

INFLUENCE OF ASSESSMENT FOR LEARNING PROFESSIONAL DEVELOPMENT IN  
RURAL GEORGIA PUBLIC SCHOOLS

A Dissertation

Presented to

The Faculty of the School of Education

Liberty University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

by

Marianne W. Cole

August 2010

Influence of Assessment for Learning Professional Development in Rural Georgia Public  
Schools

By Marianne W. Cole

APPROVED:

COMMITTEE CHAIR

Jeff Crawford, Ed.D.

COMMITTEE MEMBERS

John Bartlett, Ed.D.

Georgia G. Evans Ed.D.

CHAIR, GRADUATE STUDIES

Scott B. Watson, Ph.D.

### **Abstract**

Marianne W. Cole. INFLUENCE OF ASSESSMENT FOR LEARNING PROFESSIONAL DEVELOPMENT IN RURAL GEORGIA PUBLIC SCHOOLS. (Under the direction of Dr. Jeffery Crawford) School of Education, August, 2010.

This study investigated the effect of two models of professional development concerning Assessment for Learning on teacher perception of the effectiveness of Assessment for Learning strategies and student achievement as measured by standardized Georgia End of Course Tests. The study hypothesized that a positive relationship exists between teacher perception of the benefit of Assessment for Learning strategies and increased student achievement as measured by student performance on standardized Georgia End of Course Tests. The study hypothesized that a positive relationship exists between teacher participation in Assessment for Learning professional development and teacher positive perception of the benefit of Assessment for Learning strategies. The sample consisted of 174 teachers and 2,787 students in core content courses in grades 9-12 in rural Georgia schools. Teacher participants received training on Assessment for Learning in either a theory-based or application-based professional development class or were part of the control group not participating in Assessment for Learning professional development. Teacher participants completed a detailed survey to gather assessment perception data. Students participated in instruction and completed Georgia End of Course Tests (EOCT), standardized achievement assessments, to gather performance data. Multiple ANOVA were used to statistically analyze the data and a relationship was found between teacher participation in formative-assessment professional development and student performance on standardized Georgia End of Course Tests.

### **Dedication**

To Jimmy,

You have always encouraged me, always believed in me, and always supported me (even when it was hard). You've been there through years of study for me to reach this goal and never once complained about the extra duties you had to shoulder. I love you. Always.

To Benjerman, Christina, and Harrison,

You, more than anyone, have sacrificed to reach this goal. I am grateful to have been blessed with such wonderful children. I love you. May you benefit from this accomplishment just as much as you have sacrificed for me to reach this goal.

To Gummy,

You have always been in my corner. Your understanding and support have allowed me to be true to myself. When others wondered why I would do such a thing you have always understood. Thank you.

To MawMaw,

All my life you have provided a peaceful place to rest. This time was no different. Thanks for letting your home be mine whenever I needed it.

To Mom,

Even when you didn't understand why, you were always willing to help. Thanks for all the extras—for being my Mom and my best friend. I'd never make it without you.

## **Acknowledgments**

In addition to my family, who supported me throughout this journey, there are others who I must acknowledge or be remiss.

To Dr. Jeff Crawford, You always calmed my fears and provided a voice of reason amid the chaos. Your support and guidance were invaluable. I could not have done this without you and your prayers.

To Dr. Georgia Evans, You challenged me to do anything I set my mind to. Over the years you have inspired me to reach for my goals and met them. Your knowledge and work ethic have left me in awe. I am proud to call you my colleague and blessed to call you my friend. Thank you for being my sounding board anytime I needed you.

To Dr. John Bartlett, Your comments were always laser sharp and completely on target. Thanks for your time.

To Hannah Allison, I would never have completed this without your edits and comments. Your insights into the APA process and your instructions on good writing have improved my skills immensely.

**Table of Contents**

Abstract	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vi
Chapter One – Introduction	1
Statement of the Problem	1
Research Proposal	4
Research Questions	5
Research Hypotheses	5
Background	6
Significance	8
Terminology	10
Chapter Two – Review of Literature	16
The Current Era	16
Defining Assessment	19
Assessment in Practice	24
Assessment for Student Success	34
Need for the Study	38
Chapter Three – Methodology	41
Research Design	41
Research Questions	41

Research Hypothesis	42
Population	42
Instrumentation	48
Procedures	51
Chapter Four- Results	58
Descriptive Statistics	58
Statistical Analysis of Research Questions	66
Data Summary	71
Chapter 5- Summary and Discussion	72
Statement of the Problem	72
Review of the Methodology	73
Summary of the Findings	73
Interpretation of the Findings	75
Conclusions	77
Implications	77
Research Applications	78
Limitations	78
Recommendation for Further Research	80
Summary Thoughts	81
References	83
Appendices	89
Appendix A- Letter inviting participation to school systems	89

Appendix B- Follow-up letter to school principals	90
Appendix C- Presentations for Professional Development	91
Appendix D – Bol Questionnaire	93
Appendix E- Study Questionnaire	94
Appendix F- IRB Documents	98
Appendix G- All Data Tables	107



## **Chapter One: Introduction**

Over the last 100 years research has consistently shown that effective formative assessment techniques, used to adjust instruction and provide student feedback, can and do improve student achievement and learning. Research continues to show, despite this knowledge, teachers are not implementing effective formative assessment techniques. Multiple reasons exist for the disconnect between best practices demonstrated in research and implemented practice found in classrooms. Reasons may include issues with inadequate professional development, a negative perception of formative assessment, a lack of understanding, or apathy.

This researcher believes that if educators understood the benefits of effective formative assessment and how to implement effective formative assessment in their classrooms, then their practices would change. This researcher also believes if educators understand the impact their perceptions have on the success of formative assessment techniques, then they would work to implement effective formative assessment in their classrooms. The focus of this research was on providing data to support this assertion.

### **Statement of the Problem**

With the increased focus of society and educational agencies on standardized test scores, great scrutiny has been placed on assessment. Reeves (2005) noted,

Today, all fifty states have academic content standards and some form of testing based on those standards. The No Child Left Behind Act represents the most sweeping federal education legislation in more than three decades. Although the Act remains controversial on many counts and is certain to be a campaign

issue [...] the fact remains that more than 90 percent of members of Congress from both political parties voted for the law in 2000 [...] But whatever changes may be made to the law, the four essential elements of No Child Left Behind – standards, accountability, testing, and choice – are very likely to remain. (p 1)

Teachers are looking for ways to increase student achievement. Stiggins (1999) reminded educators that the responsibility for academic progress does not reside with the teacher, principal, superintendent, or parent, but with the learner. Assessment for Learning provides a model focused on increasing student achievement. Effective classroom assessment allows educators to examine student learning and to gain information that effects student learning (Davies, 2000). Assessment for Learning allows educators to develop a complete picture of student understanding and mastery. By identifying both the benefits of Assessment for Learning and the student achievement gains that this model provides, this researcher realized a need for assessment other than summative standardized test results. The challenge to educators is to keep students from losing confidence in themselves and to develop confidence in those who have lost it (Stiggins, 1999). Formative assessment is one means to support increased student learning and a positive self concept by providing an alternative to the traditional use of assessment to grade and rank students avoiding the side effect of “ poorer performance and lowered academic pursuits for students seeking comparative rankings” (Reeves, 2005, p 181).

Research suggests that classroom assessment regularly focused on descriptive feedback, student self assessment, and peer assessment with clear expectations and guidelines is beneficial to student achievement and that communicating this research is imperative to the effort to promote teachers’ use of Assessment for Learning. Stiggins, Arter, Chappius, and Chappius (2006) found,

Dozens of studies conducted at all levels of instruction offer evidence of strong achievement gains in student performance as measured by standardized tests (Bloom, 1984; Black & Wiliam, 1998; Black, 2003; Meisel, Atkins- Burnett, Xue, Bickel & Hon, 2003; Rodriguez, 2004). The effect of assessment for learning on student achievement is some four to five times greater than the effect of reduced class size (Ehrenberg, Brewer, Gamoran, & Willms, 2001). Few interventions in education come close to having the same level of impact as Assessment for Learning. But the most intriguing result is that, while all students show achievement gains, the largest gains accrue to the lowest achievers. Everyone wins, with those who have the most to win, winning most. (p 37)

This researcher believes, if educators understood the benefits to student performance that the Assessment for Learning techniques provide, the strategies would be incorporated into daily classroom instruction.

Currently classroom assessment sits at the intersection of instruction, classroom management, and assessment (Brookhart, 2004). This researcher believes this can lead to confusion and tension among educators as they attempt to understand and better use classroom assessment to improve student learning. Harlan (2005) noted the interaction between formative and summative assessment and the possibility for negative or positive interaction based on teacher judgment and teacher knowledge of assessment. Part of the problem surrounding the use of good, solid assessment techniques lies within the current use of assessment results by politicians, agencies, and governments. Heritage (2007) stated,

Formative assessment, if used effectively, can provide teachers and their students with the information they need to move learning forward. But after more than a

hundred years of exhortations and a significant body of research on the topic, the idea that assessment and learning are reciprocal activities is still not firmly situated in the practice of educators. Instead, assessment is often viewed as something in competition with teaching, rather than as an integral part of teaching and learning. In our current accountability environment, assessment is not regarded as a source of information that can be used during instruction. Instead, it has become a tool solely for summarizing what students have learned and for ranking students and schools. In the process, the reciprocal relationship between teaching and assessment has been lost from sight. In a context in which assessment is overwhelmingly identified with a competitive evaluation of schools, teachers, and students, it is scarcely surprising classroom teachers identify assessment as something external to their everyday practice. (p 140)

Additionally, Reeves (2001) observed,

Educators are in a vicious cycle that hurts; Standards and assessments have convinced many teachers and administrators to abandon effective curriculum and instruction and pursue mindless test drills; test scores remain unsatisfactory; policymakers demand yet more tests; administrators in turn become even more focused on test prep to the expense of thinking, reasoning, and writing and test scores get worse. (p 5)

### **Research Proposal**

The aim of this research was to further study the relationship between teacher participation in targeted professional development and teacher use of formative assessment. The research was also designed to study the relationship between teacher perception of formative

assessment and student achievement as measured by student summative assessment performance on the standardized Georgia End of Course Tests. An exhaustive review of the literature has identified a clear link between sound classroom assessment practices and student achievement gains and, consequently, increased student performance on summative assessments. Review of the literature has also indicated a disconnect for teachers in understanding the link between classroom assessment and student learning and, therefore, student gain on summative assessment measures. The belief of this researcher is that further study of the relationship between professional development, teacher perception, and student performance may provide additional information about formative assessment.

### **Research Questions**

The following questions were addressed in this research study.

RQ<sub>1</sub>. Is there a relationship between teacher participation in Assessment for Learning professional development and student achievement as measured by the Georgia End of Course Tests standardized assessment?

RQ<sub>2</sub>. Is there a relationship between teacher participation in Assessment for Learning professional development and teacher perception of the benefits of Assessment for Learning strategies?

RQ<sub>3</sub>. Does a relationship exist between teacher perception of the benefit of formative assessment and student achievement as measured by the Georgia End of Course Tests standardized assessment?

### **Research Hypothesis**

Based on a review of the literature hypotheses were developed related to the research questions.

### **Hypotheses RQ<sub>1</sub>.**

*H<sub>01</sub>*. There will be no significant difference between teachers who participated in Assessment for Learning professional development those that did not as shown by student performance on the Georgia End of Course Tests standardized assessment.

### **Hypothesis RQ<sub>2</sub>.**

*H<sub>02</sub>*. There will be no significant difference between teachers who participated in Assessment for Learning professional development as measured by teacher perception of the benefits of Assessment for Learning strategies.

### **Hypotheses RQ<sub>3</sub>.**

*H<sub>03</sub>*. No significant relationship exists between teacher positive perception of the use of formative assessment and student achievement as measured by student performance on the Georgia End of Course Tests standardized assessment.

### **Background**

This researcher's background is in elementary instruction, professional learning, curriculum development, and testing. Working in instruction and assessment provided a unique view into changes as summative assessment more directly impacted schools and teachers with the implementation of the No Child Left Behind Act and the related accountability for schools and teachers. Working with teachers to research and implement best practices this researcher saw resistance and lack of understanding exhibited to standards-based classroom assessment practices. This researcher's observations resulted in an interest in the connection between professional learning, teacher perception and the success of formative assessment practices in the classroom. Susan Bucci's comments as part of her 2002 dissertation inspired this researcher when she stated,

I suggest, as a final reflection, that we are at the threshold of a major transformation in assessment. The 21<sup>st</sup> Century presents unprecedented technological, knowledge, scientific, and moral revolutions that require human beings to reach beyond confining boundaries. Quite literally, a “different brain” is entering our classrooms (Sousa, 2001) borne from the technology that shapes our present times. The demands on this ‘new’ brain are those of capability, flexibility, creativity and efficacy. Understanding how to assess this new brain using formative assessment techniques while understanding that assessment is integral to learning is the main challenge of the classroom teacher. (pp 7-8)

A challenge was presented for this researcher to better understand formative assessment, its connection to learning, and how professional learning opportunity and teacher perception impacts the success of formative assessment techniques in the classroom. At this same time, the researcher was participating in in-depth study of the works of assessment researchers, particularly the collaborative work of several in the book *Ahead of the Curve* (Reeves, Ainsworth, Almeida, Davies, DuFour, Gregg, & Guskey, 2007). In *Ahead of the Curve* (2007) Reeves’ preface questioned the ability of educators to change practices in order to rise to the challenge of educating the current generation more effectively. Reeves (Reeves, et.al., 2007) called for an examination of effective practice versus popular practice and also presented another question which resonated with this researcher when he stated, “The essential question is not ‘What is the proof?’ but rather, ‘What is the risk if we engage in this change compared to the risk of continuing our present practice?’” (p 7).

## Significance

The point of schooling is to increase student achievement, to support students' learning and growth. Wiggins and McTighe (2008) noted that the mission of schools is not covering content but helping students "become thoughtful about, and productive with, content. It is not to help students get good at school, but rather to prepare them for the world beyond school – to enable them to apply what they have learned to issues and problems they will face in the future" (p10). Research (Bloom, 1984, Black & Wiliam, 1998, Stiggins, 2002, Stiggins, 2005, Guskey, 2005, and Wiliam, 2006) has pointed to the successful use of formative assessment in improving student learning. Black and Wiliam's 1998 research synthesis reported formative assessment produced gains with effect sizes between 0.4 and 0.7. The synthesis showed schools using formative assessment that began by offering students a clear picture of learning targets, providing students feedback related to the learning target, engaged students in self- assessment, and provided students with an understanding of specific steps to take to improve (Chappius, 2005). Despite overwhelming research to support the benefits of formative assessment, classroom teachers have not embraced or effectively implemented sound assessment practices in classrooms. Dorn (2010) noted "The literature base on using formative assessment for instructional and intervention decisions is formidable, but the history of the practice of formative assessment is spotty" (p 325). This researcher believes a better understanding of the relationship between targeted professional learning, teacher perception of formative assessment, and the implementation of formative assessment techniques in the classroom may provide additional data as researchers attempt to understand this resistance. Tierney (2006) summarized this problem when stating,

The use of classroom assessment to promote student learning is strongly



supported by current educational research. Following the seminal review by Black and Wiliam (1998), a host of empirical work has confirmed the pedagogical potential of classroom assessment (e.g. Black & Harrison, 2001; Bariitchi & Keshavarz, 2002; Orsmong et al.,2002; Coffey, 2003; Lee & Gavin, 2003; Waddel, 2004). Despite this research evidence, the sustained championing by assessment specialist [sic] (e.g., Stiggins, 1994, 2001; Stiggins & Chappius, 2005), and the increasing endorsement by professional organizations (Joint Committee on Standards for Educational Evaluation, 2003; Miller, 2005), assessment is still not widely used to promote learning in elementary and secondary classrooms. On the contrary, the summative function is emphasized, and teachers continue to use classroom assessment primarily for grading and reporting (Kehr, 1999; McNair,et al., 2003;Uchiyama, 2004). Although the rhetoric for assessment reform is strong, the way in which student learning is assessed in classrooms on a regular basis seems resistant to change. (p239)

The focus of this research was to further examine the link between professional development and the impact teacher perception concerning assessment may have on student growth in an effort to answer the challenge of assessment “gurus” in this nation. The leading researchers in assessment called for a “redirection of assessment to its fundamental purpose: the improvement of student achievement, teaching practice, and leadership decision-making” (Reeves, et al., 2007, p1). Additionally, sound assessment practice should provide stakeholders (students, parents, teachers, and supervisors) with information about how the student is doing by providing students with an opportunity to improve achievement and keeping an individual record of student achievement of standards (Reeves, 2005). Currently, however, a class often functions as follows, the teacher teaches then tests then moves on, leaving unsuccessful students to finish

last, on the premise that comparing unfavorably to others will motivate students to perform better in the future (Chappius & Stiggins, 2002). On the contrary, Assessment for Learning occurs during the teaching and learning process, providing students feedback, the time and ability to self correct, and the opportunity to receive additional support for mastery of the learning goal (Chappius & Stiggins, 2002). In Assessment for Learning teachers and students use formative assessment information to pretest and adjust instruction for individuals, analyze who needs more practice, revise instruction continually, reflect on effectiveness of teaching practices, confer with students concerning strengths and areas for improvement, and facilitate peer tutoring (Chappius & Stiggins, 2002).

The state of Georgia and particularly the West Georgia region were implementing standards-based classroom practices which included training and implementation of standards-based formative assessment within targeted classrooms. Therefore, the population was the ideal target for measuring the gains of student achievement and the relationship to teacher implementation of formative assessment practices within the classroom. At the time of the study the West Georgia area was uniquely suited to provide an ideal environment within which to examine test scores from previous years which were not influenced by teachers participating in formative assessment professional learning.

### **Terminology**

In the study of assessment certain terms are used which need clarification and explanation. The terms defined include *Assessment for Learning*, *Assessment of Learning*, *formative assessment*, *peer assessment*, *student self assessment*, and *summative assessment*.

Supporting research is provided to further clarify. Additionally, comparison charts for formative

assessment versus summative assessment and Assessment of Learning versus Assessment for Learning are included.

**Assessment for Learning.** Assessment for Learning is also known as *classroom assessment*, *formative assessment*, and *descriptive assessment*. The goal of Assessment for Learning is to provide initial feedback to the students, teacher, and other adults to result in changes in instruction, motivation, or behavior in order to impact student learning and growth. Assessment for Learning is often compared analogously to a doctor's check-up or coaches' team practice (Wiggins, 2007). Assessment for Learning involves teachers providing descriptive rather than evaluative feedback and students self-assessing and communicating their own results to others (Stiggins, Arter, Chappius, & Chappius, 2006). The focus of Assessment for Learning is the improvement of student achievement (Reeves, 2001) and the pupils learning instead of the teacher's teaching (Harris, 2007). Assessment for Learning also includes educative feedback designed to provide immediate, relevant, and useful information to the student (Reeves, 2001) and formative feedback (non-evaluative, supportive, timely, specific) designed to provide information communicated to the student to support modification of thinking or behavior to improve learning (Shute, 2008). Table 1.1 provides additional information for clarification of assessment for learning.

**Assessment of Learning.** Assessment of Learning is also known as *summative assessment* and *evaluative assessment*. The goal of Assessment of Learning is to provide a measure of student learning once instruction is complete. Traditionally, adjustments may be made to instruction for future groups of students using this data, but no changes are made to instruction for the group of students measured because the instruction is already complete. Assessment of Learning is often compared analogously to an autopsy or the team's game

(Wiggins, 2007). Assessment of Learning typically involves tasks developed by testing professionals, expressed in a quantitative score, given at the end of a set learning period, aimed to see how much a student has learned as a result of instruction, and results are often norm-referenced or criterion-referenced and used to hold teachers, learners, and schools accountable (McNamee & Chen, 2005). Table 1.1 provides additional information for clarification of assessment of learning.

Table 1.1

*Comparing Assessment for and of Learning: Overview of Key Differences*

	Assessment for Learning	Assessment of Learning
Reasons for Assessing	Promote increases in achievement to help students meet more standards, support ongoing student growth, improvement	Document individual or group achievement or mastery of standards, measure achievement status at a point in time for purposes of reporting, accountability
Audience	Students about themselves	Others about students
Focus of assessment	Specific achievement targets selected by teachers that enable students to build toward standards	Achievement standards for which schools, teachers, and students are held accountable
Place in Time	A process during learning	An event after learning
Primary Users	Students, teachers, parents	Policy makers, program planners, supervisors, teachers, students, parents
Typical Uses	Provide students with insight to improve achievement, help teachers	Certify student competence, sort students according to achievement, promotion and

	diagnose and respond to student needs, help parents see progress over time, help parents support learning	graduation decisions, grading
Teacher's Role	Transform standards into classroom targets, inform students of targets, build assessments, adjust instruction based on results, offer descriptive feedback to students, involve students in assessment	Administer the test carefully to ensure accuracy and comparability of results, use results to help students meet standards, interpret results for parents, build assessments for report card grading
Student's Role	Self- Assess and keep track of progress, contribute to setting goals, act on classroom assessment results to be able to do better next time	Study to meet standards, take the test, strive for the highest possible score, avoid failure
Primary Motivator	Belief that success in learning is achievable	Threat of punishment, promise of rewards
Examples	Using rubrics with students, student self- assessment, descriptive feedback to students	Achievement tests, final exams, placement tests, short cycle assessments

*Note.* From *Classroom Assessment for Student Learning* by R. Stiggins, J. Arter, J. Chappuis, and S. Chappuis, 2006, p. 33.

**Formative assessment.** Formative assessment is a process through which evidence of student learning is gathered and instruction is modified to increase student learning (Cauley &

McMillan, 2010). Formative assessment is used to “identify specific student misunderstandings [and] provide feedback to students to help them correct their errors” (Cauley & McMillan, 2010, p1). Formative assessment involves the process of teaching and learning (Pryor & Croussouard, 2008). Formative assessments should provide teachers and students feedback about student learning (Allen, Ort, and Schmidt, 2009). The results of formative assessment are used by students and teachers to adjust what is being done and improve learning (Colburn, 2008). Formative assessment is used by the teacher to diagnose where students are in the learning process, where gaps exist, and to help teachers and students improve learning (Perie, Marion, & Gong, 2009). Table 1.2 provides additional information clarifying formative assessment.

Table 1.2  
*Comparison of Formative and Summative Assessment*

	Formative Assessment	Summative Assessment
Purpose	To improve instruction and provide student feedback	To measure student competency
When administered	Ongoing throughout unit	End of unit or course
How students use results	To self- monitor understanding	To gauge their progress toward course or grade-level goals and benchmarks
How teachers use results	To check for understanding	For grades, promotion

*Note.* From *Checking for Understanding* by D. Fisher and N. Frey, 2007, p.4, Alexandria, VA: ASCD. Copyright 2007 by ASCD. Reprinted with permission

**Peer assessment.** Peer assessment includes feedback provided by peers to other students. It is often used for correction and student growth. Peer assessment is often guided by teacher feedback, rubrics, or checklists (Davies, 2000).

**Student self assessment.** Student self assessment includes feedback students complete themselves and use to correct their own learning. It is often guided by teacher feedback, rubrics, or checklists (Davies, 2000).

**Summative assessment.** Summative assessment is a record of current student achievement (Cauley & McMillan, 2010). Summative assessment uses tests to grade students or evaluate the effectiveness of curriculum (Pryor & Croussouard, 2008). Summative assessments are one time assessments administered at the end of the unit, semester, or year and are usually used as part of an accountability program, as part of a grading process, or to make instructional or policy decisions (Colburn, 2008, Perie, Marion, & Gong, 2009). Table 1.2 provides additional clarification of summative assessment.

Chapter One provided a general overview of the project including the statement of the problem, the research proposal, research questions, research hypotheses, background information, the significance of the research and definitions of important terminology. Chapter Two will provide a review in an effort to better understand the available information on formative and summative assessment, the student achievement gains evidenced in this research, and the disconnect between research-based best practices and teacher practices in school. The review of literature will provide an overview of the most relevant and recent research related to formative assessment and supports the need for this research study.

## Chapter Two: Review of Literature

### The Current Era

Academic Achievement and attempts to measure that achievement have been focused on with increasing popularity over the last century. In his book, *The Best Schools*, Thomas Armstrong provided an argument as part of the national discourse on academic achievement. Armstrong (2006) provided a timeline in the development of the national fixation with achievement and accountability over the last 100 years. Armstrong's (2006) timeline spans from the 1893 Committee of Ten Report to the passage of the No Child Left Behind Act in 2001. Additionally, since its passage in 2001 the spirit behind No Child Left Behind seems to be building momentum with increasing focus on national standards, international benchmarking, and more accountability (Zhao, 2009). In the current era of accountability, schools have increasingly focused on standardized test scores as a measure of student achievement and academic growth. With the advent of No Child Left Behind and the ensuing focus on Adequate Yearly Progress measured by standardized test scores all educators have become aware of the need to have students succeed on these assessments.

In this entire assessment milieu, the difference between evaluation and assessment has not been addressed. No Child Left Behind and Adequate Yearly Progress focus on evaluation of student achievement. Evaluation is a good thing. It lets educators, and others, decide if students have learned what they need to learn. However, if educators are not careful, they can focus too much on evaluation to the neglect of assessment. Summative large-scale assessments provide information for evaluation and accountability, but teachers are not concerned only with final results because their primary concern is with the process of helping students to master standards



(Guskey, 2005). Assessment allows educators to examine student learning and to gain information that affects instruction and helps students learn even more (Davies, 2000). The key to increased student achievement is to use evaluation to determine if the student has learned and to also use assessment in the classroom before the evaluation in order to make adjustments in instruction and to provide feedback so that the student can learn more.

In their recent article, Stiggins and Chappuis (2005) reviewed the research supporting gains and closing of achievement gaps which are possible if classroom assessments “focus on clear purposes, provide accurate reflection of achievement, provide students with continuous access to descriptive feedback on improvement in their work (versus infrequent judgmental feedback) and, bring students into the classroom assessment processes”(p 5). Stiggins and Chappuis summarized research findings over the last several years that show documentation of success of quality formative assessment. The strategies and tools teachers need to use to implement effective formative assessment have been identified by researchers.

Bloom (1984) found significant differences in achievement for students in classes that relied on classroom assessment to support learning. Jerald (2001) found improvement in traditionally low-performing schools by increasing the use of day-to-day classroom assessment. In 2003 Meisels, Atkins-Burke, Xue, and Bickel showed increased achievement for students involved in work-sample-based performance assessment. Black and Wiliam (1998) reviewed 250 articles and found that improving student involvement in formative assessment helps all students but shows the most gains for low achieving students. The research indicated repeatedly that effective formative assessment practices in the classroom directly impact and improve student learning as measured by classroom performance data and on standardized assessments. Nugent (2009) found a strong correlation between student performance on formative

benchmarking assessments and a content-based criterion-referenced standardized assessment. Williams (2009) found the Tennessee Formative Assessment Program (TFAP) was a predictor of student test scores on the Tennessee Comprehensive Assessment Program with the majority of teachers in the study using TFAP information to adjust classroom instruction. Analysis of the literature supports the use of quality formative assessment as having a significant impact on increased student achievement. Formative assessment is a means to gather evidence of student learning and using it in a way to maximize student learning (Stiggins, Arter, Chappius, & Chappius, 2006).

In the state of Georgia with the advent of House Bill 1187 and then in the nation as a whole with the No Child Left Behind Act, politicians brought the use of standardized test scores as a means of measuring student progress to the forefront of the education community. Stiggins (2002) set out a timeline from the early 1960s in which society as a whole placed great emphasis on the results of standardized test scores. The trend to use these evaluation results has continued and intensified, and it continues to do so at the present time. The use of these evaluative tools is not negative in and of itself; standardized testing as a measure of student achievement is perfectly appropriate. The danger lies in the failure of politicians and society to understand that standardized test scores do not paint a total picture of student achievement. The even greater danger lies in the fact that we are creating an entire generation of teachers for which standardized test scores are the most important measure. This researcher believes this theory causes the present generation of educators to conclude erroneously that standardized measures are the only measure of student achievement. Stiggins argued the need to balance assessment is vital. Stiggins' 2002 article illustrated this concept beautifully by comparing Assessment of Learning to Assessment for Learning. He promoted the need for a balanced assessment that uses both

standardized test scores and classroom assessment to promote school improvement. Educators must use Assessment of Learning (the standardized test) and Assessment for Learning (classroom assessment) to assess student achievement accurately and to affect student learning the most (Stiggins,2002). Stiggins and Chappius (2005) summarized the effects of formative versus summative assessment when they stated,

Feedback delivered once a year from a standardized district, state, national, or international assessment is far too infrequent and broadly focused to be helpful.

The evidence must come to students moment to moment through on-going classroom assessment. This places the teacher at the heart of the relation between assessment and school effectiveness. (p 1)

### **Defining Assessment**

Assessment is an attempt to determine what students know and what they still need to learn. If this understanding relies solely on tests and quizzes, a full picture is not provided. In order to develop a more complete picture of a student's knowledge, the educator must rely on more measures such as summarizing, diagramming, comparing and contrasting, and demonstrating (Strong, Silver, & Perini, 2001). This type of assessment takes time and training. However, in a curriculum focused on standards and adjustments to curriculum, more time is available because the less teachers teach then the more teachers coach (assess for transfer) , therefore, teachers are freed up to cause better results by adjusting learning. The coaching process is grounded in more frequent assessment, not more teaching (Wiggins, 2007). Standards effect classroom assessment by comparing student work to the standard as opposed to an average or other student's work, requiring students to demonstrate proficiency, clearly communicating what is expected of students, focusing on the improvement of student learning not merely the

results of an evaluation or a score (Reeves, 2001). Wiggins (1993) provided principles of assessment for better learning:

- The interests of the student shall be paramount. Assessment shall be planned and implemented in ways which maximize benefits for students, while minimizing negative effects on them.
- The primary purpose of assessment shall be to provide information which can be used to identify strengths and to guide improvement. In other words, it should suggest actions which may be taken to improve the educational development of students and the quality of educational programs.
- Assessment information should not be used for judgmental or political purposes if such would be likely to cause harm to students or to the effectiveness of teachers or schools.
- Every effort should be made to ensure that assessment and evaluation procedures are fair to all.
- Community involvement is essential to the credibility and impact of assessment and evaluation processes. All parties with a direct interest should have an opportunity to contribute fully. Self-assessment is the appropriate starting point.
- Careful consideration should be given to the motivational effects of assessment and evaluation practices.
- In the assessment of intellectual outcomes, substantial attention should be devoted to more sophisticated skills such as understanding of principles, applying skills and knowledge to new tasks, and investigating, analyzing, and discussing complex issues and problems.

- Emphasis should be given to identifying and reporting educational progress and growth, rather than comparisons of individuals or schools.
- The choices made in reporting assessment information largely determine the benefit or harm resulting from the information. For this reason, the selection, presentation, and distribution of information must be controlled by the principles outlined previously. (pp 26-27)

Once educators understand the need to balance the assessment used in the classroom, they must understand how to effectively assess in the classroom. Assessment for Learning requires teachers to re-examine their methods of assessment. Assessment for Learning is based on the major premise that two types of assessment can occur in the classroom. Evaluative assessment (such as the standardized test) provides students with information on how well they have grasped a certain concept and on their mastery of this concept in relation to a standard or compared to other students. Descriptive or formative assessment provides the student information on areas of strength and weakness, and it provides the student with opportunity to learn and to correct errors in order to improve (Davies, 2000). In order to promote student achievement and academic growth, the most effective educator provides both types of assessment in the classroom. Educators should understand the difference between the format of evaluative assessment and formative assessment. The state test (evaluative, summative) is a quick audit of how students are performing against standards; there is a cause and effect relationship (Wiggins, 2007). Wiggins explained this difference eloquently with this allegory: “We are confusing the yearly doctor’s physical exam with the day to day ‘test’ of ‘being healthy’. You don’t get healthier by practicing the physical. You pass the physical by doing healthy tasks year-round (p.2)”. McTighe and O’Conner (2005) identified seven assessment

practices teachers can implement to enhance effective student learning including using summative assessment to frame meaningful performance goals, showing criteria and models in advance, assessing before teaching, offering appropriate choices, providing feedback early and often, encouraging self- assessment and goal setting, and allowing new evidence of achievement to replace old evidence. Stiggins, Arter, Chappius, and Chappius (2006) outlined best practices in assessment as understanding informational needs of the student and designing the assessment to meet those needs, having a clear sense of achievement expectations, translating learning targets into student-friendly language, insuring accuracy by selecting proper assessment method for each context. Other strategies suggested by Stiggins, Arter, Chappius, and Chappius (2006) included developing the method and using it well, being aware of and working to counteract things that can go wrong and lead to inaccuracy of assessment, and communicating results of the assessment to the users (students, parents, teachers).

See Table 2.1 for additional indicators of sound classroom assessment practices.

Table 2.1  
Indicators of Sound Classroom Assessment Practice

<p>Why Assess? Assessment Processes and Results Serve Clear and Appropriate Purposes</p>	<ul style="list-style-type: none"> <li>• Teachers understand who the users and uses of classroom assessment information are and know their information needs.</li> <li>• Teachers understand the relationship between assessment and student motivation and craft assessment experiences to maximize motivation.</li> <li>• Teachers use classroom assessment processes and results formatively (Assessment for Learning).</li> <li>• Teachers use classroom assessment results summatively (Assessment of Learning) to inform someone beyond the classroom</li> </ul>
--	--

	<p>about students' achievement as of a particular point in time.</p> <ul style="list-style-type: none"> <li>• Teachers have a comprehensive plan over time for integrating assessment for and of learning in the classroom.</li> </ul>
<p>Assess What? Assessments Reflect Clear and Valued Student Learning Targets</p>	<ul style="list-style-type: none"> <li>• Teachers have clear learning targets for students; they know how to turn broad statements on content standards into classroom-level targets.</li> <li>• Teachers understand the various types of learning targets they hold for students.</li> <li>• Teachers select learning targets focused on the most important things student need to know and be able to do.</li> <li>• Teachers have a comprehensive plan over time for assessing learning targets.</li> </ul>
<p>Assess How? Learning Targets Are Translated Into Assessments That Yield Accurate Results</p>	<ul style="list-style-type: none"> <li>• Teachers understand what the various assessment methods are.</li> <li>• Teachers chose assessment methods that match intended learning targets.</li> <li>• Teachers design assessments that serve intended purposes.</li> <li>• Teachers sample learning appropriately in their assessments.</li> <li>• Teachers write assessment questions of all types well.</li> <li>• Teachers avoid sources of bias that distort results.</li> </ul>
<p>Communicate How? Assessment Results Are</p>	<ul style="list-style-type: none"> <li>• Teachers record assessment information accurately, keep it confidential, and appropriately combine and summarize it for reporting (including grades). Such summary accurately reflects current level of student learning.</li> </ul>

<p>Managed Well and Communicated Effectively</p>	<ul style="list-style-type: none"> <li>• Teachers select the best reporting option (grades, narratives, portfolios, conferences) for each context (learning targets and users).</li> <li>• Teachers interpret and use standardized test results correctly.</li> <li>• Teachers effectively communicate assessment results to students.</li> <li>• Teachers effectively communicate assessment results to a variety of audiences outside the classroom including parents, colleagues, and other stakeholders.</li> </ul>
<p>Involve Students How? Students Are Involved in Their Own Assessment</p>	<ul style="list-style-type: none"> <li>• Teachers make learning targets clear to students.</li> <li>• Teachers involve students in assessing, tracking, and setting goals for their own learning.</li> <li>• Teachers involve students in communicating about their own learning.</li> </ul>

*Note.* From *Classroom Assessment for Student Learning: Doing It Right- Using It Well* by R. J. Stiggins, J. Arter, J. Chappius and S. Chappius, 2006 p.27

### **Assessment in Practice**

Formative or descriptive feedback focuses on several different theories about student learning. The first theory is that students must be involved in a continuous assessment cycle (Davies, 2000). Students are actively involved in assessing their own and each others' work. This assessment leads to improvement and then to further assessment, and the cycle continues. By being directly involved in the assessment process, students are able themselves to make important decisions about their learning. Student involvement allows students to decide if they are capable. Effective assessment practices in the classroom keep students believing in themselves (Stiggins, Arter, Chappius, and Chappius, 2006). Part of this self and peer



assessment focuses on the need for a standard to which the students can compare their work. Students need to see realistic samples that focus on what they need to do but are not too far from where they are currently in their ability. Exemplars of student work at differing ability levels should be posted throughout the classroom and reviewed with students. This allows students to see how to get from where they are to where they need to be. Part of this continuous assessment cycle is the opportunity for students to practice and to learn from their errors and successes. Students are able, through practice, to take the learning to a deeper level (Davies, 2000). As students practice, they further self assess. This self assessment process decreases the evaluative feedback while increasing the descriptive feedback. Black and Wiliam (1998) found that through self assessment all students show gains, but typically those with the least gains prior to self assessment show the largest gains with self assessment. Once the practice time has been allowed and students are comfortable with their work, they need time to present their work to others. By sharing their work with others, all students have the opportunity to learn from each other (Davies, 2000).

Teachers wanting to incorporate more descriptive, formative assessment into the classroom need to understand descriptive feedback. According to Davies descriptive feedback comes during as well as after learning, is easily understood and related directly to learning, is specific so performance can improve, involves choice on the part of the learner as to the type of feedback and how to receive it, is part of an ongoing conversation about the learning, is in comparison to models, exemplars, samples or descriptions, and is about the performance of the work, not the person (Davis, 2000, p13).

Descriptive feedback can be provided by the teacher, other adults in the school and at home, other students, or the student themselves through self assessment. Feedback relates

descriptions and suggestions to student work and provides information to teachers and students about the student's work (Brookhart, 2008). Shute's 2008 meta-analysis of research on formative feedback provided guidelines for providing feedback including focusing on the task instead of the learner, providing elaborated feedback to enhance learning, presenting feedback in manageable units, being specific and clear, keeping information as simple as possible, providing information to clarify the difference between the performance and the goal, providing unbiased, objective feedback in writing, and using feedback to set the goal as learning rather than performance.

According to Brookhart (2008), feedback strategies vary in timing, amount, mode, audience, focus, comparison, function, valence, clarity, specificity, and tone. Variations in timing relate to when feedback is given and how often and good feedback is provided immediately for knowledge and facts, delayed for more comprehensive reviews, provided in time to make a difference for the student, and provided as often as possible for all major assignments (Brookhart, 2008). How many points are made and how much information is provided about each point are variations in amount. Good feedback addresses the most important points, addresses points that relate to major learning goals, and considering the student's developmental level (Brookhart, 2008). Variations in mode include differences between written, visual, and oral feedback. Good feedback occurs when teacher are selecting the best mode for the message, being interactive when possible, using written feedback on written work, and using demonstration when students need an example (Brookhart, 2008). Audience variations are relate to providing feedback to the individual or the group with good feedback recognizing that individual feedback sends a message of valuing the student's work and group feedback is appropriate when most of the class needs support with the same concept (Brookhart, 2008).

Comparison variations can be accomplished by comparing to the criteria, other students, or the student's past performance. Good feedback is criterion based to provide feedback on the work, norm referenced when providing feedback on process, and self-referenced when student's need to see progress (Brookhart, 2008). Variations in function of feedback are either descriptive or evaluative with good feedback describing instead of judging (Brookhart, 2008). Feedback should be clear using vocabulary and concepts students understand and tailoring amount and content to the student's developmental level (Brookhart, 2008). Variations in specificity relate to feedback being nitpicky, appropriate, or overly general with good feedback tailored to the student and the task, being specific enough for students to know what to do but not doing it for them, and identifying errors without correcting every one (Brookhart, 2008).

Additionally, type of feedback influences students motivation to learn, with feedback that is descriptive of the work, criterion-based, emphasizes that the learning is what is important not looking good and points out the strengths in the work as well as addressing areas in need of improvement (Stiggins, Arter, Chappius, and Chappius, 2006). Formative feedback can identify help students identify the gap between the current work and the desired level of performance (Shute, 2008). Additionally, providing specific feedback which highlights the actions needed for improvement can result in improved performance (Shute, 2008).

The student must have a standard to relate the work to in order for descriptive feedback to be meaningful and for an opportunity for growth toward the exemplar or standard. Cauley and McMillan (2010) suggested using the standard to provide a clear learning target, providing feedback about progress toward meeting the target, attributing student results to effort, and encouraging student self-assessment. In Georgia, the Department of Education has developed content standards for the state. These standards form the basis for student learning at each grade

level. Currently, the standards are presented in four distinct parts that support one another. The standard itself presents the major overlying concept with elements attached to that standard to further define the standard for each grade level. Tasks for each standard allow teachers to have sample activities that could be used to introduce the standards. Student work samples with teacher commentary provide both teachers and students with an exemplar of work. By providing the four components, the state itself is providing a valuable tool for teachers to use in providing descriptive feedback to students. The standards and elements provide an outline for teachers and students as they measure progress toward their goal. The student work samples provide a document for teachers and students to use in order to compare individual work to the standard and to assess individual progress toward the goal.

It is this researcher's opinion that the standards developed in Georgia have great promise for use in the classroom. The standards provide teachers and students with both a guide and a final destination. Students and teachers that know the final destination are more apt to be successful (Davies, 2000). Samples can be used to develop criteria with students, to assess student work, and to help others understand learning (Gregory, Cameron, & Davies, 1997). The teacher commentary attached to the samples can provide educators with a sample of descriptive feedback. These standards and the work samples that demonstrate mastery should be publicly displayed in as many ways as possible through the use of walls, halls, and newsletters (Carr & Harris, 2001).

Educators must also collect evidence of student learning as they move through this assessment process. Lincoln and Guba (1984) described a process of triangulation as a way to collect a balanced representation of student learning. Triangulation includes three types of evidence collecting: observation of process, collection of products, and conversations. Teachers

need to collect evidence in all of these areas. Teachers can observe the process by examining student work or by observing as the student completes the work. A collection of products could be completed by the teacher or the student. By allowing the student to develop the collection of products, the educator encourages further self assessment as the student reexamines the work to determine whether or not to include it in the collection. Conversations with students could be ongoing, and are a further way to collect evidence of student learning. All of these opportunities should encourage the educator to focus on multiple ways of evaluation and provide opportunities for students of differing abilities to demonstrate mastery.

Strong, Silver, and Perini (2001) shared the concept of graduated difficulty as a way to assess students and to allow for diversity in the classroom. Their graduated difficulty strategy focuses on four goals that the student and teacher examine together. The focus of each assessment should allow students to have flexibility in selecting the difficulty level, completing the task (with teacher support as needed), evaluating the performance (allowing for self assessment), and goal setting (planning the next move) (Strong, Silver, & Perini, 2001).

In addition to addressing diversity needs in the classroom, assessment must also support progress toward the standards. It should be ongoing and relevant to the learning goals in the classroom. The assessment needs to be comprehensive, inclusive, and technically sound (Carr & Harris, 2001). Comprehensive assessment is relative to the whole purpose of the classroom, diverse in nature, and allows for student strengths and weaknesses. Inclusive assessment is developmentally and culturally appropriate, addresses learning styles and multiple intelligences, and involves self-assessment. Technically sound assessment is continuous, valid, and reliable (Carr & Harris, 2001).

Multiple forms of assessment should be used in the classroom as part of Assessment for Learning in order to gain an accurate picture of student learning. Tools like performance checklists, scales, tests, quizzes, and student work samples should be used continually throughout the learning process (Carr & Harris, 2001). Before determining the use of an assessment to evaluate students, the educator should determine if the assessment provides feedback in relation to the standard, is sufficient to gather information and to document the standard, and is of high quality (Carr & Harris, 2001).

High quality formative assessment answers questions about where students are going, where student are now, and how students can close the gap (Stiggins, Arter, Chappius, & Chappius, 2006). Lewis, Berghoff, and Pheeney (1999) expressed questions when related to a specific assessment as what will be on it, what the teacher wants and what the students have to do. Stiggins, Arter, Chappius, and Chappius provide seven strategies to answer these questions:

Strategy 1- Provide a Clear and Understandable Vision of the Learning Target.

Share learning targets in advance of teaching, giving assignments, or completing activities. Convert targets to student-friendly language. Provide scoring guides for students to evaluate their own work.

Strategy 2 – Use Examples and Models of Strong and Weak Work. Share anonymous students work, work for outside sources, or teacher’s work. Have students analyze and discuss samples. Model beginning, correcting, and revising work.

Strategy 3- Offer Regular Descriptive Feedback. Provide feedback on strengths and weaknesses in relationship to the set criteria. Narrow comments to address specific areas for improvement.

Strategy 4- Teach Students to Self-Assess and Set Goals. Provide students guidance and time to identify their own strengths and weaknesses. Allow students to offer descriptive feedback to others. Use feedback to identify what needs to be worked on and develop future goals.

Strategy 5- Design Lessons to Focus on One Aspect of Quality at a Time. Build competence one block at a time. Introduce components of larger skills sets individually.

Strategy 6 – Teach Students Focused Revision. Model for students how the teacher would revise and provide them with an example to revise. Allow students to peer assess.

Strategy 7 – Engage Students in Self- Reflection, and Let Them Keep Track of and Share Their Learning. Engage students in tracking, reflecting, and communicating on their work. Provide prompts for reflection on student work. (pp 42-45)

To answer these questions Chappius and Stiggins (2002) also suggested that “students need clearly articulated, concise learning targets [...] can practice comparing their works to models of high-quality work [...] and know what to do to move from their current position to the final goal” (p 43). Lewis, Berghoff, and Pheeny (1999) suggested strategies including sharing test specifications, marking expectations to set the standard expected by brainstorming criteria with students, defining criteria to develop the grading rubric, and generally making public what is to be judged. Harris (2007) noted that providing specific ongoing feedback should help students see the gap between their current knowledge and the expected knowledge and abilities and also help them to identify actions needed to achieve the expectations. Shute (2008) suggested two major actions needed to provide relevant feedback as verification and elaboration. Verification is the act of visually acknowledging, through highlighting, a checkmark, or some other mark, that the information is correct (Shute, 2008). Elaboration is the act of

acknowledging the correct answer, explaining why a response is wrong and providing information on what the correct answer should be and why (Shute, 2008).

Once evidence has been collected, the evaluation and reporting of student work can occur. This process is very different from the traditional averaging of grades in order to determine a score to record on a report card. The point of evaluating and reporting is not to obtain a grade (although that is a side effect). Evaluating and reporting should answer four major questions for all stakeholders involved—the student, the teacher, and the parent letting them know what the student knows and is able to do, identifying what requires further attention, exploring what ways the student’s learning can be supported, and showing how the student is doing in relation to the standard (Davies, 2000).

In addition to providing individual student feedback, assessment results should provide curricular feedback for the teacher. Fisher and Frey (2007) noted,

As Tomlinson (1999) so aptly stated, “Assessment always has more to do with helping students grow than with cataloging their mistakes” (p 11). We couldn’t agree more. Tests and assessments can and should be used to check for understanding with the goal of increasingly precise instruction for individual students. Although we acknowledge that tests and assessments will be used for other purposes-report cards, grading, and public accountability to name a few - it is crucial that we also use the information we gather through testing to plan our instruction. (p 119)

In his keynote address to curriculum directors at the 2007 Georgia Association for Curriculum and Instructional Supervisors Fall Conference, Grant Wiggins explained the role of formative assessment in curricular decision making. Wiggins (2007) compared teachers to



coaches and supported educators making curricular adjustments based on the strengths and weaknesses of the students in each class in comparison to the set standard. Teachers cannot plan in a vacuum and adhere to a strict pacing guide. Educators must plan to adjust because initial planning will never adequately predict the complex reality that exists in every classroom (Wiggins, 2007). The flexibility and responsibility to assess student mastery and adjust instruction based on the results of the assessment is imperative to support student growth. Time must be allowed for educators to make adjustments based on the results of formative assessment in their classrooms. This is not time for teachers to “wing it,” but is time for intelligently planned instruction based on adjustments needed as identified by assessment results (Wiggins, 2007).

Finally, assessment should build for transfer. Transferability is based on three cornerstones which impact the content’s relationship to authentic real world tasks. In order to build for transfer, educators must address through their curriculum what it means to “do” the subject (to have abilities “tested” in the real world), what authentic options, constraints, and opportunities exist when doing such work, and what the key transfer tasks at the heart of each subject are (Wiggins, 2007). Transferability is related to Blooms’ taxonomies’ application level in that application is different from simple comprehension. The student is asked to think in new situations not reply with specific knowledge. The assessment must involve situations new to students in order to assess transfer and determine what a student has learned to apply in a practical way (Wiggins, 2007). The importance of transfer is its relationship to the core priority of all learning which is to serve us in the future. Bruner, in 1960, stated

The first object of any act of learning, over and beyond the pleasure it may give, is that it should serve us in the future... In essence, it consists in learning initially

not a skill but a general idea which can then be used as a basis for recognizing subsequent problems... This type of transfer is at the heart of educational process- the continual broadening and deepening of knowledge in terms of ...ideas. (p 17)

Transfer is found when assessment focuses on fewer quizzes (recall) and more performance tasks (application). Assessment should be focused not on content alone, but on thoughtful and effective use of the content where transfer and personal meaning making are required, and not on knowledge and skills, but on important accomplishments requiring big ideas, knowledge, and skills (Wiggins, 2007).

As authenticity, complexity, performance effectiveness, and autonomy increase in student tasks and as prompts, cues, and scaffolds decrease, students move toward more autonomous transfer (Wiggins, 2007). Formative assessment allows students to participate and receive feedback on authentic performance assessments which build for transfer into real life situations. Not only does formative assessment address what education should be about to begin with, which is preparing students for success in the future, it also results in better performance on summative, standardized assessments. Standardized assessments mostly require transfer, not rote practice, because the items are unknown, they change from year to year, and only if the student understands the concept can they cope with novelty (Wiggins, 2007).

### **Assessment for Student Success**

By examining classroom assessment, today's educator can provide a system in which students can find greater success. What educators have done is not working; It is hurting, not helping (Biddle and Berliner, 1998). Davies summarized this best when she writes, "making classroom assessment work means reframing the conversation from one about ranking and

sorting students to one about assessing learning in the context of our students' futures" (Davies, 2000, p 77).

In her 2008 study, Sarah McMannus used a theory approach to determine emerging themes when teachers implemented formative assessment and new behaviors of students were observed as a result of the use of formative assessment. McMannus found that teachers' views were more inclusive, and students were viewed as partners in the process. Additionally, students' self-efficacy, commitment to learning, and engagement increased (McMannus). Stiggins' (2007) work noted the historical role of assessment as highlighting differences and ranking students to produce winners and losers. Stiggins supports Assessment for Learning as a process that "turns day-to-day assessment into a teaching and learning process that enhances (instead of merely monitoring) student learning" (p 22). Dylan Wilam (2006) noted that as long ago as 1969 Benjamin Bloom noted a distinction between evaluation (summative assessment) and the use of formative assessment to provide feedback and correction. In her 2002 study Patricia Bucci, even developed a new term "assesslearnment" based on her case study of the connection between formative assessment practices in the classrooms of elementary teachers and the learning and growth of students. Bucci found a connection between the beliefs exhibited by outstanding educators about the positive benefits of formative assessment and increased student learning.

Educators are at a unique time in the history of education in the nation and in the state of Georgia. Now, more than at any time in the past, society, government, and communities are focused on increasing student achievement and improving schools. State legislation and societal agencies continue to focus on improving education. The current Georgia State School Superintendent, Kathy Cox, promises that Georgia will lead the nation in student achievement.

The current Governor of Georgia, Sonny Purdue, supports the Superintendent's charge. The national media and the government, through No Child Left Behind, are providing greater focus and challenge as they argue all students should be proficient in reading and math by 2014. The emphasis on assessment and student evaluation is increasingly great. This focus and attention has created an urgency in the field of education that is unique to this time. Teachers, administrators, and researchers involved directly in the educational setting have reacted to this urgency with a focus on best practices and effective techniques and strategies that is countered by the need to have quick and significant gains on standardized assessment measures. However, this researcher believes politicians and government have missed the mark and have failed to realize the potential of classroom assessment as a means of increasing student achievement. Willis (1993) examined the inconsistency between policy statements and theory that is applied in most classrooms when he argues, "The rhetoric of curriculum reform with its references to the development of understanding, and lifelong learning is meaningless. Those objectives are unlikely to be achieved unless the accompanying assessment reflects the same theoretical principals" (p 384).

If states, systems, and schools continue to focus on standardized assessments as the diviner of evidence of learning, without promoting the belief that quality formative classroom assessment best impacts learning and growth of students, then educators will have missed the opportunity to most impact student achievement. Wiggins (2007) supported standardized assessment and recognized its purpose in education. Standardized assessment is a corollary instrument meant to provide an indicator of student success. Designers of standardized assessments understand that the items used to gauge student understanding should have a high correlation to student understanding of the content and at best are proxy in nature (Wiggins).

Educators in the classroom will come to understand the significant difference between standardized and formative assessments. Standardized assessments are a measure of student understanding as a whole and should be used as a guide to understanding the performance of groups overall. Formative assessments provide information of a specific nature for individual students and can be used to provide feedback and instruction to individuals or to adjust instruction and curriculum as needed. “The distinction between formative and summative purposes of assessment should be maintained, while assessment systems should be planned and implemented to enable evidence of students’ ongoing learning to be used for both purposes” (Harlan, 2005, p207). Key to student achievement gains is the understanding that quality formative assessment leads to increased student learning which will lead to increased performance on summative measures.

This researcher believes what educators have done in the past is not sufficient and that the movement in Georgia toward a standards-based education system, with the advent of the Georgia Performance Standards, is a step in the right direction. If the system, school, and classroom reenergize their classroom assessment with a focus on Assessment for Learning, they will show a great increase in student achievement. Grant Wiggins (2007) summarized the need to understand and use formative assessment in the classroom when he said, “It is not the teaching that causes the results; it is the adjustments that cause the results.” Teachers must have a plan to follow, but it is believed if part of this plan does not include formative assessment, and adjustment of instruction because of the results of this assessment, then teachers are merely covering content not teaching standards. Stiggins, Arter, Chappius, and Chappius (2006) stated,

Used with skill, assessment can motivate the unmotivated, restore the desire to learn, and encourage students to keep learning, and it can actually create – not

simple measure – increased achievement. None of this happens if assessment functions solely as an accountability measure, as it does in the case of standardized testing and in determining grades. Because we now understand that assessment can work in positive ways to benefit learning, the time is right to add to our definition of good teaching the skillful use of assessment – doing it right and using it well. (p 3)

### **Need for the Study**

After a careful review of the research including examination of the current educational climate, definitions of assessments, research that supports best practices in assessment, and recommendations for further investigation, this researcher believed there was room for more research concerning assessment. The current political and societal focus on student achievement and resulting focus on standardized summative assessment can overshadow the need for quality formative assessment. Rodriguez and Bellanca (2007) provided a compelling summary of the need to develop authentic formative assessments when stating,

When assessment goes beyond the limits of the #2 pencil standardized test and examines authentic learning, multiple views emerge regarding what a student knows and is able to do... By forming standards and criteria for success and by using new tools to challenge the multiple intelligences in the classroom, the teacher can access the knowledge and the student's ability to use the knowledge in meaningful ways. (p 23)

Contributing to the research base on formative assessments adds to the information available on ways to increase teacher reliance on assessments that impact student learning. Rodriguez and Bellanca (2007) argued that professional development is key to providing teachers with the

resources needed to implement quality assessment strategies that impact student learning in their classrooms. Understanding the relationship between professional development and teacher perception of formative assessment as well as the correlation to student achievement data will shed further light on teachers' understanding of the relationship between professional development, formative assessment, teacher perception and student learning. Very few recent studies were found related to the relationship between professional development and Assessment for Learning or formative assessment. A study conducted by Nash in 2008 was found to investigate the relationship between teacher perception of formative assessment and the teachers' use of goal setting and descriptive feedback in the classroom. In a study of 730 teachers in the state of Kansas, Nash (2008) found that "teachers with a more positive perception regarding efficacy of the formative assessments also scored higher on the variable that measured teachers' use of goals and feedback in the formative process" (p 49). A study by Gilson in 2009 investigated the relationship between a professional learning community and how the type of professional development provided can increase the effectiveness of Assessment for Learning (formative assessment) use in the classroom. Gilson (2009) concluded that teachers benefited from professional development and felt positively about the benefits of using the assessment techniques in their classrooms. This researcher found the study of formative assessment compelling, and found focusing on the relationship between professional development, teacher perception of formative assessment, and student achievement as an opportunity to further the literature in this area.

In Chapter Two, a review of the research provided evidence of the existence of decades of data supporting the assertion that formative assessment supports increased student learning. Additionally, Chapter Two highlighted research which pointed to the lack of implementation on

best practices in assessment in many classrooms and the link between professional development and best practice implementation. Chapter Three will provide details on the research design, questions, hypotheses, and population examined in this study and outline the instrumentation and procedures used.



### **Chapter Three: Methodology**

Chapter Three includes the methodology used to complete the study. It includes descriptions of the research design, a statement of the research questions, a summary of the hypotheses to be examined, characteristics of the population to be studied, descriptions of the instruments used to gather data, and details of the procedures to be carried out.

#### **Research Design**

First, the research within this study was designed to examine, in the specific population, the relationship between participation in targeted professional learning and student achievement gains on standardized Georgia End of Course Tests. Second, this research was designed to examine the relationship between participation in targeted professional learning and teacher perception of formative assessment techniques. Finally, this research was designed to examine the relationship between teacher perception of formative assessment techniques and student achievement gains on standardized Georgia End of Course Tests. There are several questions and hypotheses related to this research design.

#### **Research Questions**

As a part of the research design of this study, questions were asked and data gathered and analyzed to further examine these questions.

- RQ<sub>1</sub>.** Is there a relationship between teacher participation in Assessment for Learning professional development and student achievement as measured by the Georgia End of Course Tests standardized assessment?
- RQ<sub>2</sub>.** Is there a relationship between teacher participation in Assessment for Learning professional development and teacher perception of the benefits

of Assessment for Learning strategies?

- RQ<sub>3</sub>.** Does a relationship exist between teacher perception of the benefit of formative assessment and student achievement as measured by the Georgia End of Course Tests standardized assessment?

### **Research Hypothesis**

Based on a review of the literature hypotheses were developed related to the research questions.

#### **Null Hypotheses RQ<sub>1</sub>.**

*H<sub>01</sub>*. There will be no significant difference between teachers who participated in Assessment for Learning professional development those that did not as shown by student performance on the Georgia End of Course Tests standardized assessment.

#### **Null Hypothesis RQ<sub>2</sub>.**

*H<sub>02</sub>*. There will be no significant difference between teachers who participated in Assessment for Learning professional development as measured by teacher perception of the benefits of Assessment for Learning strategies.

#### **Null Hypotheses RQ<sub>3</sub>.**

*H<sub>03</sub>*. No significant relationship exists between teacher positive perception of the use of formative assessment and student achievement as measured by student performance on the Georgia End of Course Tests standardized assessment.

### **Population**

The population for the study was ninth, tenth, eleventh, and twelfth grades students in rural Georgia public high school. Ninth, tenth, eleventh, and twelfth grades teachers and students in core content areas participated in the study. For the purposes of this study, three high

schools in rural Georgia were identified to participate. The schools were similar demographically and all three were located in the rural west Georgia area. All of the schools in the study were accredited by the Southern Association of Colleges and Schools and the Georgia Accrediting Commission. All schools served similar size and demographic populations. For the purposes of anonymity each school will be referred to as Application-Based High School, Theory-Based High School, and Control High School.

Teachers at Application-Based High School participated in application-based professional development on formative assessment. Teachers at Theory-Based High School participated in theory-based professional development on formative assessment. Teachers at Control High School (the control group) did not participate in targeted professional development. All teachers within the school provided survey data.

Teachers within the school who provided instruction during the 2008-2009 year in core content areas of 9<sup>th</sup> Grade Literature, American Literature, Physical Science, Biology, United States History, and Economics participated in the study to provide classroom Georgia End of Course Test data for that year alone. Teachers within the school who provided instruction during the 2007-2008 school year and provided instruction in the same content area for the 2008-2009 school year in core content areas of 9<sup>th</sup> Grade Literature, American Literature, Physical Science, Biology, US History, and Economics participated in the study to provide classroom Georgia End of Course Test data for both years to compare gains.

The student and teacher sample at Application-Based High School, Theory-Based High School, and Control High School consist of minimally diverse demographic groups. Table 3.1 provides summary data for student demographic information for the population of all three

schools. Table 3.2 provides summary data for teacher demographic information for the population of all three schools.

Table 3.1

*Demographic Data for Student Study Population*

School	Application-Based High School	Theory-Based High School	Control High School
N of Students	632	556	1,599
% Asian	0	1	0
% black	11	55	23
% Hispanic	2	1	2
% multiracial	3	1	2
% white	85	42	73
% alternative program	1	4	3
% economically disadvantaged	53	56	30
% gifted	8	0	11
% remedial	4	0.5	1
% students with disabilities	6	22	6
% vocational labs	64	72	58
N EOCT completed in 2008-2009	818	831	2,246

The student sample at Application-Based High School consists of 632 9<sup>th</sup>-12<sup>th</sup> grade students. The racial demographics of the students were 11% black, 2% Hispanic, 3% multiracial, and 85% white. Additionally, the sample student population consists of 1% alternative program, 53% economically disadvantaged, 8% gifted, 4% remedial, 6% students with disabilities, and 64% vocational labs. Students at Application-Based High School completed 818 Georgia End of Course Tests, with some students testing in multiple subjects, during the 2008-2009 school year.

Table 3.2

*Demographic Data for Teacher Study Population*

	Application-Based High School	Theory-Based High School	Control High School
N Teachers	41	41	92
N Female Teachers	18	19	61
N Male Teachers	23	22	31
% black	5	17	10
% Hispanic	5	2	0
% white	90	80	90
Average Years Experience	12.41	8.68	12.74
% Holding Bachelors or Masters Degree	88	80	85

The teacher population at Application-Based High Schools consists of 41 full-time teachers with 23 males and 18 females with a racial makeup of 5% black, 90% white, and 5% Hispanic. The average years experience for teachers is 12.41 years with 88% holding a bachelors or masters degree. Thirteen teachers at Application-Based High School provided instruction in Georgia End of Course Test subjects during the 2008-2009 school year. Application-Based High School met Adequate Yearly Progress as defined by the Georgia Office of Student Achievement for 2007-2008 and 2008-2009.

The student sample at Theory-Based High School consists of 556 9<sup>th</sup>-12<sup>th</sup> grade students. The racial demographics of the students were 1% Asian, 55% black, 1% Hispanic, 1% multiracial, and 42% white. Additionally, the sample student population consists of 4% alternative programs, 56% economically disadvantaged, 0.5 % remedial, 22% students with disabilities, and 72 % vocational labs. Students at Theory-Based High School completed 831 Georgia End of Course Tests, with some students testing in multiple subjects, during the 2008-2009 school year.

The teacher population at Theory-Based High Schools consists of 40 full-time teachers and one part-time teacher with 22 males and 19 females with a racial makeup of 17% black, 80% white, and 2% Hispanic. The average years experience for teachers is 8.68 years with 80% holding a bachelors or masters Degree. Fourteen teachers at Theory-Based High School provided instruction in Georgia End of Course Test subjects during the 2008-2009 school year. Theory-Based High School did not meet Adequate Yearly Progress as defined by the Georgia Office of Student Achievement for 2007-2008 and 2008-2009.

The student sample at Control High School consists of 1,599 9<sup>th</sup>-12<sup>th</sup> grade students. The racial demographics of the students were 23% black, 2% Hispanic, 2% multiracial, and 73%

white. Additionally, the sample student population consists of 3% alternative programs, 30% economically disadvantaged, 11% gifted, 1% remedial, 6% students with disabilities, and 58% vocational labs. Students at Control High School completed 2,246 Georgia End of Course Tests, with some students testing in multiple subjects, during the 2008-2009 school year.

The teacher population at Control High School consists of 86 full-time teachers and six part-time teachers with 31 males and 61 females with a racial makeup of 10% black and 90% white. The average years experience for teachers is 12.74 years with 85% holding a bachelors or masters degree. Thirty-seven teachers at Control High School provided instruction in Georgia End of Course Test Subjects during the 2008-2009 school year. Control High School did not meet Adequate Yearly Progress as defined by the Georgia Office of Student Achievement for 2007-2008 but did meet Adequate Yearly Progress for 2008-2009.

Students were heterogeneously grouped throughout classrooms and were of mixed academic abilities based on standard class assignment procedures for each school. All students received instruction using the same curriculum based on Georgia Performance Standards. Standards, elements, performance tasks, exemplars, and curriculum guides were available to all teachers from the Georgia Department of Education at [www.georgiastandards.org](http://www.georgiastandards.org).

The intervention focus was on professional learning concerning assessment strategies to promote Assessment for Learning and formative assessment. Teachers at Application-Based High School were provided application-based professional learning from a Regional Educational Services Agency incorporating theory related to effective formative assessment practices. Teachers at Theory-Based High School were provided theory-based professional development from a national expert on assessment strategies. Teachers at Control High School did not participate in any professional development related to assessment strategies and functioned as the

control group. (See Appendix C for Power Points from trainings.) Classroom instruction techniques could have been adjusted based on assessment strategies learned in the professional development.

The population and participants invited to participate in the study provided a unique opportunity to study the effects of implementation of formative assessment practices. The implementation of standards-based classroom practices in the state of Georgia and the focus on formative assessment practices implementation within school in the West Georgia area at the time of the study were unique characteristics of schools within the study area. Future study of the topic in the same area would be limited and would not provide as detailed and specific information concerning formative assessment. The participants identified were ideal for the study at the time.

### **Instrumentation**

A survey was used in the Spring of 2009 designed to gather data on perception of use of assessment in the classroom. The survey used was from work completed by Bol, Stephenson, and O'Connell (1998) as part of a study for the University of Memphis. (See Appendix D for Bol survey. See Appendix E for actual study survey.) The survey was used with permission from the University of Memphis. The original survey was used to gather information on the influence of teaching experience, grade level, and subject area on assessment practice. The Bol survey was useful to the current study because of the survey data gathered on previous year use of assessments (summative and formative), teacher preparedness for using differing assessment methods, and teacher perception of the usefulness of particular assessment methods (summative and formative) in determinations of student learning and progress.



The survey is divided into several sections. Section One provides background data on teacher prior knowledge of assessment strategies. Section Two provides data on teacher perception of the benefit of formative assessment techniques. Demographic data provided information on individual teachers. Section Two of the survey was scored numerically, and two scale scores were identified (one for traditional assessment and one for formative assessment). The scale score was calculated as the mean range obtained across the items comprising each scale. A high scale score indicated positive perception, and a low scale score indicated a negative perception. Bol et al. (1998) found the reliability coefficient for the survey at .49 for the traditional assessment questions and .75 for the formative assessment questions (Bol, 1998). For the purposes of this study the formative assessment questions were analyzed and the .75 reliability coefficient found by Bol et al. yielded a strong reliability score at .75. Additionally, Bol (1998) indicated construct validity of the scales supported through factor analysis. For the purpose of this study the entire survey was administered to preserve construct validity. However, the formative assessment questions were the only ones examined. Responses to questions 4, 5, 6, 7, 18, 19, 20, 21, 25, 26, 27, and 28 were assigned a numeric value on a 1-5 scale and a total scale score was calculated for each survey. When all formative assessment questions were scored at a positive end of the 1-5 scale score a scale score of at least 36 would be attained. Therefore, a scale score of 36 or more would indicate a positive perception of formative assessment strategies on the survey. The scale score was analyzed in relationship to the predetermined positive scale score to determine a positive or negative perception. All scores over 36 were considered indicative of positive perception. All scales 36 and under were considered indicative of negative perception.

Student achievement data was gathered from school and system reports provided by the Georgia Department of Education and individual school systems for the standardized Georgia End of Course Test for 9th Grade Literature, American Literature, Geometry, Physical Science, Biology, United States History, and Economics for the 2007-2008 and 2008-2009 school years. The Georgia End of Course Tests were given each year to all students in the state of Georgia at the completion of the course for eight required high school courses which were 9th Grade Literature, American Literature, Algebra I, Geometry, Physical Science, Biology, United State History, and Economics. For the purposes of this study Algebra I and Geometry data was not analyzed due to the change in Mathematics course requirements for the state from 2007-2008 to 2008-2009 causing inconsistencies in the number of students in 2008-2009 participating in Algebra I and Geometry coursework. Individual students were given a numerical score on a 500 point scale. Additionally scale scores were converted to grade equivalencies on a 100 point scale with below 70 indicated a student does not meet the standard, above 70 indicating the student does meet the standard, and above 90 indicating the student exceeds the standard. The Georgia End of Course Tests were developed by Pearson Educational Measurement and vetted for validity and reliability through the national testing company's rigorous standards.

Additionally, test security measures were required at all Georgia schools to ensure the security of the test including documented signatures for receipt and return of test materials from the system test coordinator and at each level to the student using the materials for the tests. In order to preserve security, the tests were stored in a locked room except during administration. All test sessions were administered by a certified teacher and the room was arranged in order to avoid cheating. Each school was required to administer the test during the three week window assigned by the state at the end of each semester. All students completing the required eight

courses were required to complete the assessment. Each subject test consisted of two sections given on the same testing date with a short break between sections. Students were provided a minimum of 60 minutes to complete each section with the test session being extended as long as needed for students requiring additional time. Students with accommodations and modifications as indicated as part of their Individualized Education Plan, Special Education Students, or Individualized Adaptation Plan, Section 504 Students, received additional modification as allowed for in the state testing guidelines.

### **Procedures**

Once the sample was identified, the first step was to contact the school systems meeting population requirements as rural Georgia schools. As mentioned earlier in the chapter, for the purposes of anonymity schools were labeled as Application-Based High School, Theory-Based High School, and Control High School. School system administration was contacted in Fall 2008 and invited to participate in the study. (See Appendix A for contact letter.) Once school systems agreed to participate, each school was contacted individually and the researcher met with the school leader to discuss the research in the Fall of 2008. (See Appendix B for principal contact letter). The researcher provided each school leader with a brief overview of the purpose of the study. Intervention schools and their teachers participated in the intervention during the 2008-2009 school year.

Teachers at Application-Based High School participated in professional learning sessions in the Fall/Winter of 2008 by taking part in direct in-service instruction that was application-based in October 2008 and January 2009 provided by the Regional Education Services Agency. Professional Learning for this group was applications based and focus on practical strategies related to implementing formative assessment strategies in the classroom. All

teachers at Application-Based High School participated in both professional learning sessions. Teachers at Theory-Based High School participated in a single professional learning session provided by a national expert in formative assessment theory that was theory-based in January 2009. (See Appendix C for in-service handouts.)

Teachers indicated the type and date of in-service participation as part of the demographic information gathered as part of the survey. Teachers may have had some prior knowledge or may have previously implemented some formative assessment techniques as a result of independent study, therefore, teacher prior knowledge information was gathered as part of the survey. Classroom implementation of Assessment for Learning strategies was indicated as part of the teacher survey as well.

Teacher surveys were administered in Spring 2009, and surveys were administered as part of a scheduled faculty meeting at each school. Teachers were assured of anonymity and confidentiality as their data was only to be seen by the researcher. Teachers were allowed to opt out of participation as required under International Review Board regulations. Application-Based High School teachers responded to the survey and 76% of the faculty participated. Theory-Based High School teachers responded to the survey and 85% of the faculty participated. Control High School Teachers responded and 26% of the faculty participated. Teachers were encouraged to be candid in their responses. Teacher demographic and perception data was gathered using the survey tool in Appendix E.

Student Achievement data was gathered in Spring 2009 by request to the school leaders or his/her designee. Data gathered included Standardized Georgia End of Course Test Data for 9<sup>th</sup> Grade Literature, American Literature, Geometry, Physical Science, Biology, US History, and Economics for both the 2007-2008 and 2008-2009 school year. Data was gathered and indicated

as belonging to Teacher 1, Teacher 2, etc. for Application-Based High School, Theory-Based High School, or Control High School.

For the purpose of this study, multiple research designs were used. Student Data was identified as being from students at Application-Based High School, Theory-Based High School or Control High School and also matched to individual teachers at each school. Teacher survey data was classified as positive or negative based on the scaled score and was identified as from Application-Based High School, Theory-Based High School, or Control High School. For classes which were taught by the same teachers for the two years of the data gathered, Gain in Georgia End of Course Test Data from 2007-2008 to 2008-2009 was calculated and classified as from a teacher at Application-Based High School, Theory-Based High School or Control Based High School.

Data Analysis included analysis of descriptive statistics including measures of central tendency and measures of variability. Descriptive statistics were examined related to student performance on the Georgia End of Course Tests. The examination of descriptive statistics allowed the researcher to gather generalized information to assist in describing the results as a whole. The mean score was determined for all Georgia End of Course Test results and for results by individual school (Application-Based High School, Theory-Based High School, and Control Based High School) for 2007-2008 and 2008-2009. The mean score was determined for Georgia End of Course Test classroom results in each individual teacher's class at all three high schools for 2007-2008 and 2008-2009. The standard deviation was determined for all Georgia End of Course Tests results and for results by individual school (Application-Based High School, Theory-Based High School, and Control High School) for 2007-2008 and 2008-2009. The

standard deviation was also determined for classrooms results for each individual teacher at all three high schools for 2007-2008 and 2008-2009.

Once all descriptive statistics were analyzed more sophisticated statistical analysis included application of one-way ANOVA to test the significance of group differences. Analysis of Variance was ideal for this study due to the assumption on normality, assumption of homogeneity of variance, and assumption of independence of observations used as the primary underlying assumptions to test for significance of the findings. Testing using ANOVA allowed the researcher to use the variance of the scores for the entire group of participants independent of the truth or falsity of the null hypothesis compared to the variance of the scores dependent upon the null hypothesis to either reject or accept the null hypothesis. Ideally, the number of scores in each group would be equal however using additional calculations and adjusting the  $F$  ratio using degrees of freedom ( $df$ ) and a correlated Critical Values of  $F$  Distribution Chart ANOVA can provide sound statistical analysis. The ANOVA described the variance between three means as an  $F$  ratio providing the relationship between the between-group variability and the error variance of within group variability (Mertler & Vannatta, 2005). Using the value of the  $F$  ratio the research can determine if the differences between two groups are due to the treatment/intervention or simply due to chance (Mertler & Vannatta, 2005) after the critical  $F$  value was determined using the Critical value of the  $F$  Distributions: Alpha = .05 Chart (Howell, 2008). Application of the ANOVA was followed by post hoc tests to allow for comparison of differences between two individual groups. Multiple comparisons were then done using the Bonferroni tests to identify which groups a significant difference was observed between. The research questions were examined and the null hypothesis answered through data analysis as follows:

**RQ<sub>1</sub>** Is there a relationship between teacher participation in Assessment for Learning professional development and student achievement as measured by the Georgia End of Course Tests standardized assessment?

**H<sub>01</sub>**. There will be no significant difference between teachers who participated in Assessment for Learning professional development those that did not as shown by student performance on the Georgia End of Course Tests standardized assessment.

**Data analysis.** ANOVA identified the difference in groups and additional analysis was applied to determine which groups were different for a significant *F* value. For H<sub>01</sub> School Effect showed the amount of variation between schools (Application, Theory, and Control), Time Effect showed the difference before and after the intervention, and the School By Time Interaction showed the difference between the schools (Application, Theory, and Control) over time.

**RQ<sub>2</sub>** Is there a relationship between teacher participation in Assessment for Learning professional development and teacher perception of the benefits of Assessment for Learning strategies?

**H<sub>02</sub>**. There will be no significant difference between teachers who participated in Assessment for Learning professional development as measured by teacher perception of the benefits of Assessment for Learning strategies.

**Data analysis.** To evaluate the third hypothesis teacher participation in Assessment for Learning Professional development was used as the independent variable with teacher perception of the benefits of Assessment for Learning used as the dependent variable. A one-way ANOVA was applied with a *F* ratio greater than the critical *F* ratio from the Critical Values of *F* Distribution Chart: Alpha =

.05 (Howell, 2008) indicating statistical significance. If a significantly different  $F$  ratio was obtained the null hypothesis was rejected and additional follow up post hoc Bonferroni tests were run to identify which group combinations were significantly different.

**RQ<sub>3</sub>.** Does a relationship exist between teacher perception of the benefit of formative assessment and student achievement as measured by the Georgia End of Course Tests standardized assessment?

**H<sub>03</sub>.** No significant relationship exists between teacher positive perception of the use of formative assessment and student achievement as measured by student performance on the Georgia End of Course Tests standardized assessment.

**Data analysis.** To evaluate the fourth hypothesis teacher perception was used as the independent variable with mean teacher classroom Georgia End of Course Test scores used as the dependent variable. Person product-moment correlation coefficient ( $r$ ) was calculated to determine the correlation between teacher perception of formative assessment and student performance.

Chapter Three provided an outline for the study including the design, questions, hypotheses, population to be studied and the instrumentation, procedures, and methodology to be used. Chapter Four and Five will examine the data gathered during the course of the study and the discussion and conclusions to be gained from examination of this data. The study gathered data and analyzed it in an effort to better understand the relationship between professional learning, teacher perception of formative assessment techniques, and student achievement on



standardized Georgia End of Course Tests. A detailed presentation of the data collected and analyzed is included in Chapter Four in tables and narrative texts.

## **Chapter Four: Results**

The purpose of this study was to compare the differences in student performance as measured by the Georgia End of Course Tests standardized assessment based on teacher participation in two different types of professional development on Assessment for Learning and formative assessment. Additionally, this researcher sought to examine the effects of participation in two different types of professional development on Assessment for Learning and formative assessment on teacher perception of the benefits of formative assessment based on. Finally, the researcher sought to compare the difference in student performance as measured by the Georgia End of Course Tests standardized assessment based on teacher perception of the benefits of formative assessment. The study was based on the population of three rural west Georgia schools including 174 teachers who participated in application-based, theory-based, or no professional learning related to formative assessment and 2,787 students who completed standardized Georgia End of Course Tests during the 2008-2009 school year.

Chapter Four is organized around general descriptive statistics and provides more advanced statistical analysis of data related to the three research questions and five null hypotheses presented in Chapter One and Chapter Three. It presents the statistical analysis as outlined in Chapter Three including descriptive statistics as well as ANOVA and Pearson product-moment correlation analysis. Results are presented in tables and text. A summary concludes the chapter.

### **Descriptive Statistics**

Information on student participation in Georgia End of Course Test standardized assessments, teacher retention at each school and teacher response to the survey are provided as

part of this chapter. Additionally, general descriptive statistics including measures of central tendency and measures of variability describe the population within the study. For each school differing numbers of students participated in courses which required a Georgia End of Course Test Assessment and some students completed multiple assessments. Table 4.1 provides summary data of student participation in Georgia End of Course Test Assessment.

At Application-Based High School 632 students received instruction from 13 core content teachers and completed 818 Georgia End of Course Tests in 2008-2009. At Theory-Based High School 556 students received instruction from 14 core content teachers and completed 831 Georgia End of Course Tests in 2008-2009. At Control High School 1,599 students received instruction from 37 core content teachers and completed 2,246 Georgia End of Course Test in 2008-2009. Georgia End of Course Test Data was available for classrooms at all three schools for 2007-2008 as well. At Application-Based high School 12 teachers provided core content instruction and students completed 1,042 Georgia End of Course Tests in 2007-2008. At Theory-Based High School 8 teachers provided core content instruction and students completed 847 Georgia End of Course Tests in 2007-2008. At Control High School 34 teachers provided instruction and students completed 2,242 Georgia End of Course Tests in 2007-2008.

For each school only a portion of the core content teachers provided instruction for both years of the study. At Application-Based High School 12 teachers, 92% of the core content staff in 2008-2009, provided instruction to students at the school in 2007-2008 and 2008-2008 . Students in their classes completed 1,021 Georgia End of Course Tests in 2007-2008 and 796 Georgia End of Course Tests in 2008-2008. At Theory-Based High School four teachers, 29% of the core content staff in 2008-2009, provided instruction to students at the school in 2007-2008 and 2008-2009. Students in their classes completed 432 Georgia End of Course Tests in

2007-2008 and 226 Georgia End of Course Tests in 2008-2009. At Control high School 20 teachers, 54% of the core content staff in 2008-2009, provided instruction to students in 2007-2008 and 2008-2009.

Table 4.1

*Participation in EOCT 2008-2009 & 2007-2008*

	Application-Based High School	Theory-Based High School	Control High School
<i>N</i> Students in EOCT Core Content Classes 2008-2009	632	556	1,599
<i>N</i> Core Content Teachers 2008-2009	13	14	37
<i>N</i> EOCT Completed 2008-2009	818	831	2,246
<i>N</i> Core Content Teachers 2007-2008	12	8	34
<i>N</i> EOCT Completed 2007-2008	1,042	847	2,242
% Core Content Teachers in Same Subject 2007-2008 & 2008-2009	92	29	54

Students in their classes completed 1,779 Georgia End of Course Tests in 2007-2008 and 1,696 Georgia End of Course Test in 2008-2009. (See Appendix G for individual End of Course Test scores).

All teachers at the three schools were presented with the opportunity to participate in the formative assessment survey. The survey was distributed to 41 teachers at Application-Based High School, 41 teachers at Theory-Based High School, and 92 teachers at Control High School. Thirty-seven teachers, 76%, completed the survey at Application-Based High. Thirty-five teachers, 85%, completed the survey at Theory-Based High. Twenty-four teachers, 26%, completed the survey at Control High. Table 4.2 provides summary data for teacher participation in the formative assessment survey.

Table 4.2

*Teacher Participation/Return Rate for Formative Assessment Survey*

	Application-Based High School	Theory-Based High School	Control High School
<i>N</i> Teachers	41	41	92
% Teachers Responded	76	85	26

Using a scale score at more than 36 total on the survey as providing a positive response concerning formative assessment use in the classroom and 35 or less total on the survey as providing a negative response concerning formative assessment use in the classroom. At Application-Based High School, 27 teachers, 87%, indicated a positive perception concerning formative assessment use in the classroom. At Theory-Based High School, 25 teachers, 71%,

indicated a positive perception concerning formative assessment use in the classroom. At Control Based High School, 23 teachers, 96%, indicated a positive perception concerning formative assessment use in the classroom. Table 4.3 provides summary information for survey response.

Table 4.3

*Formative Assessment Survey Summary Information*

	Application-Based High School	Theory-Based High School	Control High School
<i>N</i> Respondents	37	35	24
% Positive Perception	87	71	96

Mean results for each question concerning formative assessment, those used as part of this study's data analysis, are provided in Table 4.4. Table 4.4 shows that all schools had a mean score indicating an overall positive perception of formative assessment in the classroom.

Questions 4, 18, and 25 related to use of portfolio assessment (a collection of assignments, work samples) in the classroom. When asked in Question 4 how frequently they use portfolio assessment teachers at Control High responded with most frequent use of portfolios, followed by teachers at Theory-Based High, with teachers at Application-Based High using portfolios the least. Teachers at Control High also felt portfolio usage had more of an impact on student achievement than did teachers at Theory-Based High or Application-Based High. Teachers at Control High ranked above the norm and felt most prepared to use portfolios,

teachers at Application-Based High felt the next most prepared, with teachers at Theory-Based High feeling the least prepared.

For observation (evaluating participation, group work), Control High ranked the most frequent usage with Application-Based High the next frequent, and Theory-Based High the least frequent. Also, Control High ranked first in feeling prepared to implement observation as an assessment, with Application-Based High ranking second in preparedness, and Theory-Based High feeling least prepared. When asked to rank the degree of impact of observation on student achievement Application-Based High had the highest score for observation with Control High second, and Theory-Based High ranking the lowest for observation.

For all other forms of formative assessment the school rankings were identical to each other. For performance task (assessment of students as they work on a problem or a task) and self-assessment by students Control High ranked highest in usage, preparedness, and likeliness of effecting students. Application-Based High ranked second in usage, preparedness, and likeliness of effecting students for both. Theory-Based High ranked least in usage, preparedness, and likeliness of effecting students.

All three schools ranked observation as the most often used form of formative assessment. Application-Based High ranked observation as the formative assessment strategy they felt most prepared to implement. Control High and Theory-Based High, both, ranked performance task as the formative assessment strategy they felt most prepared to implement. All three schools ranked performance task as the most likely to have an impact on student learning.

Overall Control High scored formative assessment more positive with a mean survey score of 48.08. Although not as strong as Control High, Application-Based High scored formative assessment as positive, with a mean survey score of 44.55. Finally, although not as

strong as Control High or Application-Based High, Theory-Based High scored formative assessment as positive, with a mean survey score of 42.37.

Table 4.4

*Formative Assessment Perception Survey Results: Mean Responses and Standard Deviation by Question for Study Groups*

SCHOOL	APPLICATION HIGH		THEORY HIGH		CONTROL HIGH		OVERALL	
<i>N</i>	31		35		24		90	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Q4	2.65	1.23	2.77	1.31	2.79	1.22	2.73	1.24
Q5	4.19	0.95	3.86	1.12	4.54	0.72	4.16	0.99
Q6	3.94	1.03	3.74	1.29	4.46	0.59	4.00	1.08
Q7	2.81	1.05	2.51	1.34	3.04	1.16	2.76	1.20
Q18	3.52	1.36	3.43	1.20	3.96	1.16	3.60	1.25
Q19	4.55	0.68	4.20	1.05	4.63	0.58	4.43	0.84
Q20	4.32	0.98	4.26	0.92	4.67	0.56	4.39	0.87
Q21	3.52	1.36	3.11	1.28	3.96	0.95	3.48	1.27
Q25	3.55	1.31	3.69	1.13	4.04	0.93	3.73	1.16
Q26	4.10	1.08	3.89	0.93	4.04	0.91	4.00	0.97
Q27	4.29	1.01	4.03	1.10	4.38	0.65	4.21	0.97
Q28	3.13	1.26	2.89	1.08	3.91	0.95	3.24	1.18
TOTAL	44.55	8.29	42.37	8.45	48.08	6.52	44.64	8.16

In addition to this data, general descriptive information on student performance on the Georgia End of Course Test was gathered. Table 4.5 provides summary information of the mean scale score and grade conversion for all Georgia End of Course Test given at Application-Based High School, Theory-Based High School, and Control High School for the 2007-2008 and 2008-2009 school year. Standard Deviations for the set of scores are also provided.

The descriptive statistics in Table 4.5 allow for a general picture of the performance of students at Application-Based High, Theory-Based High, and Control High on the standardized Georgia End of Course Test for the school year 2007-2008, prior to intervention, and the school year 2008-2009, after intervention. The statistics in Table 4.5 provide information on all



students and all teachers at each school during the period of the study. In 2007-2008, the mean scale score and grade conversion for Application-Based High was higher than the mean scale score and grade conversion for Control High, with Theory-Based High scoring the lowest. In 2008-2009 Application-Based High still had the highest mean scale score and grade conversion, and Theory-Based High had the lowest mean scale score and grade conversion.

To gather a more detailed picture of the performance of students at Application-Based High, Theory-Based High, and Control High for those teachers who were teaching the same classes during 2007-2008 and 2008-2009 additional descriptive statistical analysis was performed and is summarized in Table 4.6. When comparing gains for all three schools overall scale score gains were highest at Application-Based High School with a gain of 15.74 , second highest at Theory-Based High School with a gain of 9.10, and lowest at Control High School with a gain of 4.79. All three schools had gains in mean scale score from 2007-2008 to 2008-2009. The same pattern was seen in grade conversion with Application-Based High gaining 3.87 in grade conversion, Theory-Based High gaining 2.72 in grade conversion, and Control High gaining 1.57 in grade conversion.

Table 4.5

*Georgia End of Course Test Mean Scale Score and Grade Conversion with Standard Deviation for All Scores*

School	Year	<i>N</i>	<i>MSS</i>	<i>SS SD</i>	<i>M</i> Grade Conversion	<i>SD</i> Grade Conversion
Application High	2007-2008	1042	423.70	41.06	77.25	11.77
	2008-2009	818	437.72	45.22	80.75	11.27
Theory High	2007-2008	847	403.72	39.21	71.04	12.86
	2008-2009	831	399.54	37.86	69.88	12.82
Control High	2007-2008	2242	420.75	42.93	76.34	12.40
	2008-2009	2246	428.43	41.83	78.81	12.02

Table 4.6

*Georgia End of Course Test Mean Scale Score and Grade Conversion with Standard Deviation for Repeat Teachers*

School	Year	<i>N</i>	<i>M SS</i>	<i>SS SD</i>	<i>M Grade Conversion</i>	<i>SD Grade Conversion</i>
Application High	2007-2008	1021	422.65	77.02	77.02	11.74
	2008-2009	796	428.44	45.44	80.89	11.26
	Gain		15.74		3.87	
Theory High	2007-2008	432	406.28	39.28	72.06	12.60
	2008-2009	226	415.38	38.39	74.78	11.72
	Gain		9.10		2.72	
Control High	2007-2008	1779	426.38	41.64	78.11	11.45
	2008-2009	1696	431.17	41.14	79.68	11.64
	Gain		4.79		1.57	

### Statistical Analysis of Research Questions

Descriptive statistical analysis including information on measures of central tendency and measures of variability provided a general picture of the data gathered in this study.

Additionally, in depth statistical analysis was performed related to each research question and null hypothesis to provide more concrete information for the study and to further determine if the difference seen in the descriptive statistical analysis were statistically significant. The more complex data analysis is presented for each research question and corresponding null hypotheses.

**RQ<sub>1</sub>.** Is there a relationship between teacher participation in Assessment for Learning professional development and student achievement as measured by the Georgia End of Course Tests standardized assessment?

**H<sub>01</sub>.** No significant difference exists between teacher participation in Assessment for Learning professional development and student achievement as measured by student performance on the Georgia End of Course Tests standardized assessment.

To evaluate the first null hypothesis teacher participation in professional development was used as the independent variable with Georgia End of Course Test scores used as the dependent variable. Specifically, analyses examined trends in standardized assessments before and after training. Because different students took the Grade 9 Georgia End of Course Tests in 2007-2008 and 2008-2009, it was not possible to use the 2007-2008 scores as a covariate in an ANCOVA analysis of 2008-2009 scores. Instead, scores were analyzed within an Analysis of Variance framework in which three levels of School (Theory, Application, and Control) were crossed with two levels of time (2007-2008 and 2008-2009). If teacher professional development made a difference to students' performance on the Georgia End of Course standardized assessments, then the two way interaction between School and Time should be significant. A significant two-way interaction would indicate that the trend in test scores over time differed between Theory, Application, and Control schools. If professional development increased performance on the Georgia End of Course test, then further examination of the trends should show that scores increased more over time in the Theory and Application schools than in the Control schools

**Data Analysis.** Scores on the Georgia End of Course Assessments increased significantly between 2007-2008 and 2008-2009 and differed significantly between schools (see Table 4.7). Most critically for the test of Hypothesis One, the two-way School by Time interaction was significant. Though scores increased in all three schools, they increased most in the Application School (+15 points) followed by the Theory School (+9 points) and Control School (5 points). Thus, the pattern of findings supports the view that participation in Assessment for Learning professional development is associated with improvements in students' performance on the Georgia End of Course Assessment.

Table 4.7

*Means and ANOVA for Application, Theory and Control High Georgia End of Course Tests Scores for 2007-2008 and 2008-2009*

School	Year						Change
	2007-2008			2008-2009			
	M	(SD)	n	M	(SD)	n	
Application	423	(40)	1021	438	(45)	796	+15
Theory	406	(39)	432	415	(38)	226	+ 9
Control	426	(42)	1779	431	(41)	1696	+ 5
		Sum of Squares	df	Mean Square	F		
School		186927	2	93463	54.180 ***		
Time		86697	1	86697	50.258 ***		
School X Time		35452	2	17726	10.276 ***		

Note

\*\*\*  $p < .0001$

**RQ<sub>2</sub>.** Is there a relationship between teacher participation in Assessment for Learning professional development and teacher perception of the benefits of Assessment for Learning strategies?

**H<sub>0</sub><sup>2</sup>.** No significant difference exists between teacher participation in Assessment for Learning professional development and teacher

positive perception of the benefits of Assessment for Learning strategies.

Table 4.8  
*ANOVA and Bonferroni Test of Application, Theory, and Control High Mean Perception Score*

School	M	(SD)	n	SE	Pooled SE
Application	44.5	8.3	31	1.49	1.42
Theory	42.4	8.5	35	1.43	1.34
Control	48.1	6.5	24	1.33	1.62

  

	Sum of Squares	df	Mean Square	F	p
Groups	464.9	2	232.5	3.7	0.0
Residual	5465.7	87	62.8		287
Total	5930.6	89			

  

Bonferroni Contrast	Difference	95% CI
Application vs. Theory	2.2	-2.6 to 6.9
Application vs. Control	-3.5	-8.8 to 1.7
Theory vs. Control	-5.7	-10.8 to 0.6 (significant)

**Data analysis.** To evaluate the second hypothesis teacher participation in Assessment for Learning Professional development was used as the independent variable with teacher perception of the benefits of Assessment for Learning used as the dependent variable. A one-way ANOVA

was applied with an  $F$  ratio greater than the critical  $F$  ratio from the Critical Values of  $F$  Distribution Chart: Alpha = .05 (Howell, 2008) indicating statistical significance. For null hypotheses two critical  $F_{.05}(2, 87) = 3.38$ . As presented in Table 4.8 analysis produced  $F = 3.7$  which is statistically significant therefore,  $H_02$  was rejected supporting a difference in teacher perception of formative assessment. Additional follow up post hoc Bonferroni tests were run to identify which group combinations were significantly different.

The Bonferroni post hoc test was conducted to determine which professional development group categories were significantly different in teacher perception of formative assessment. Results, included in Table 4.8, did not show any significant difference between Application-Based High and Theory-Based High nor was a difference found between Application-Based High and Control High. However, post hoc tests showed significant differences existed between Theory-Based High and Control High and that difference between Theory-Based High and Control High was negative in nature yet again.

**RQ<sub>3</sub>.** Does a relationship exist between teacher perceptions of the benefit of formative assessment and student achievement as measured by the Georgia End of Course Tests standardized assessment?

**H<sub>03</sub>.** No significant relationships exist between teacher positive perception of the use of formative assessment and student achievement as measured by student performance on the Georgia End of Course Tests standardized assessment.

**Data analysis.** To evaluate the third null hypothesis, teacher perception was used as the independent variable with 2008-2009 Georgia End of Course Test scores used as the dependent variable. Teacher participation in professional development was not considered. To examine the association between perceptions of professional development and scale scores, a Pearson

Product-Moment correlation was computed. This coefficient shows the strength of association between a binary variable (positive versus negative perception) and a continuous one (end of course scores). In the present sample, the correlation between perceptions of the professional development program and students' performance on the Georgia End of Course Assessments was not statistically significant ( $r(659) = .054; ns$ ), so the Null Hypothesis is retained.

### **Data Summary**

Descriptive statistics provide an overall picture of Georgia End of Course Test results for Application-Based High School, Theory-Based High School, and Control High School. ANOVA results indicated a statistically significant differences for Georgia End of Course Test results over time between teachers participating in Application-Based Formative Assessment professional development, Theory-Based Formative Assessment professional development, and the control group. ANOVA results and follow up Bonferroni post hoc tests also indicated a statistically significant difference in teacher perception between teachers participating in Theory-Based Formative Assessment professional development and the control group. A negative difference was found between the Theory-Based group and the Control group. Pearson-Product Moment correlations did not reveal a statistically significant difference to teacher perception of formative assessment based on participation in professional development. Chapter Five will present conclusions, discussion, and implications from the study.

## Chapter Five: Summary and Discussion

### Statement of the Problem

The aim of this research was to further study the relationship between teacher participation in targeted professional development, teacher perception of formative assessment, and student performance as measured by student summative assessment performance defined as student scores on the standardized Georgia End of Course Tests. Review of the literature identified a clear link between sound classroom assessment practices and student achievement gains and , consequently, increased student performance on summative assessments. Review of the literature also indicated a disconnect for teachers in understanding the link between classroom assessment and student learning and , therefore, student gain on summative assessment measures. It was thought that further study of the relationship between participation in targeted professional development, teacher perception of formative assessment, and student performance would provide additional data related to understanding more clearly the topic of formative assessment. The study focused on three major questions.

- RQ<sub>1</sub>. Is there a relationship between teacher participation in Assessment for Learning professional development and student achievement as measured by the Georgia End of Course Tests standardized assessment?
- RQ<sub>2</sub>. Is there a relationship between teacher participation in Assessment for Learning professional development and teacher perception of the benefits of Assessment for Learning strategies?
- RQ<sub>3</sub>. Does a relationship exist between teacher perceptions of the benefit of formative assessment and student achievement as measured by the



## Georgia End of Course Tests standardized assessment?

### Review of the Methodology

As stated in chapter three, three high schools were identified to participate. The schools are similar demographically and all three are located in the rural west Georgia area. Teachers at Application-Based High School participated in application-based professional development on formative assessment. Teachers at Theory-Based High School participated in theory-based professional development on formative assessment. Teachers at Control High School did not participate in targeted professional development (the control group). Teachers within the school provided survey data. Teachers within the school who provided instruction in core content areas of 9<sup>th</sup> Grade Literature, American Literature, Physical Science, Biology, United States History, and Economics during the 2008-2009 year participated in the study to provide classroom Georgia End of Course Test Data for that year alone. Teachers within the school who provided instruction in core content areas of 9<sup>th</sup> Grade Literature, American Literature, Physical Science, Biology, US History, and Economics during the 2007-2008 school year and provided instruction in the same content area for the 2008-2009 school year participated in the study to provide classroom Georgia End of Course Tests data for both years to compare gains.

### Summary of the Findings

**RQ<sub>1</sub>.** Is there a relationship between teacher participation in Assessment for Learning professional development and student achievement as measured by the Georgia End of Course Tests standardized assessment?

**RQ<sub>1</sub> findings.** ANOVA results, presented in Table 4.4, show a significant difference in the Georgia End of Course Test Scores between the three participating groups over time ( $F=10.276$  with critical  $F_{.05}=3.01$ ) .

Study results supported research from Stiggins and Chappius (2005), Bloom (1984), and Black and Wiliam (1998), showing that strong, effective assessment techniques (such as those demonstrated in the Application-Based High professional development) lead to increased student achievement. This study adds to the current body of research confirming in another way that formative assessment is linked to increased student learning. The study also supports Rodriquez and Ballanca's 2007 findings supporting the positive link between professional development and teacher implementation of quality assessment strategies that increase student learning.

**RQ<sub>2</sub>.** Is there a relationship between teacher participation in Assessment for Learning professional development and teacher perception of the benefits of Assessment for Learning strategies?

**RQ<sub>2</sub> findings.** ANOVA results, presented in Table 4. 6, showed a significant difference in teacher perception ( $F = 3.7$  with critical  $F_{.05} = 3.38$ ) between the three groups. The Bonferroni post hoc test results revealed that significant negative difference is found between Theory-Based High and Control High teacher perception of formative assessment ( Difference = -5.7 with 95% Confidence Interval). This study highlights again the need for further study of the relationship between professional learning and teacher perception. Gilson (2009) found a positive relationship between a professional learning community type of professional development and teacher implementation of formative assessment strategies. The current study found no significant difference between the perception of those participating in application-based professional development and the control group. However, this study found a negative difference between the perceptions of teacher participating in theory-based professional development when compared to the control. Overall these findings suggest that further study in this area would be beneficial.

**RQ<sub>3</sub>.** Does a relationship exist between teacher perceptions of the benefit of formative assessment and student achievement as measured by the Georgia End of Course Tests standardized assessment?

**RQ<sub>3</sub> findings.** Person Product-Moment Correlations show no significant differences exist in Georgia End of Course Test score for classrooms where the teacher had a positive perception compared to classrooms where the teacher had a negative perception. Inadequate data was available through the study data collection process to compare Georgia End of Course Test scores for 2007-2008 and 2008-2009 for teachers with a negative perception of formative assessment. Nash (2008) found that teachers with a positive perception of formative assessment used more formative assessment strategies in the classroom. This study focused on the direct relationship between perception and performance and did not find a relationship that was statistically significant.

### **Interpretation of the Findings**

The research findings suggest that teacher participation in focused professional development on formative assessment had a relationship with their student's scores on the Georgia End of Course Tests. Data from test scores compared over time for the different groups were significantly different based on professional development participation with an F ratio of 10.276 (See Table 4.7).

This study found that teachers who participated in application-based formative assessment professional development had student who scored better, when compared to a control group, on standardized achievement tests. The study also found that teachers who participated in theory-based professional development did not score as well, when compared to a control group, on standardized achievement tests.

When examining the relationship between professional development and teacher perception of formative assessment, the research supports a difference in perception based on the group teachers participated in (application-based professional development, theory-based professional development, control). Table 4.1 provides additional information on survey data by school. Those participating in application-based professional development had a mean scale score on the perception survey of 44.5. Those participating in theory-based professional development had a mean scale score on the perception survey of 42.4. Those participating in the control group had a mean scale score on the perception survey of 48.1. A small difference ( $F=3.7$ ) was found between the three groups. Finally, when examining the relationship between positive and negative perception and student scores on the standardized Georgia End of Course Test the research did not find a significant relationship between teacher perception of formative assessment and student performance on standardized achievement tests such as the Georgia End of Course Test ( $r(659) = .054:ns$ ). The students at Application-Based High had higher scale scores on the Georgia End of Course Test than those at Control High or at Theory-Based High. Overall, students with teachers participating in application-based professional development on formative assessment did show a greater gain in scale score (see Table 4.6 and 4.7) than those whose teachers participated in theory-based professional development on formative assessment or the control group and when comparing group difference over time a statistical significance was found (See Table 4.7). The teachers at all three schools participating in the study had an overall positive perception of formative assessment with those teachers at Control High ranking their usage, preparedness, and belief in the impact of formative assessment highest of all three groups. However, when comparing the differences for significance none was found.

## **Conclusions**

Professional development related to formative assessment for teachers does have a relationship to student performance. Students whose teachers have participated in application-based formative assessment professional development show increased test scores and increase those scores more than students whose teachers have participated in theory-based professional development and more than students whose teachers have not participated in professional development related to formative assessment. Schools concerned with increasing student performance and student learning should investigate the idea of formative assessment and provide concrete, practical, application-based professional development for their teachers on implementing formative assessment in their classrooms.

Initial descriptive statistics support the assertion that formative assessment makes a difference in the learning of students in the classroom. The results of this study when viewed overall could suggest that initial perception (such as the high positive perception of the control group) may be just as effective as professional development participation and contribute to similar gains as those achieved by students whose teachers have sound professional development. However, noting the impact of professional development over time the ANOVA results show that even though both the control group and the Application-Based Group both made improvements, the improvements in test scores were significantly higher for the Application-Based group (See Table 4.7).

## **Implications**

Schools working to improve student performance, especially in this era of high stakes testing, would benefit from spending time gathering data on their teachers' perception of formative assessment and studying that data to determine the overall perception and use of

formative assessment in the school. Providing professional development related to formative assessment could benefit teacher's classroom practices and result in improved student scores on standardized assessments such as the Georgia End of Course Test. Application-Based Professional Development throughout the year appears to be more effective in impacting resulting student performance and teacher positive perception than one shot professional development from national experts.

Although this study is limited to the high school population in rural west Georgia, and generalizations cannot be made to all students, the findings suggest that some relationship exists between understanding formative assessment techniques and practical applications based strategies for implementation and improved student performance on standardized assessments.

### **Research Applications**

The findings in this research provide additional information to add to the already large volume of work on formative assessment. This study points to the existence of a relationship between teacher perception and student performance and raises questions about the possibility of this perception superseding the effects of professional development in impacting improved student achievement because of use of formative assessment. Careful study of the data within this research raised more questions as opposed to providing any concrete answers. Application of this research can best be accomplished by studying the limitations and identifying areas of further study.

### **Limitations**

The study was limited in several ways as it was not purely experimental and required the use of preexisting schools, classes, teachers, and the accompanying history and culture of the location. Threats to internal validity may have occurred including:

- As the study was conducted in real schools differences in the schools may have impacted the findings. Application-Based High and Theory-Based High are more similar in size and almost half the size of Control High. However, Application-Based High and Control High had more similar student populations demographically and both were the only high school in their district, where Theory-Based High was not.
- Prior teacher knowledge related to formative assessment or other instructional strategies may have influenced the findings. With the focus on No Child Left Behind, school improvement, and increased student achievement, it is feasible that individual teachers in all study groups may have participated in prior learning activities related to formative assessment.
- Demographics of each school may have impacted the study. The researcher attempted to find school similarly situated demographically and geographically. Theory-Based High's student population consisted of more economically disadvantaged students and more minority students than the other two schools. Application-Based High's student population was most similar to Control High's population but had more economically disadvantaged students. However, as mentioned in the participants section, all schools are different and the demographic differences between the schools participating in the study could have influenced the findings.
- In addition to the professional development providing as part of the study, other improvement efforts at the three schools could have impacted the study. As it was not feasible to isolate the three schools and organize the schools improvement efforts related to the subject of the study, other efforts in the schools could have effected student performance and impacted the findings.

- Survey results were not received from all teacher participants. As survey participation was voluntary teacher choice to complete the survey may have been influenced by extreme negative or positive perception or by other unrelated factor. Survey participation or non-participation could have influenced the findings.

### **Recommendation for Further Research**

Based on the findings of the study many areas of further research could be explored.

Recommendations include:

- The study indicated a relationship between professional development participation and student achievement on standardized Georgia End of Course Tests. The study was conducted in three schools in rural west Georgia. Replicating the current study with a larger group of schools including more suburban and urban schools could be beneficial in verifying results with a larger population base.
- The study indicated a relationship between professional development participation and student achievement on standardized Georgia End of Course Tests. This study was limited to implications for students in grades 9-12. Broadening the study to include professional development for teachers in grades 1-8 and examining the related standardized assessment such as the Georgia Criterion Referenced Competency Test could be beneficial in verifying results with a larger population.
- The study was limited and the researcher noted concerns about teacher prior perception influencing the findings. Replicating the study and gathering perception data prior to the study and having pre and post intervention perception data could be beneficial in addressing this limitation.



- The study was limited due to the possible other improvement activities being conducted at the schools. Replicating the study and locating schools willing to focus on formative assessment professional development as their primary improvement goal for the duration of the study could be beneficial in controlling for this limitation.
- The study was limited due to possible other improvement activities being conducted at the schools. Replicating the study and cross populating the professional development groups to include participants from different schools could be beneficial in controlling for this limitation. This would entail identifying teachers for across multiple schools willing to participate in different professional learning groups (i.e. School A- 1/3 faculty application-based learning, 1/3 faculty theory-based learning, 1/3 faculty control; School B – 1/3 faculty application-based learning, 1/3 faculty theory-based learning, 1/3 faculty control).
- The study examined participation in professional development but did not examine implementation of professional development strategies in the classroom. Further research could be conducted examining teacher use of formative assessment in the classroom and its impact on student achievement on standardized assessments.

### **Summary Thoughts**

The information gathered over the course of this research provided data to examine student performance in three rural Georgia high schools. Survey information in this study also provided information on the perception teachers have about formative assessment in these same three schools. Considering the amount of research available to support the assertion that formative assessment does impact student performance on standardized test this study, ideally, would have found supporting data. The data gathered through this research showed some

significant difference in student performance between schools and also based on teacher perception. The suggestions for further research provided should certainly be considered as areas worthy of further exploration.

Decades of research support the link between sound practices in formative assessment impacting student achievement. This study found some support for this link but raised questions in other areas. The need to investigate more the link between perception of formative assessment and increased student achievement is obvious. Initially a clear assumption was made between professional development and implementation in the classroom. This study did not investigate this aspect teacher growth. The need to investigate more the link between learning and doing for teachers is obvious. This researcher has changed from a mind set of showing the best practice and assuming other see it to wanting to investigate more the link between showing the research based best practice and connecting that to a belief that implementing it will improve student learning. There is more to study, investigate, and analyze, as educators work together to improve instruction so students can learn more.

## References

- Allen, D., Ort, S.W., & Schmidt, J. (2009). Supporting classroom assessment practice: Lessons from a small high school. *Theory Into Practice, 48*(1), 72-80.
- Armstrong, T. (2006). *The best schools*. Alexandria, Virginia: Association for Supervision and Curriculum Development.
- Biddle, B., & Berliner, D. (1998). *The manufactured crisis*. Don Mills, Ontario: Addison-Wesley Publishing Company.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Educational Assessment: Principle Policy and Practice, 5*(1), 7-74.
- Bloom, B. (1984). The search for methods of group instruction as effective as one-to-one tutoring. *Educational Leadership, 41*(8), 4-17.
- Bol, L., Stephenson, P., O'Connell, A., & Nunnery, J. (1998, July). Influence of experience, grade level, and subject area on teachers' assessment practices. *Journal of Educational Research, 91*(6), 323. Retrieved from Academic Search Complete database.
- Brookhart, S.M. (2004). Classroom assessment: Tensions and intersections in theory and Practice. *Teachers College Record, 10* (3), 429-458.
- Brookhart, S.M. (2008). *How to give effective feedback to your students*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Bruner, J. (1960). *The process of education*. Cambridge, Massachusetts: Harvard University Press.
- Bucci, P. (2002). *Teacher knowledge, beliefs and practices of classroom assessment:*

- From the perspective of five elementary teachers* (Ed.D. dissertation). Retrieved from Dissertations & Theses: Full Text database. (Publication No. AAT NQ74612)
- Carr, J.F., & Harris, D.E. (2001). *Succeeding with standards*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Cauley, K.M. & McMillan, J.H. (2010). Formative assessment techniques to support student motivation and achievement. *Clearing House*, 83(1), 1-6.
- Chappius, J. (2005, November). Helping students understand assessment. *Educational Leadership*, 39-43.
- Chappius, S. & Stiggins, R.J. (2002, September). Classroom assessment for learning. *Educational Leadership*, 40-43.
- Colburn, A. (2009). An assessment primer. *Science Teacher*, 76(4), 10.
- Davies, A. (2000). *Making classroom assessment work*. Canada: Hignell Printing Limited.
- Dorn, S. (2010). The political dilemma of formative assessment. *Exceptional Children*, 76(3), 325-337.
- Fisher, D. & Frey, N. (2008). *Checking for understanding: Formative assessment techniques for your classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Gilson, R. (2009). *Professional development in assessment for learning* (Ed.D. Dissertation). Retrieved from Dissertations & Theses: Full Text Database. (Publication No. AAT3353674)
- Gregory, K., & Cameron, C., & Davies, A. (1997). *Knowing what counts: Setting and*

- using criteria*. Merville, British Columbia: Connections Publishing.
- Guskey, T.R. (2005, November). Mapping the road to proficiency. *Educational Leadership*, 32- 38.
- Harlan, W. (2005). Teachers' summative practices and assessment for learning – Tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- Harris, L. (2007). Employing formative assessment in the classroom. *Improving Schools*, 10(3), 249-260.
- Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi-Delta Kappan*, 89(2), 140-146.
- Jerald, C.D. (2001). *Dispelling the myth revisited*. Washington, D.C.: Education Trust.
- Lewis, R., Berghoff, P., & Pheeneey, P. (1999). Focusing students: Three approaches for learning through evaluation. *Innovative Higher Education*, 23(3), 181-196.
- Lincoln, Y., & Guba, E. (1984). *Naturalistic inquiry*. Beverly Hills, CA: Sage Publications.
- McManus, S,M, (2008). *A study of formative assessment and high staked testing: Issues of student efficacy and teacher views in the mathematics classroom* (Ph.D. dissertation) Retrieved from Dissertations & Theses: Full Text database. (Publication No. AAT 3306610)
- McNamee, G.D. & Chen, J. (2005, November). Dissolving the line between assessment and teaching. *Educational Leadership*, 72-76.
- McTighe, J. & O'Conner, K. (2005, November). Seven practices for effective learning. *Educational Leadership*, 10-17.
- Meisels, S., Atkins-Burke, S., Xue, Y., & Bickel, D.D. (2003). Creating a system of

- accountability: The impact of instructional assessment on elementary children's achievement scores. *Educational Policy Analysis Archives*, 11(9), 19.
- Mertler, C.A. & Vannatta, R.A. (2005). *Advanced and multivariate statistical methods: Practical application and interpretation*. Glendale, CA: Pyrczak Publishing.
- Nash, B. (2008). *Perceptions and use of a formative assessment system* (Dissertation). Retrieved from Dissertations & Theses: Full Text Database. (Publication No. AAT1450481)
- Nugent, C.J. (2009). *Predicting standardized state achievement scores from measured of academic progress* (Ed.D. Dissertation) Retrieved from Dissertations & Theses: Full Text Database. (Publication No. AAT3366127)
- Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues & Practice*, 28(3), 5-13.
- Pryor, J. & Croussouard, B. (2008). A socio-cultural theorization of formative assessment. *Oxford Review of Education*, 34(1), 1-20.
- Reeves, D.B. (2001). Standards make a difference: The influence of standards in classroom assessment. *NASSP Bulletin*, 85(5), 5- 12.
- Reeves, D.B. (2005). *Accountability in action: A blueprint for learning organizations* Englewood, CO: Advanced Learning Press.
- Reeves, D., Ainsworth, L., Almeida, L., Davies, A., DuFour, R., Gregg, L., Guskey, T., et.al., (2007). *Ahead of the curve: The power of assessment to transform teaching and learning*. Bloomington, IN: Solution Tree.
- Rodriguez, E.R. & Bellanca, J. (2007). *What is it about me you can't teach? An*

- instructional guide for the urban educator*. Thousand Oaks, CA: Corwin Press.
- Shute, V.J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(2), 153-189.
- Stiggins, R.J. (1999). Assessment, student confidence, and school success. *Phi-Delta Kappan*, 81(3), 191.
- Stiggins, R. (2002). Assessment crisis: the absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765.
- Stiggins, R. (2005) From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*. 87(4), 324-328.
- Stiggins, R. (2007, May). Assessment through the student's eyes. *Educational Leadership*, 22-26.
- Stiggins, R.J., Arter, J.A., Chappius, J., & Chappius, S. (2006). *Classroom assessment for student learning: Doing it right – using it well*. Portland, OR: Educational Testing Service.
- Stiggins, R. & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory Into Practice*, 44(1), 1-7.
- Strong, R.W., Silver, H.F., & Perini, M.J. (2001). *Teaching what matters most: Standards and strategies for raising student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tierney, R.D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education*. 13(3), 239-264.
- Wiggins, G. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco, CA: Jossey-Bass.

- Wiggins, G. (2007, September). *Schooling by design: Assessing what matters (not what is easy or feared)*. Paper presented at the annual meeting of the Georgia Association of Curriculum and Instructional Supervisors, Athens, GA.
- Wiggins, G. & McTighe, J. (2008). Put understanding first. *Educational Leadership*, 65(8), 36-41.
- Wiliam, D. (2006). Formative assessment: Getting the focus right. *Educational Assessment*, 11(3 &4), 283-289.
- Williams, T.R. (2009). *A correlation study of the Tennessee Formative Assessment Program (TFAP) to the Tennessee Comprehensive Assessment Program (TCAP) on fifth grade reading/language arts* (Ed.D. Dissertation). Retrieved from Dissertation & Theses:Full Text Database.(Publication No. AAT3376911)
- Willis. (1993). Learning and assessment: Exposing the inconsistencies of theory and practice. *Oxford Review of Education*, 19(3), 383-402.
- Zhao, Y. (2009). *Catching up or leading the way: American education in the age of globalization*. Alexandria, VA: Association for Supervision and Curriculum Development.



## Appendices

### Appendix A : Letter Inviting Participation to School Systems

Date

Name

Title

School System

Address

City, State, Zip

Dear Name,

I am currently a doctoral student at Liberty University in Lynchburg, Virginia and am working on my doctoral dissertation. My work is focused on a better understanding of formative assessment and the link between teacher participation in professional development, teacher perception of the use and benefits of formative assessment and standardized assessment scores. I am interested in studying this subject in relation to rural Georgia schools and students. In an effort to broaden my population, I would like to be able to include at least three different system's students and teachers in grades 9-12 courses which have an EOCT. I would need access to student standardized test scores from 2007-2008 and 2008-2009 and teacher survey data from spring 2009. I would not need any individually identifiable information such as student name or id number. I would like to have access to student demographic information such as age, gender, socio-economic status (if available), and disability (if applicable). I would need student data disaggregated by teacher for each year in order to compare student growth based on teacher assignment/perception. We can certainly work out a coding system to address identity protection on survey and test data.

I am hopeful that you will be willing to work with me and allow access to your student data and to provide time with your faculty for survey completion so I may better study this subject with a broad population. I will certainly be happy to provide you with a copy of my findings. I would welcome the opportunity to discuss my study further in detail as needed. Please respond via mail, e-mail or phone at your convenience. I look forward to hearing from you.

Sincerely,  
Marianne Cole

Marianne Cole  
Assistant Superintendent  
Heard County Schools  
PO Box 1330  
Franklin, GA 30217  
706-675-3320  
[mcole@heard.k12.ga.us](mailto:mcole@heard.k12.ga.us)

**Appendix B: Follow-up Letter to School Principals**

Date  
Name  
Title  
School System  
Address  
City, State, Zip

Dear Name,

I am currently a doctoral student at Liberty University in Lynchburg, Virginia and am working on my doctoral dissertation. My work is focused on a better understanding of formative assessment and the link between teacher participation in professional development, teacher perception of the use and benefits of formative assessment and standardized assessment scores. I am interested in studying this subject in relation to rural Georgia schools and students. In an effort to broaden my population, I would like to be able to include at least three different system's students and teachers in grades 9-12 courses which have an EOCT. I have previously contacted your system administration and they have graciously agreed to allow me to contact you to request your help. I would need access to student standardized test scores from 2007-2008 and 2008-2009 and teacher survey data from spring 2009. I would not need any individually identifiable information such as student name or id number. I would like to have access to student demographic information such as age, gender, socio-economic status (if available), and disability (if applicable). I would need student data disaggregated by teacher for each year in order to compare student growth based on teacher assignment/perception. We can certainly work out a coding system to address identity protection on survey and test data.

I am hopeful that you will be willing to work with me and allow access to your student data and to provide time with your faculty for survey completion so I may better study this subject with a broad population. I will certainly be happy to provide you with a copy of my findings. I would welcome the opportunity to discuss my study further in detail as needed. Please respond via mail, e-mail or phone at your convenience. I look forward to hearing from you.

Sincerely,

Marianne Cole

Marianne Cole  
Assistant Superintendent  
Heard County Schools  
PO Box 1330  
Franklin, GA 30217  
706-675-3320  
[mcole@heard.k12.ga.us](mailto:mcole@heard.k12.ga.us)

**Appendix C: Presentations for Professional Development**

**See Supplemental Attachment File for Application-Based Training Power Points -**

**October 2008 and January 2009** Reprinted with permission from West Georgia

Regional Education Services Agency Grantville, GA

**See Supplemental Attachment File for Theory-Based Training Handout – January**

**2009** Reprinted with permission from Tom Guskey, Ph.D. Georgetown College

Georgetown, KY

Additional Readings Provided as Part of Dr. Guskey’s Handouts

Guskey, T. R. (2007). The rest of the story. *Educational Leadership*. 65 (4), 28-35.

Guskey, T. R. (2006). Making high school grades meaningful. *Phi Delta Kappan*. 87 (9), 670-675.

Guskey, T. R. (2004). Are zeros your ultimate weapon. *Principal Leadership*. November 2004, 30-35.

Guskey, T. R. (2004). The communication challenge of standards- based reporting. *Phi Delta Kappan*. December 2004, 326-329.

Guskey, T. R. (2003). How classroom assessments improve learning. *Educational Leadership*. 60(5), 6-11.

Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership*. 59 (1), 20-27.

Guskey, T. R. (2000). Deciding what’s important to learn. *News & Notes*, Summer 2000, 3-7

Guskey, T. R. (2000). Grading policies that work against standards and how to fix them. *NASSP Bulletin*. 84 (620), 20-29.

Guskey, T. R. (2000). Twenty questions? Twenty tools for better teaching, *Principal Leadership*. 1 (3), 5-7.

Guskey, T. R. (1999, April 1). Inflation not the issue; focus on grades purpose. *Lexington Herald-Leader*. p. A19.

Guskey, T. R. (1999). Making standards work. *School Administrator*. October 1999,

Guskey, T. R. (1998, January 19). Good teachers can overcome effects of poverty on learning, Lexington *Herald-Leader*. p. A9.

Guskey, T.R. (1994). Making the Grade: What Benefits Students. *Educational Leadership*, 52 (2), 14-20.

**Appendix D- Bol Questionnaire****See Supplemental Attachment File for Original Questionnaire**

## Appendix E- Study Questionnaire

### Instructions

The following questions ask you to provide information about your experiences with student assessment and your feelings about a variety of assessment methods. The data you provide in this questionnaire will help in research conducted in an effort to better understand Georgia teacher's practices and perceptions related to assessment. Please take a few minutes of your time to respond carefully to each question. In some questions you are asked about your assessment practice **last year**. If you are a first-year teacher, respond in reference to what you did as a student teacher. *Your responses will be treated confidentially.*

### SECTION ONE

Check one box for each item. Please rate how frequently you used the following assessment methods in your classroom **last year**.

		1- Never		5- Frequently		
Close-ended exams, quizzes, or other assignments (e.g. multiple choice, matching, or true-false items)	1	2	3	4	5	
Open-ended exams or quizzes or other assignments (e.g., short answer or essay items)	1	2	3	4	5	
Written assignments (e.g. essays, term papers, reports, journals)		2	3	4	5	
Portfolio assessment (a collection of assignments, work samples)	1	2	3	4	5	
Observations (e.g. evaluating participation, group work)	1	2	3	4	5	
Performance Task (e.g. assessment of students as they work on a problem or task)		2	3	4	5	
Self-assessment by students	1	2	3	4	5	

Check one box for each item. **Last year**, to what extent did your assessment methods demand:

	1- Never			5 -Frequently	
Basic knowledge or comprehension of Information	1	2	3	4	5
Selection of important vs. unimportant Information	1	2	3	4	5

Integration of information from different Sources	1	2	3	4	5
Application of information	1	2	3	4	5
A focus on facts or details	1	2	3	4	5
A focus on terms or definitions	1	2	3	4	5
A focus on concepts or principles	1	2	3	4	5

## SECTION TWO

Check one box for each item. Please rate how **well prepared** you feel in developing and administering the following assessment methods.

	Not at all			5- Very	
Closed-ended exams, quizzes, or other assignments (e.g. multiple choice, matching, or true false items)	1	2	3	4	5
Open-ended exams or quizzes or other assignment (e.g. short answer or essay items)	1	2	3	4	5
Written assignments (e.g. essays, term papers, reports, journals)	2	3	4	5	
Portfolio assessment (a collection of assignments, work samples)	1	2	3	4	5
Observation (e.g. evaluating participation, group work)	1	2	3	4	5
Performance Task (e.g. assessment of students as they work on a problem or task)	1	2	3	4	5
Self-assessment by students	1	2	3	4	5

Check one box for each item. Please rate how confident you are that the following assessment methods **accurately reflect student achievement and progress**.

	Not at all			5- Very	
Close-ended exams, quizzes, or other assignments (e.g. multiple choice, matching, true-false items)	1	2	3	4	5
Open-ended exams or quizzes or other assignments (e.g. short answer or essay items)	1	2	3	4	5
Written assignments (e.g., essays, term papers, reports, journals)	1	2	3	4	5
Portfolio assessment (a collection of assignments, work samples)	1	2	3	4	5
Observation (e.g. evaluating participation, group work)	1	2	3	4	5
Performance task (e.g. assessment of students as they work on a problem or task)	1	2	3	4	5
Self- assessment by students	1	2	3	4	5

### SECTION THREE

Please comment on the training experiences you have had pertaining to assessment in terms of quality, usefulness, adequacy, etc.

So far, what are you doing differently in your school as a result of the assessment training?

At this point, what are your feelings about formative assessment in terms of its likely effect upon students, teachers, and/or parents/guardians at your school?



On a scale of 1 to 10, with 1 being very low and 10 being very high, please rate each of the following items.

- \_\_\_\_\_ Overall quality of the training conducted related to assessment
- \_\_\_\_\_ Adequacy of training to prepare you to implement assessment strategies
- \_\_\_\_\_ School/District support for implementation
- \_\_\_\_\_ Enthusiasm of teachers in your school for implementing assessment strategies

#### DEMOGRAPHIC INFORMATION

Please provide the following demographic information.

Years of teaching experience: \_\_\_\_\_

Grade level(s) you teach: \_\_\_\_\_

Subject areas(s) you teach: (Check all that apply)

- \_\_\_\_\_ Mathematics
- \_\_\_\_\_ Social Science
- \_\_\_\_\_ Fine Arts
- \_\_\_\_\_ Science
- \_\_\_\_\_ English/Language Arts
- \_\_\_\_\_ Physical Education
- \_\_\_\_\_ Other (Please specify: \_\_\_\_\_)

School: \_\_\_\_\_

Teacher ID: \_\_\_\_\_

Assessment In-service participated in this year: (check all that apply)

- \_\_\_\_\_ October 2008 and January 2009 (RESA)
- \_\_\_\_\_ January 2009 (Guskey)
- \_\_\_\_\_ Other (Please specify: \_\_\_\_\_)

THANK YOU FOR COMPLETING THE QUESTIONNAIRE!

## Appendix F – IRB Application

11/06Ref. # \_\_\_\_\_

### APPLICATION TO USE HUMAN RESEARCH SUBJECTS Liberty University Committee On The Use of Human Research Subjects

1. Project Title: An Analysis of the Influence of Assessment for Learning Professional Learning Models in Rural Georgia Public Schools
2. Full Review  Expedited Review  X
3. Anticipated Funding Source: Self
4. Principal Investigator:  
Marianne Cole, Doctoral Candidate  
GA 30217  
770-328-6217, [mwcole@liberty.edu](mailto:mwcole@liberty.edu)  
2460 Armstrong Mill Rd. Franklin,
5. Faculty Sponsor (if student is PI), also list co-investigators below Faculty Sponsor, and key personnel:  
Dr. Jeff Crawford  
Professor, Dissertation Chair  
Education, [jcrawford@liberty.edu](mailto:jcrawford@liberty.edu)
6. Non-key personnel:  
Name and Title  
Dept, Phone, E-mail address
7. Consultants:  
Name and Title  
Dept., Phone, E-mail address
8. The principal investigator agrees to carry out the proposed project as stated in the application and to promptly report to the Human Subjects Committee any proposed changes and/or unanticipated problems involving risks to subjects or others participating in approved project in accordance with the Liberty Way and the [Confidentiality Statement](#). The principal investigator has access to copies of [45 CFR 46](#) and the [Belmont Report](#). The principal investigator agrees to inform the Human Subjects Committee and complete all necessary reports should the principal investigator terminate University association. Additionally s/he agrees to maintain records and keep informed consent documents for three years after completion of the project even if the principal investigator terminates association with the University.

\_\_\_\_\_  
Principal Investigator Signature

\_\_\_\_\_  
Date

\_\_\_\_\_  
Faculty Sponsor (If applicable)

\_\_\_\_\_  
Date

**Submit the original request to: Human Subjects Office, Liberty University, 1971 University Blvd., IRB Chair, Suite 2400 CN, Lynchburg, VA 24502**

## APPLICATION TO USE HUMAN RESEARCH SUBJECTS

10. This project will be conducted at the following location(s): (please indicate city & state)
- Liberty University Campus
- Other (Specify): Harris County High School Hamilton, GA, Heard County High School Franklin, GA, and Manchester High School Manchester, GA.
- 
11. This project will involve the following subject types: (check-mark types to be studied)
- |   |   |
|---|---|
| <input checked="" type="checkbox"/> Normal Volunteers (Age 18-65)<br><input type="checkbox"/> In Patients<br><input type="checkbox"/> Out Patients<br><input type="checkbox"/> Patient Controls<br><input type="checkbox"/> Fetuses<br><input type="checkbox"/> Cognitively Disabled<br><input type="checkbox"/> Physically Disabled _____<br><input type="checkbox"/> Pregnant Women | <input type="checkbox"/> Subjects Incapable Of Giving Consent<br><input type="checkbox"/> Prisoners Or Institutionalized Individuals<br><input checked="" type="checkbox"/> Minors (Under Age 18)<br><input type="checkbox"/> Over Age 65<br><input type="checkbox"/> University Students (PSYC Dept . subject pool ____)<br><input type="checkbox"/> Other Potentially Elevated Risk Populations _____ |
|---|---|
12. Estimated number of subjects to be enrolled in this protocol: 3,000
13. Does this project call for: (check-mark all that apply to this study)
- Use of Voice, Video, Digital, or Image Recordings?
- Subject Compensation? Patients \$\_\_\_\_\_ Volunteers \$\_\_\_\_\_
- Participant Payment Disclosure Form
- |   |   |
|---|---|
| <input type="checkbox"/> Advertising For Subjects?<br><input type="checkbox"/> More Than Minimal Psychological Stress?<br><input type="checkbox"/> Confidential Material (questionnaires, photos, etc.)?<br><input type="checkbox"/> Extra Costs To The Subjects (tests, hospitalization, etc.)?<br><input type="checkbox"/> The Exclusion of Pregnant Women?<br><input type="checkbox"/> The Use of Blood? | <input type="checkbox"/> More Than Minimal Risk?<br><input type="checkbox"/> Alcohol Consumption?<br><input type="checkbox"/> Waiver of Informed Consent?<br><input type="checkbox"/> VO2 Max Exercise?<br>Total Amount of Blood _____<br>Over Time Period (days) _____ |
|---|---|
- The Use of rDNA or Biohazardous materials?
- The Use of Human Tissue or Cell Lines?
- The Use of Other Fluids that Could Mask the Presence of Blood (Including Urine and Feces)?
- The Use of Protected Health Information (Obtained from Healthcare Practitioners or Institutions)?
14. This project involves the use of an **Investigational New Drug (IND)** or an **Approved Drug For An Unapproved Use**.
- YES       NO
- Drug name, IND number and company:
15. This project involves the use of an **Investigational Medical Device** or an **Approved Medical Device For An Unapproved Use**.
- YES       XNO
- Device name, IDE number and company:
16. The project involves the use of **Radiation or Radioisotopes**:
- YES       NO
17. Does investigator or key personnel have a potential conflict of interest in this study?
- YES       NO

## EXPEDITED/FULL REVIEW APPLICATION NARRATIVE

### A. PROPOSED RESEARCH RATIONALE

Over the last 100 years research has consistently shown that effective formative assessment techniques, used to adjust instruction and provide student feedback, can and do improve student achievement and learning. Despite this, research continues to show teachers are not implementing effective formative assessment techniques. Is this a perception problem? Is this a professional development problem? Is this a lack of knowledge problem? Is this a lack of understanding problem? Is this a lack of caring problem?

This researcher believes that should educators understand the benefits of effective formative assessment, know how to implement effective formative assessment in their classroom, and understand the impact their perception has on the success of formative assessment techniques that most educators would work to implement effective formative assessment in their classrooms. The focus of this research is on providing data to support this assertion.

### B. SPECIFIC PROCEDURES TO BE FOLLOWED

All students will receive instruction using the same curriculum based on Georgia Performance Standards. Standards, elements, performance tasks, exemplars, and curriculum guides will be available to all teachers from the Georgia Department of Education at [www.georgiastandards.org](http://www.georgiastandards.org).

The intervention focus is on professional learning concerning assessment strategies to promote assessment for learning. Some teachers will be provided application based professional learning from a Regional Educational Services Agency incorporating theory related to effective formative assessment practices. Some teachers will be provided theory based professional development from a national expert on assessment strategies. Some teachers will not participate in any professional development related to assessment strategies and will function as the control group. Classroom instruction techniques may be adjusted based on assessment strategies learned in the professional development.

A survey will be used in the Spring of 2009 designed to gather data on perception of use of assessment in the classroom. The survey used is adapted from work completed by Bol, Stephenson, and O'Connell (1998). See Appendix D for Bol survey. See Appendix E for study survey. The survey is adapted and used with permission from Dr. Bol. The original survey was used to gather information on the influence of teaching experience, grade level, and subject area on assessment practice. The usefulness of the survey to the current study relates to the survey data gathered on previous year use of assessments (summative and formative), teacher preparedness for using differing assessment methods, and teacher perception of the usefulness of particular assessment methods (summative/traditional and formative) in determinations of student learning and progress.

The survey is divided into several sections. Section One will provide background data on teacher prior knowledge of assessment strategies. Section Two will provide data on teacher perception of the benefit of formative assessment techniques. Demographic data will provide information to identify teacher participation in specific professional development or membership in the control group and to match survey data to End of Course Test data for specific classes.

Section Two of the survey will be scored numerically and two scale scores will be identified (one for summative/traditional assessment and one for formative assessment). The scale score will be calculated as the mean range obtained across the items comprising each scale. A high scale score will indicate positive perception and a low scale score will indicate a negative perception. Bol et.al. found the reliability coefficient for the survey at .49 for the traditional assessment questions and .75 for the formative assessment questions (Bol, 1998). For the purpose of this study the formative

assessment questions will be examined and yield a high reliability. Additionally, Bol (1998) indicated construct validity of the scales supported through factor analysis.

Student achievement data will be gathered from school and system reports provided by the Georgia Department of Education and individual school systems for the standardized End of Course Test for 9th Grade Literature, American Literature, Geometry, Physical Science, Biology, United States History, and Economics for the 2007-2008 and 2008-2009 school years. Individual students are given a numerical score on a 100 point scale. The test is developed by Pearson Educational Measurement and vetted for validity and reliability through the national testing company's rigorous standards.

Teachers in the intervention group will participate in professional development related to effective use of formative assessment in the Fall/Winter of 2008 by participation in direct in-service instruction that is application based in October 2008 and January 2009 or direct in-service instruction that is theory based in January 2009. See Appendix C for in-service information. Teachers will indicate the type and date of in-service participation as part of the demographic information gathered as part of the survey. Teachers may have some prior knowledge or have previously implemented some formative assessment techniques as a result of independent study therefore teacher prior knowledge information will be gathered as part of the survey. Classroom implementation of Assessment for Learning strategies will be indicated as part of the teacher survey as well.

Teacher surveys will be administered in Spring 2009. Surveys will be administered as part of a scheduled faculty meeting at each school. Teachers will be assured of anonymity and confidentiality as their data will only be seen by the researcher. Teachers will be encouraged to be candid in their responses.

Student Achievement data will be gathered in Spring 2009 by request to the school leaders or their designee. Data gathered will include Standardized End of Course Test Data for 9<sup>th</sup> Grade Literature, American Literature, Geometry, Physical Science, Biology, US History, and Economics for both the 2007-2008 and 2008-2009 school year. Data will be gathered and indicated as belonging to School A Teacher 1, School A Teacher 2, School B Teacher 1, School B Teacher 2, etc. Each Teacher will be assigned a letter code to represent themselves by the school leader and will use this code for teacher perception/survey data as well.

### **C. SUBJECTS TO BE INCLUDED**

The population for the study is ninth, tenth, eleventh, and twelfth grade teachers and students in rural Georgia public high school. Ninth, Tenth, Eleventh, and Twelfth Grade teachers and students in core content areas will participate in the study. For the purposes of this study three high schools in rural Georgia will be identified to participate. Teachers will participate in application based professional development on formative assessment, theory based professional development on formative assessment, or not participate in targeted professional development (the control group). Teachers within the school who have provided instruction in core content areas of 9<sup>th</sup> Grade Literature, American Literature, Geometry, Physical Science, Biology, US History, and Economics during the 2007-2008 school year and are providing instruction in the same content area for the 2008-2009 school year will be identified to participate in the study. The student sample will consists of 3,216 students with about 49% male and 51% female. Economically disadvantaged students will account for approximately 46% of the population. Students with disabilities will account for approximately 12 % of the population. Student ethnicity is approximately 63% white, 35% African American, 1% Hispanic, and 1% multiracial. English Language Learners will account for less than 1% of the population. Students will be heterogeneously grouped throughout classrooms and will be of mixed academic abilities based on standard class assignment procedures for each school. Only students participating in classes of teachers teaching 9<sup>th</sup> Grade Literature, American Literature, Geometry, Physical Science, Biology, US History, and Economics during the 2007-2008 and 2008-2009 school year will be invited to participate.

**D. RECRUITMENT OF SUBJECTS AND OBTAINING INFORMED CONSENT**

Once the sample is identified the first step will be to contact the school systems which meeting population requirements as rural Georgia schools. For the purpose of anonymity, the participating schools will be labeled School A, School B, and School C. School system administration will be contacted in Fall 2008 and invited to participate in the study. Once school systems have agreed to participate, each school will be contacted individually and the researcher will meet with the school leader to discuss the research in the Fall of 2008. The researcher will provide each school leader with a brief overview of the purpose of the study. Intervention schools and participating teachers and students will participate in the intervention, complete surveys, and complete state required standardized End of Course Tests during the 2008-2009 school year.

**E. PROCEDURES FOR PAYMENT OF SUBJECTS**

No payment will be included.

**F. CONFIDENTIALITY**

As described above teacher data will be identified as School A teacher 1, etc. Student data will be identified by a sequential numbering code with the only identifiable information being a link to School A Teacher 1, etc.

All data collected in the study will be kept under lock and key in a filing cabinet in the researcher's office which is located in a secure office building. Any data with original identifying information or referencing which school data is obtained from will be kept in a separate file cabinet under lock and key in the same office.

Data records will be kept on file for 3 years.

**G. POTENTIAL RISKS TO SUBJECTS**

The risk associated with participation in this study is minimal and no more than that anticipated in daily activity.

**H. BENEFITS TO BE GAINED BY THE INDIVIDUAL AND/OR SOCIETY**

Students participating in this study may obtain academic benefit due to the increased knowledge and expertise in assessment strategies of their participating teacher. Teachers participating in this study may gain an array of improved professional practice due to participation in the professional development activities associated with this study.

The overall benefit to be gained by society is related to the focus of this research. The focus of this research is to further examine the link between professional development and the impact teacher perception concerning assessment may have on student growth in an effort to answer the call of assessment "gurus" in this nation as they call for a "redirection of assessment to its fundamental purpose: the improvement of student achievement, teaching practice, and leadership decision-making" (Reeves, et.al., 2007, p.1). Additionally, sound assessment practice should provide stakeholders (students, parents, teachers, and supervisors) with information about how the student is doing by providing students with an opportunity to improve achievement and keeping an individual record of student achievement of standards (Reeves, 2005). Currently, however, a class often functions as follows, the teacher teaches then tests then moves on leaving unsuccessful students to finish last founded on the premise that comparing unfavorable to others will motivate students to perform better in the future (Chappuis & Stiggins, 2002). On the contrary, assessment for learning occurs during the teaching and learning process providing students feedback and the time and ability to self correct and receive additional support for mastery of the learning goal (Chappuis & Stiggins). In assessment for learning teachers and students use formative assessment information to pretest and adjust instruction for individuals, analyze who needs more practice, revise

instruction continually, reflect on effectiveness of teaching practices, confer with students concerning strengths and areas for improvement, and facilitate peer tutoring (Chappuis & Stiggins). The state of Georgia and particularly the West Georgia region are currently implementing standards based classroom practices which include training and implementation of standards based formative assessment within targeted classrooms. Therefore, the population is the ideal target for measuring the gains of student achievement and the correlation to teacher implementation of formative assessment practices within the classroom. This area at this time is uniquely suited to provide an ideal environment within which to examine test scores from previous years which were not influenced by teachers implementing precise, thoughtful formative assessment techniques supported by detailed professional learning.

- I. **INVESTIGATOR'S EVALUATION OF THE RISK-BENEFIT RATIO**  
**As risk is minimal and not above that associated with daily activity the risk to benefit ratio is heavily in favor of the benefit.**
- J. **WRITTEN INFORMED CONSENT FORM** (*see attached*)
- K. **WAIVER OF INFORMED CONSENT OR SIGNED CONSENT:** No waiver is requested.
- L. **SUPPORTING DOCUMENTS** (*no supporting documentation is attached*)
- M. **COPIES:** 4 copies are included.

## CONSENT FORM

### An Analysis of the Influence of Assessment for Learning Professional Learning Models in Rural Georgia Public Schools

Marianne W. Cole

Liberty University

Education Department

You are invited to be in a research study to investigate the effect of two models of professional development concerning Assessment for Learning on teacher perception of the effectiveness of Assessment for Learning strategies and student achievement as measured by standardized End of Course Tests. You were selected as a possible participant because you are a teacher in grades 9-12 in a rural Georgia school. Additionally in some cases you have received training on Assessment for Learning in either a theory based or application based professional development class or you will be part of the control group not participating in this type of training. We ask that you read this form and ask any questions you may have before agreeing to be in the study.

This study is being conducted by: Marianne W. Cole, doctoral candidate, Education Department

#### **Background Information**

The purpose of this study is study two specific questions will gather and analyze data to further examine these questions.

1. Is there a correlation between teacher perception of the benefit of formative assessment and student achievement as measured by the Georgia End of Course Tests standardized assessment?
2. Is there a correlation between teacher participation in Assessment for Learning professional development and teacher perception of the benefits of Assessment for Learning strategies?

#### **Procedures:**

If you agree to be in this study, we would ask you to do the following things:

1. Participate in training provided by your school or system related to Assessment for Learning.
2. Administer all standardized End of Course Tests as directed by your school or system.
3. Complete a brief survey and answer honestly and confidentially.

#### **Risks and Benefits of being in the Study**

Risks of participating in this study are no more than the participant would encounter in everyday life.

No individual benefits from this study are predicted to occur. However, information gained may be beneficial and informative to the educational profession as a whole.

#### **Confidentiality:**



The records of this study will be kept private. In any sort of report we might publish, we will not include any information that will make it possible to identify a subject. Research records will be stored securely and only researchers will have access to the records.

Information and data gathered in the course of this study will only be accessed by the researcher and will be kept under lock and key in the researcher's office.

**Voluntary Nature of the Study:**

Participation in this study is voluntary. Your decision whether or not to participate will not affect your current or future relations with the Liberty University or with your school or system. If you decide to participate, you are free to not answer any question or withdraw at any time without affecting those relationships.

**Contacts and Questions:**

The researcher conducting this study is Marianne Cole. You may ask any questions you have now. If you have questions later, **you are encouraged** to contact her at 770-328-6217, [mwcole@liberty.edu](mailto:mwcole@liberty.edu). The faculty advisor for this research is Jeff Crawford. You may contact him at [jcrawford@liberty.edu](mailto:jcrawford@liberty.edu).

If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher(s), **you are encouraged** to contact the Institutional Review Board, Dr. Fernando Garzon, Chair, 1971 University Blvd, Suite 2400, Lynchburg, VA 24502 or email at [fgarzon@liberty.edu](mailto:fgarzon@liberty.edu).

***You will be given a copy of this information to keep for your records.***

**Statement of Consent:**

I have read the above information. I have asked questions and have received answers. I consent to participate in the study.

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

Signature of parent or guardian: \_\_\_\_\_ Date: \_\_\_\_\_  
(If minors are involved)

Signature of Investigator: \_\_\_\_\_ Date: \_\_\_\_\_

IRB Approval 699.030809: An Analysis of the Influence of Assessment for Learning Prof... Page 1 of 1

**IRB Approval 699.030809: An Analysis of the Influence of Assessment for Learning Professional Learning Models in Rural Georgia Public Schools**

Institution Review Board

**Sent:** Thursday, April 16, 2009 9:45 AM  
**To:** Cole, Marianne W; Crawford, Jeffrey; Garzon, Fernando L.  
**Cc:** "Institution Review Board."  
**Attachments:** Annual Review Form.doc (29 KB) ; Change in Protocol.doc (29 KB)

---

Dear Marianne,

We are pleased to inform you that your above study has been approved by the Liberty IRB. This approval is extended to you for one year. If data collection proceeds past one year, or if you make changes in the methodology as it pertains to human subjects, you must submit an appropriate update form to the IRB. Attached you'll find the forms for those cases.

Thank you for your cooperation with the IRB and we wish you well with your research project. We will be glad to send you a written memo from the Liberty IRB, as needed, upon request.

Sincerely,

Fernando Garzon, Ph.D.  
 IRB Chair, Liberty University  
 Center for Counseling and Family Studies Liberty University  
 1971 University Boulevard  
 Lynchburg, VA 24502-2269  
 (434) 597-4054  
 Fax: (434) 522-0477

----- End of Forwarded Message

**Appendix G- All Data Table**  
**See Supplemental Attachment File for All data Tables**