to reject the null hypothesis.

The researcher attributes the failure to reject the null hypothesis to two potential factors—a Type II error and the lack of a significant difference in use of formative assessment attributes between the control and experimental group teachers.

A Type II error occurs when there is a change in data but is not shown to be significant. In this case, the sample size for the experimental group was too small (n=16) and suffered atrophy by the end of the study (n=12) to register a statistically significant change. The experimental group showed an increase in the area of Eagerness to Learn but that increase failed to show as significant.

Teacher data was also collected to determine the teacher’s value and frequency of formative assessments in the classroom. Teachers from the control and experimental groups completed a formative assessment questionnaire at the beginning and end of the
study. The results of the questionnaire show that the use and value of formative assessments by the experimental teachers was not significantly different from that of the control group teachers. Both groups showed a minimal increase in use. The amount of professional development and time frame of the study may have had an impact on the use of formative assessment attributes by the experimental teachers.

*Research Question Three*

The study sought to answer the research question “is there a significantly significant association between students’ (control and experimental) perception of Clear Learning Target and Academic Efficacy?” The collected data shows an association with students perceiving Clear Learning Targets and the students’ report of Academic Efficacy. Students who reported perceiving Clear Learning Targets also reported a higher level of Academic Efficacy.

The null hypothesis, “there will be no statistically significant association between the student’s perception of Clear Learning Targets and Academic Efficacy,” is rejected at the .05 level of significance. The data resulting from the Spearman Rho test of Correlation indicates there is a strong statistically significant association between student perception of Clear Learning Targets and students’ Academic Efficacy.

The conclusion can be drawn that students felt they were more able to learn when they perceived Clear Learning Targets. Learning targets are the descriptions of skills, concepts, or knowledge students are expected to learn. According to S. McManus (2008), this attribute is essential for students to reach the next level of understanding. This data supports the work of Althoff et al. (2007). Althoff et al.’s study stated that
students who reported an increase in awareness of learning targets also reported an increase in their level of comfort in facing assessments. Chappuis and Stiggins (2002) make the claim that students need to respond to assessment information with the belief that they are able to continue to learn. The collected data is able to determine that, in this setting, there is an association with student responses to perceiving Clear Learning Targets and their Academic Efficacy of students in grades sixth through eighth.

Research Question Four

The study sought to answer the research question “is there a statistically significant association between students’ (control and experimental) perception of Progress Monitoring Information and Academic Efficacy?” The collected data shows an association with students’ perception of Progress Monitoring Information and the students’ report of Academic Efficacy. Students who reported perceiving Progress Monitoring Information as descriptive feedback also reported a higher level of Academic Efficacy.

The null hypothesis, “there will be no statistically significant association between the students’ perception of Progress Monitoring Information and Academic Efficacy,” is rejected at the .05 level of significance. The data resulting from the Spearman Rho test of Correlation indicates there is a strong statistically significant association between student perception of Progress Monitoring Information and students’ Academic Efficacy.

Feedback is most effective when it provides information and focus for the student on his or her progress towards mastery of an educational task and thus enhances self-efficacy and encourages effort (Black, McCormick et al., 2006; Black & Wiliam, 1998b;
Crooks, 1988; Stiggins & Chappuis, 2005). The conclusion can be drawn that students who perceived receiving more progress monitoring information as descriptive feedback felt as though they can be more successful in school. The negative can also be concluded—students who perceived receiving little to no progress monitoring information felt less likely to be successful in school. Black and Wiliam (1998b) claimed that the increase in formative assessment attributes, including progress monitoring information, has a positive impact on student academic performance—especially for struggling students. The collected data suggests that this dramatic increase in student achievement for struggling learners comes from the association between students’ perception of Progress Monitoring Information and student Academic Efficacy for students in grades sixth through eighth.

Research Question Five

The study sought to answer the research question “is there a statistically significant association between the students’ (control and experimental) perception of Clear Learning Target and Eagerness to Learn?” The collected data shows an association with students’ perception of Clear Learning Targets and their Eagerness to Learn. According to the collected data, students that reported understanding the learning targets also responded with a higher level of Eagerness to Learn.

The null hypothesis, “there will be no statistically significant association between the students’ perception of Clear Learning Target and Eagerness to Learn,” is rejected at the .05 level of significance. The data resulting from the Spearman Rho test of
Correlation indicates there is a strong statistically significant association between student perception of Clear Learning Targets and students’ Eagerness to Learn.

An extremely important issue that teachers and school leaders discuss is how to motivate students to learn. According to this data, there is an association with knowing the learning target and being motivated to reach that target. Chappuis and Stiggins (2002) claim that a student can hit the target as long as the student knows what it looks like and that it remains still. According to Campus and O’Hern (2007), students who are unclear about the learning targets show a lack of responsibility for their learning. Crooks (1988), Stiggins and Chappuis (2008), and Wiliam et al. (2004) state that the use of learning targets will stimulate the intrinsic motivation of the student. The collected data adds to the claims of these researchers by determining that, in this setting, there is an association with students’ perception of Clear Learning Targets and their Eagerness to Learn for students in grades sixth through eighth.

Research Question Six

The study sought to answer the research question “is there a statistically significant association between the students’ (control and experimental) perception of Progress Monitoring Information and Eagerness to Learn?” The collected data shows an association with students’ perception of Progress Monitoring Information and their Eagerness to Learn. According to the collected data, students that reported receiving Progress Monitoring Information through descriptive feedback also responded with a higher level of Eagerness to Learn.
The null hypothesis, “there will be no statistically significant association between the students’ perception of Progress Monitoring Information and Eagerness to Learn,” is rejected at the .05 level of significance. The data resulting from the Spearman Rho test of Correlation indicates there is a strong statistically significant association between student perception of Progress Monitoring Information and students’ Eagerness to Learn.

The role of progress monitoring information is to increase student achievement and student eagerness to learn. Black and Wiliam (1998a, 1998b, 2004) and Stiggins (1999, 2008) have shown research evidence to support the positive impact accurate, specific, and timely feedback about student progress can have on student achievement and motivation. The common belief was that grades motivated students to want to perform better. Research by Guskey (2002), Carlton (1992), and Davinroy et al. (1994) show that grades are often misunderstood and tend to motivate only the high achieving students while discouraging the struggling students. The collected data shows, in this setting, an association with Progress Monitoring Information and student Eagerness to Learn for students in grades sixth through eighth.

**Implications on Teaching and Learning**

Although the data was unable to show that the systematic use of formative assessments had an impact on student affect, it is this researchers belief that the inclusion of formative assessment as a process not only has an impact on student affect but also on student academic performance. The study did show associations between the two affective variables (academic efficacy and eagerness to learn) and two formative assessment attributes (progress monitoring information and clear learning targets).
Although the data is unable to show that increasing the use of these attributes will impact the affective variables, it is this researcher’s belief that good teaching pedagogy will include the student in the learning process. This inclusion addresses the student as a user of assessment information and helps refocus the student to the learning process instead of the point-collecting process.

The study did support the association of clear learning targets and progress monitoring information. As a response to this data, teachers should strive to increase the identification of learning targets to their students in a manner that students understand what skill or concept they are expected to be learning. Teachers should also increase the amount of opportunities for students to receive feedback through progress monitoring. Feedback that identifies where the student is on the continuum combined with a clear learning target can help redirect students to what they are “learning” and not just “doing.”

**Limitations of the Study**

This study was limited by the size and location of the site. The data was collected from one middle school and thus a narrow demographic. Without a large sampling from various schools with varying demographics, the results are only applicable to similar sites and student make-up. The data was represented from grades 6 through 8 in a middle school setting. While the results are promising for additional grades, additional data will need to be collected to attribute these results to different grade levels.

The study was unable to show a change in student perception by using a quasi-experimental design. The school setting itself limited the ability to create randomized groups. The researcher was able to utilize convenience grouping. The quasi-
experimental design relied on teachers to receive professional development and then apply that training to classroom instruction. The researcher had to rely on the classroom teachers to administer the intervention—formative assessment attributes—to the experimental group of students through class instruction. As it was determined through the examination of teacher questionnaires, those teachers did not use or value formative assessments more than the control group. The schedule of professional development was administered early during the semester in which the study was conducted. By providing the professional development during the study, the students received the full treatment for a shorter period of time. The teachers were not able to develop the attributes for classroom use in enough time to implement them for the full study.

The study was also limited by the duration. The study was in place for 18 weeks. This length of time is common to schools as one semester. An increase in the duration of the variable may have shown a significant change in student perception form pre administration of the survey to post administration.

While the study received a large number of student responses for the correlation study (n=337), the sample size of the experimental group was below the number acceptable to draw conclusions to a larger population (n=12). The researcher received 16 consenting participants for the experimental group. This number decreased to 12. The ideal size of 30 was not achieved.
Recommendations for Future Research

Due to the limitations of demographic information, it is recommended that this study be replicated in a variety of grade levels and diverse settings. The study was conducted with middle school students in grades sixth through eighth. It is recommended to replicate the study with both elementary and high school age students. Additional studies should examine for trends in socio-economic, ethnic, age, and environmental factors.

It is also recommended that the quasi-experimental study should be replicated utilizing a larger experimental group. This study was unable to establish an appropriate sample size for the experimental group to yield statistically significant results. It is recommended to replicate this study with an experimental sample size of 30 or more. It is also recommended that this study utilize a longer period for professional development. The results of the Formative Assessment Questionnaire for Teachers showed that the experimental teachers in this study that received the professional development did not utilize the attributes more than the control group. The increase in professional development may impact the fidelity of the treatment. The recommendation is for the professional development to last one school year.

It is recommended that the teachers in the experimental group receive the professional development in advance of the study. This researcher recommends providing the professional development during the first semester and collecting the final student data at the end of the second semester. Teachers would be provided with a full semester to learn and then begin to implement during the second semester.
The duration of this study was approximately 18 weeks. It is recommended to continue the study for a longer period of time—36 weeks. The longer period of time may allow for the change in student perception and affect to exhibit. The longer time frame for the study will allow teachers to perform the intervention with more fidelity and frequency.

Additional studies should be conducted to determine associations with different formative assessment attributes and academic efficacy and eagerness to learn. It is recommended to examine peer assessment, self assessment, and learning progressions to determine any association with academic efficacy or/and eagerness to learn.

**Conclusions**

According to the examination of data, it can be concluded that students in grades sixth through eighth, attending middle school with similar demographics who perceive Clear Learning Targets and Progress Monitoring Information will report a higher level of Eagerness to Learn and a higher level of Academic Efficacy.

A large amount of research has shown an association with the daily use of formative assessment attributes and an increase in student achievement on assessments. The collected data from this study suggests the possible reason that students who receive formative assessment attributes perform better on summative assessments is due to an increase belief that they can learn—Academic Efficacy—and the desire to continue to learn—Eagerness to Learn.

The collected data from this study adds to the existing body of knowledge surrounding the impact student perception of formative assessment attributes used by
classroom teachers can have on student affect. The data supports the efforts of teachers and school leaders to shift away from a heavy reliance on assessment of student learning and towards an increase in using assessments to help students learn.

The current methods of assessment and instruction do not desirably impact the nation’s struggling students. The nation has continued to lose ground in assessment results when compared to developed countries. Our current practice is most effective to only the top performing students. Despite a plethora of research and data suggesting a shift in pedagogy and assessment practices, schools have continued to administer a large amount of summative assessments and very infrequent formative assessments. In response to this practice, students have learned to play the game and “earn” enough points to pass or obtain a grade, while others have given up on education altogether. Students have shifted their focus from learning and placed it on collecting points and performing behaviors. If schools want to have an impact on student academic performance, they need to shift their philosophy to one that increases the use of formative assessments, encourages teachers to utilize formative assessment attributes in the classroom, and decrease the amount of summative assessments. Education should not be about the collection of “rights” and “wrongs”, but should rather be about students learning.
APPENDIX A

STUDENT AFFECT INVENTORY
Directions for Administration Beginning of Study

Today I want you to complete a brief inventory dealing with your experience in school. Please think about all of your classes when answering the questions. You will complete the inventory's items anonymously, so just enter your responses—don't put your name on the inventory or write anything else. There will be no way to tell how you responded, so please answer as honestly as you can.

There is a sample of how you are to respond given at the top of the inventory. As you can see, you are to circle the response-letters for the items depending on how you personally feel about each statement.

When you are finished, please place your completed inventory in the collection envelope that has been provided. Thank you for helping with this.
Directions for Administration End of Study

Today I want you to complete a brief inventory dealing with your classes this year—and only this year. Please think about all of your classes for this year while you answer the questions. You will complete the inventory's items anonymously, so just enter your responses—don't put your name on the inventory or write anything else. There will be no way to tell how you responded, so please answer as honestly as you can.

There is a sample of how you are to respond given at the top of the inventory. As you can see, you are to circle the response-letters for the items depending on how you personally feel about each statement.

When you are finished, please place your completed inventory in the collection envelope that has been provided. Thank you for helping me with this.
Please circle the correct answer

What grade are you currently in: 6th grade 7th grade 8th grade
Are you: Male or Female

THE WAY I SEE SCHOOL
Inventory for Grades 3-6

Directions: Please indicate how you feel about each statement as follows:
SA=Strongly Agree  A=Agree  U=Uncertain  D=Disagree  SD=Strongly Disagree
For example, here is how you would respond if you agree with the statement at the left:
I like to watch TV.                  SA  A  U  D  SD
When you complete the form, please place it in the large collection-envelope that’s provided.

1. I usually understand what I am supposed to learn.                   SA  A  U  D  SD
2. If I’m asked to learn new things, even if it’s difficult, I know I can learn it.  SA  A  U  D  SD
3. Typically, I don’t know if I am making progress as fast as I should.  SA  A  U  D  SD
4. I’m excited about learning new things in school.                     SA  A  U  D  SD
5. Very often, I’m not certain about what I supposed to be learning.  SA  A  U  D  SD
6. Lots of the time, I don’t look forward to learning new things in school.  SA  A  U  D  SD
7. Even if I get lots of help and plenty of time, it is hard for me to learn new things.  SA  A  U  D  SD
8. I get plenty of information to help me keep track of my own learning growth.  SA  A  U  D  SD

Thank you for completing this form.
Please circle the correct answer

What grade are you currently in: 6th grade  7th grade  8th grade

Are you: Male  Female

THE WAY I SEE SCHOOL
Inventory for Grades 7-9

Directions: Please indicate how you feel about each statement below as follows:
SA = Strongly Agree  A = Agree  U = Uncertain  D = Disagree  SD = Strongly Disagree
For example, here is how you would respond if you agree with the statement at the left:

I like to watch TV.  

When you complete the form, please place it in the large collection-envelope that’s provided.

1. I usually have a pretty good idea about what I am expected to learn in this class.  
SA  A  U  D  SD

2. When I’m asked to learn something new in this class, even if it’s difficult, I know I can learn it.  
SA  A  U  D  SD

3. Typically, I don’t have a very good idea if I am making enough progress in this class.  
SA  A  U  D  SD

4. I’m really excited about learning new things in this class.  
SA  A  U  D  SD

5. I often don’t have a clear idea in this class about what I am supposed to be learning.  
SA  A  U  D  SD

6. Most of the time, I don’t look forward to learning new things in this class.  
SA  A  U  D  SD

7. Even with help and plenty of time, I’m going to have difficulty learning new things in this class.  
SA  A  U  D  SD

8. In this class, I get enough information to keep track of my own learning achievement.  
SA  A  U  D  SD

Thank you for completing this form.
APPENDIX B

FORMATIVE ASSESSMENT QUESTIONNAIRE FOR TEACHERS
Appendix 1: Questionnaire

Formative Assessment

It would be useful to know the following information about you:

I teach in  
<table>
<thead>
<tr>
<th>Ks1</th>
<th>Ks2</th>
<th>Ks3</th>
<th>Ks4</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Please delete as appropriate

I have been teaching for  

| 0-2 | 2-10 | 10+ years |

Please delete as appropriate

Introduction

Much recent research indicates that formative assessment raises standards and in practice it could be working for you.

Please complete these sentences, by giving a reason.

Assessment is working well:

Assessment hinders:

Part one

Please circle the number and letter that most closely matches your opinion of the following strategies.

<table>
<thead>
<tr>
<th>Please circle how highly you value the following strategies:</th>
<th>Please circle how often you use the following strategies:</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
<td>1 = most lessons</td>
</tr>
<tr>
<td>B = valuable</td>
<td>2 = most days</td>
</tr>
<tr>
<td>C = no strong view</td>
<td>3 = weekly</td>
</tr>
<tr>
<td>D = of little value</td>
<td>4 = rarely</td>
</tr>
<tr>
<td>E = of no value</td>
<td>5 = never</td>
</tr>
</tbody>
</table>

Involving pupils in their learning

1. Telling pupils what you hope they will learn and (sometimes) why they are learning it

A B C D E 1 2 3 4 5

2. Inviting and building on pupils’ contributions

A B C D E 1 2 3 4 5

3. Setting up tasks designed to enable pupils to 'get on' by themselves

A B C D E 1 2 3 4 5

LSW14/12011
<table>
<thead>
<tr>
<th>4. Getting pupils to collaborate in groups on a joint outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5. Spurring pupils on by making encouraging but specific, focused comments, e.g. they are on the right lines and in what way</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. Getting a pupil to help another pupil</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

**Modelling quality**

<table>
<thead>
<tr>
<th>1. Choosing and showing pupils examples of pupils' work for learning purposes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Getting a pupil to show you how s/he had gone about something so you can diagnose error</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Getting a pupil to demonstrate to the class how s/he did something</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Getting a pupil to suggest ways something can be improved</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5. Providing formats and structures for writing or recording findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. Showing pupils a range of other pupils' work to make a judgement about performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. Showing pupils a range of other pupils' work to make a judgement about progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8. Showing pupils a range of other pupils' work to model (or exemplify) criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

**Giving feedback**

<table>
<thead>
<tr>
<th>1. Using probing questions to diagnose the extent of the pupils' learning</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Analysing completed work to work out why a pupil has or has not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Giving rewards only when achievement is satisfactory for that pupil (with specific comments referring to pupil's success)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Expressing approval when achievement is satisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = very valuable</td>
</tr>
<tr>
<td>B = valuable</td>
</tr>
<tr>
<td>C = no strong view</td>
</tr>
<tr>
<td>D = of little value</td>
</tr>
<tr>
<td>E = of no value</td>
</tr>
<tr>
<td>1 = most lessons</td>
</tr>
<tr>
<td>2 = most days</td>
</tr>
<tr>
<td>3 = weekly</td>
</tr>
<tr>
<td>4 = rare</td>
</tr>
<tr>
<td>5 = never</td>
</tr>
<tr>
<td>A B C D E</td>
</tr>
</tbody>
</table>
6. Making a conscious decision to avoid saying a pupil is wrong
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

7. Telling pupils what they have achieved with specific reference to their learning
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

8. Telling pupils what they have not achieved with specific reference to their learning
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

9. Describing why an answer is correct
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

10. Specifying a better/different way of doing something
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

11. Writing an evaluative note on a pupil's work for the pupil
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

**Self assessment**

1. Getting pupils to suggest ways they can improve
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

2. Negotiating a route to improve something
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

3. Providing time for pupils to reflect and talk about their learning
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

4. Getting pupils to review their own work and record their progress
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

5. Helping pupils to understand their achievements and know what they need to do next to make progress
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never

6. Providing opportunities for pupils to assess their own and one another's work and give feedback to one another
A: very valuable  1: most lessons
B: valuable       2: most days
C: no strong view 3: weekly
D: of little value 4: rarely
E: of no value    5: never
APPENDIX C

EMAIL AUTHORIZATION TO USE STUDENT AFFECT INVENTORY
Ruland, Joshua

From: Siggins, Richard [RSiggins@ETS.ORG]
Sent: Monday, May 18, 2009 11:16 AM
To: Ruland, Joshua
Subject: RE: Assessing Student's Affect

Hi Josh

You do have the current version. We have not been conducting ongoing technical research on the instrument. So that may be work for you to complete in your research. But you certainly do have our permission to proceed. The instrument is in the public domain and available for this kind of work.

Rick

From: Ruland, Joshua [mailto:rueland@manteno5.org]
Sent: Monday, May 18, 2009 8:53 AM
To: Siggins, Richard
Subject: Assessing Student's Affect

Dr. Siggins,

I am the curriculum director for Manteno CUSD #6, a rural K-12 district South of Chicago. I was hired last year and am moving our district towards Assessment for Learning practices. I am in complete agreement with this philosophy of assessment and know it is what is best for students. I have several teachers in each of our buildings who have formed learning teams and are piloting strategies aligned with Afl. I would like to use the survey "The Way I See School" for my teachers to collect information as well as myself. I will be asking all students in the grades levels with pilot teachers (variable and control) to administer the survey. They will again administer it in December and May. While it will be helpful for the teachers to see the results, I am also looking forward to the results to help reinforce the benefit of Afl to other teachers. I understand that you may be revising the survey. I would like to use the most up to date version. I have the version released through The Council of Chief State Officers. Is there a more current version?

I was in attendance at the Solution-Tree Assessment Summit in Atlanta, GA with several of my teachers. It was an extremely beneficial trip for us. I am currently working on my dissertation on assessment. My research question is essentially about the impact Afl strategies have on students' perceptions of school and learning. I was very excited to see this tool. I have read much literature on the impact of student achievement but am more interested how it affects student attitudes. I would like to use this survey in my dissertation research as well. Do you have any reliability statistics generated as of yet? I would also like to share with you the results from my research when I complete it (possible December-May 2011). I will be collecting research from a middle school setting in which a middle school team incorporates Afl, attributes to their practices. I will be comparing their students' results to the remaining teams that will be utilizing the traditional methods. I am interested in the impact the attributes and strategies will have on student affect and perception on learning. I would like to use your survey along with student and teacher interviews to collect information among other methods. I look forward to sharing my results.

Thank you for being an advocate for our students. I have heard the call and have take to action.

Joshua Ruland
Curriculum and Assessment Director
Manteno CUSD #6
815929-7024

This e-mail and any files transmitted with it may contain privileged or confidential information.
APPENDIX D

AUTHORIZATION LETTER TO USE FORMATIVE ASSESSMENT QUESTIONNAIRE
25/03/10
LICENSEE

Joshua Ruland
Loyola University Chicago
553 Barry Turn
Manteno
IL
60550

The Qualifications and Curriculum Authority has pleasure in granting the Licensee permission in this licence agreement to reproduce without any alternation or omission the materials below in the Licensee’s work on the terms specified. The permission granted, which includes transcription into Braille, is non-transferable, and does not include the right to issue reprinted editions of the Licensee’s work under another imprint or to publish new or revised editions of it in any format whether or not they contain the material different from that contained in the original edition.

LICENSOR’S WORK(S):
Report on Teachers Perception of Formative Assessment

MATERIAL TO BE REPRODUCED:
Extract of the above title

LICENSEE’S WORK:
Educational Research for Doctoral Dissertation

TERRITORIAL RIGHT GRANTED:
For use in the above research study

FEE:
£0.00 (free of charge) exclusive of VAT, which will be charged at the prevailing rate.

If a fee has been incurred, please do not send payment now. Payment should be made on invoice and should be remitted within one calendar month of invoice date, failure which the permission granted will automatically terminate.

ACKNOWLEDGEMENT:
As per licensors work(s)

REFERENCE COPY:
One reference copy of the licensee’s work should be sent to the licensor on publication.

Lee Lennon
APPENDIX E

TEACHER PARTICIPANT CONSENT LETTER
CONSENT TO PARTICIPATE IN RESEARCH

Project Title: The Effect of Formative Assessment on Student Academic Efficacy and Eagermess to Learn
Researcher(s): Joshua Ruland
Faculty Sponsor: Dr. Brigid Schultz

Introduction:
You are being asked to take part in a research study being conducted by Joshua Ruland for a doctoral dissertation under the supervision of Dr. Brigid Schultz in the Department of Education at Loyola University of Chicago.

You are being asked to participate because you are a member of the teaching staff in charge of instructing students utilizing formative assessment attributes in classroom instruction.

Please read this form carefully and ask any questions you may have before deciding whether to participate in the study.

Purpose:
The purpose of this study is to discover the impact using formative assessment attributes in daily classroom instruction has on students' academic efficacy and eagerness to learn.

Procedures:
If you agree to be in the study, you will be asked to:
- Complete a Formative Assessment Questionnaire for Teachers at the beginning of the study and at the end of the study after approximately 18 weeks. The Questionnaire will take about 30 minutes to complete and you will have three days to complete the questionnaire at the beginning of the study and three days from the end of the study.
- If you are assigned to the variable group, participate in professional development meetings (5 sessions) at no cost to you. Variable group is determined by the grouping of students in the looping method assignment.

Risks/Benefits:
There are no foreseeable risks involved in participating in this research beyond those experienced in everyday life.

The nature of this research will benefit society by adding to the body of evidence surrounding formative assessment. The focus of this study will help future educators and researchers address pedagogy by examining the impact instructional practices may have on student affect.

Participants in the variable group will gain professional development for classroom instruction utilizing formative assessment.
Confidentiality:
- Questionnaires will be coded by variable and control group. No names or specifically identifying information will be collected.
- Questionnaires will be completed via an online service. This service meets all the requirements for security and protection of privacy as established by Loyola University’s Institutional Review Board.
- Questionnaire data will be kept securely in the online database and in a locked file cabinet at the researcher’s office. Only the researcher will have access to the questionnaire responses.

Voluntary Participation:
Participation in this study is voluntary. If you do not want to be in this study, you do not have to participate. Even if you decide to participate, you are free not to answer any question or to withdraw from participation at any time without penalty.

Contacts and Questions:
If you have questions about this research project or interview, feel free to contact Joshua Ruland at Jruland@lac.edu or 815/928-7024 or the faculty sponsor Dr. Brigid Schultz at Bshultz@lac.edu.

If you have questions about your rights as a research participant, you may contact the Loyola University Office of Research Services at (773) 508-2689.

Statement of Consent:
Your signature below indicates that you have read the information provided above, have had an opportunity to ask questions, and agree to participate in this research study. You will be given a copy of this form to keep for your records.

Participant’s Signature          Date

Researcher’s Signature          Date
APPENDIX F

PARENTAL CONSENT LETTER
CONSENT TO PARTICIPATE IN RESEARCH
( Parental Consent)

Project Title: The Effect of Formative Assessment on Student Academic Efficacy and Eagermess to Learn
Researcher(s): Joshua Ruland
Faculty Sponsor: Dr. Brigid Schultz

Introduction:
You are being asked to take part in a research study being conducted by Joshua Ruland for a doctoral dissertation under the supervision of Dr. Brigid Schultz in the Department of Education at Loyola University of Chicago.

Your child is being asked to participate because your school is implementing instructional methods that have been shown through research to have a positive impact on student achievement. This method includes the student in the assessment process and as a user of assessment information. This study is collecting information about how students feel about certain classroom teaching methods.

Please read this form carefully and ask any questions you may have before deciding whether to allow your child to participate in the study.

Purpose:
The purpose of this study is to examine the impact of including students in the process of assessment has on students' eagerness to learn and belief that they can learn.

Procedures:
If you agree to allow your child to be in the study, he/she will be asked to:
• Answer approximately 8 questions at the beginning of the year and again at the end of the study (about 18 weeks)
• The student affect inventory is approximately 8 questions and will take less than 10 minutes to complete. The questions require students to select an answer on a scale of 1-5 about how they feel about a particular form of classroom instruction.

Risks/Benefits:
There are no foreseeable risks involved in participating in this research beyond those experienced in everyday life.

There are no direct benefits to your child from participation, but the nature of this research will benefit society by adding to the body of evidence surrounding formative assessment. The focus of this study will help future educators and researchers address classroom instruction by examining the impact instructional practices may have on how students feel about assessments.
Confidentiality:
- No identifiable information will be collected from you. You will complete the inventories anonymously.
- Student affect inventories will be coded for gender, grade level, and team of teachers. You will complete the inventory anonymously.
- The completed inventories will be kept in a secure file cabinet in the researcher’s office. The researcher will be the only one with access to the inventory results.

Voluntary Participation:
Participation in this study is voluntary. If you do not want to be in this study, you do not have to participate. Even if you decide to participate, you are free not to answer any question or to stop participating at any time without penalty. You will have no negative consequences from either the researcher or the school if you choose to stop.

Contacts and Questions:
If you have questions about this research project, feel free to contact Joshua Ruland at Jruoland@luc.edu or 815/928-7024 or the faculty sponsor Dr. Brigid Schultz at Bschull@luc.edu.

If you have questions about your rights as a research participant, you may contact the Compliance Manager in Loyola’s Office of Research Services at (773) 508-2689.

Statement of Consent:
Your signature below indicates that you have read and understood the information provided above, have had an opportunity to ask questions, and agree to allow your child to participate in this research study. You will be given a copy of this form to keep for your records.

Student’s Signature ____________________________ Date ____________

Researcher’s Signature ____________________________ Date ____________
APPENDIX G

STUDENT PARTICIPANT ASSENT LETTER
CONSENT TO PARTICIPATE IN RESEARCH
(Student Assent)

Project Title: The Effect of Formative Assessment on Student Academic Efficacy and Eagerness to Learn
Researcher(s): Joshua Ruland
Faculty Sponsor: Dr. Brigid Schultz

Introduction:
You are being asked to take part in a research study being conducted by Joshua Ruland for a doctoral dissertation under the supervision of Dr. Brigid Schultz in the Department of Education at Loyola University of Chicago.

You are being asked to participate because your school is implementing instructional methods that have been shown through research to have a positive impact on student achievement. This method includes students in the assessment process and as a user of assessment information. This study is collecting information about how students feel about certain classroom teaching methods.

Please read this form carefully and ask any questions you may have before deciding whether to participate in the study.

Purpose:
The purpose of this study is to examine the impact of including students in the process of assessment has on how students feel about assessment and different types of classroom lessons.

Procedures:
If you agree to be in the study, you will be asked to:
• Answer approximately 8 questions at the beginning of the year and again at the end of the study (about 18 weeks)
• The student affect inventory is approximately 8 questions and will take less than 10 minutes to complete. The questions require students to select an answer on a scale of 1-5 about how they feel about a particular form of classroom instruction.

Risks/Benefits:
There are no foreseeable risks involved in participating in this research beyond those experienced in everyday life.

There are no direct benefits to you by participating, but the nature of this research will benefit society and schools by adding to more information about how students react to different classroom lessons. The focus of this study will help future educators and researchers address classroom instruction by examining the impact instructional practices may have on how students feel about assessments.
Confidentiality:
- No identifiable information will be collected from students. Students will complete the inventories anonymously.
- Student affect inventories will be coded for gender, grade level, and team of teachers. Students will complete the inventory anonymously.
- The completed inventories will be kept in a secure file cabinet in the researcher’s office. The researcher will be the only one with access to the inventory results.

Voluntary Participation:
Participation in this study is voluntary. If you do not want your child to be in this study, he/she does not have to participate. Even if you decide to allow your child to participate, he/she is free not to answer any question or to withdraw from participation at any time without penalty. Your child will have no negative consequences from either the researcher or the school if you or your child chooses to withdraw from the study.

Contacts and Questions:
If you have questions about this research project or interview, feel free to contact Joshua Ruland at jruland@luc.edu or 815/928-7024 or the faculty sponsor Dr. Brigid Schultz at Bschultz@luc.edu.

If you have questions about your child’s rights as a research participant, you may contact the Compliance Manager in Loyola’s Office of Research Services at (773) 508-2689.

Statement of Consent:
Your signature below indicates that you have read and understood the information provided above, have had an opportunity to ask questions, and agree to allow your child to participate in this research study. You will be given a copy of this form to keep for your records.

<table>
<thead>
<tr>
<th>Parent's/Guardian's Signature</th>
<th>Date</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Researcher's Signature</th>
<th>Date</th>
</tr>
</thead>
</table>
APPENDIX H

FORMATIVE ASSESSMENT SCHEDULE OF SESSIONS
Formative Assessment Professional Development

**Session 1  Full Day In-Service  09/01/2010**

9:00-11:00

Overview of the five attributes and related research concerning formative assessment as a process

11:00-12:00

Lunch

12:00-2:00

Classroom policies and procedures that encourage learning and assessment as a tool for learning

**Session 2  2 hours  09/22/2010**

1:00-3:00

Articulating learning targets to students

Creating assessments aligned to targets

Strategies to collect information to make instructional decisions

**Session 3  2 hours  09/29/2010**

1:00-3:00

Identify learning targets to students

Provide descriptive feedback through progress monitoring

Learning continuum
Session 4  2 hours  09/20/2010
Involving students in the assessment, peer assessment, and self assessment
Strategies for students to receive feedback and make decisions about learning and studying

Session 5  2 hours  10/06/2010
Strategies to develop a collaborative classroom climate
Strategies to encourage student engagement in the decision-making process and to minimize the atmosphere of fear

Resources


VITA

Joshua Ruland is the son of S. D. Ruland and Judy Rae Brandt. He was born in Ashtabula, Ohio on June 3, 1978. He currently resides in a rural community south of Chicago with his wife and three children.

Joshua Ruland attended a rural public school. He graduated from the University of St. Francis in 2000 with a Bachelor of Arts in Psychology and English. In 2003 Joshua Ruland earned a Master of Education degree in Teaching from the University of St. Francis. In 2008, he completed a Type 75 School Administration Certificate Program at Loyola University.

Joshua Ruland has worked in the field of education for the past 10 years. He began his career as an educator as a middle school language arts and social studies teacher in a small rural school district and soon moved to teaching high school English in a large suburban school district. He has worked for Manteno School District No.5 for the last four years as the Director of Curriculum.

Joshua Ruland is an active member of ASCD and ILASCD. He is an active presenter and speaker at numerous curriculum and education renewal conferences and workshops.
DISSENTATION COMMITTEE

The Dissertation submitted by Joshua Ruland has been read and approved by the following committee:

Brigid Schultz, Ed.D., Director
Clinical Assistant Professor, School of Education
Loyola University Chicago

Marla Israel, Ed.D.
Associate Professor, School of Education
Loyola University Chicago

Ann Marie Ryan, Ph.D.
Associate Professor, School of Education
Loyola University Chicago