ABSTRACT

MCMANUS, SARAH MCMILLIAN. A Study of Formative Assessment and High Stakes Testing: Issues of Student Efficacy and Teacher Views in the Mathematics Classroom. (Under the direction of Dr. Lee V. Stiff).

In the age of high stakes testing, research shows that the focus on summative assessment is overshadowing the importance and use of formative assessment. According to Black and Wiliam (2005), interim assessments coupled with large-scale assessments represent substantial and well-entrenched obstacles that prevent the growth of formative assessment use in the United States. Exploring ways to overcome these obstacles, this research introduced teachers to the concept of a comprehensive balanced assessment system that values accountability statewide assessment (summative), interim/benchmark assessment (summative), and classroom assessment (summative and formative).

The methodology of the study involved four teachers from two high schools in North Carolina who participated in a 16-week Formative Assessment (FA) Project led by the researcher. The study included professional development consisting of four modules, application in practice, and participation in a learning community. Formative assessment, in short, was defined as a process used during instruction to adjust teaching and learning.

A grounded theory approach was employed to determine (1) what themes emerge as teachers implemented formative assessment in the age of high stakes testing; and (2) what new behaviors and/or attitudes students exhibit as a result of the
implementation of formative assessment. Data was collected using field observations, interviews, and artifact reviews.

The data supported the conclusion that teacher’s views about assessment were changed to become more inclusive of students as partners in the assessment process. Students’ self-efficacy increased as evidenced by their increased commitments to the learning process, use of metacognitive strategies, and levels of engagement. Based on the six themes that emerged from the data, the researcher identified four steps in the formative assessment process and three essential elements that need to be present for successful implementation in high school mathematics classrooms.
A Study of Formative Assessment and High Stakes Testing: Issues of Student Efficacy and Teacher Views in the Mathematics Classroom

by
Sarah McMillian McManus

A dissertation submitted to the Graduate Faculty of North Carolina State University
In partial fulfillment of the Requirements for the degree of Doctor of Philosophy

Mathematics Education

Raleigh, North Carolina

2008

APPROVED BY:

Dr. Lee V. Stiff
Chair of Advisory Committee

Dr. Karen F. Hollebrands

Dr. Sarah B. Berenson

Dr. Thomas Gerig
DEDICATION

This dissertation could not have been completed without Jesus Christ, my Lord and Savior, who ordered my steps. It is dedicated to Dr. Cecil H. McManus, my husband, and to my two children, Briana Callie McManus and Britney Ann McManus. Their love and support sustained me throughout this process. With deep appreciation, I thank them for their willingness to sacrifice time and attention to accommodate my professional and educational goals. This is a victory for all of us!
ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge and thank my mother, Sarah Frances McMillian, my father, Alfonso McMillian, and my five siblings – LaJune Thomas, Alfonso McMillian, Jr., Gwendolyn Reilly, Michael McMillian, and Stephanie McMillian – who all influenced my love of education, my dedication to life-long learning, my pursuit of excellence, and most of all my persistence.

I would like to acknowledge North Carolina State University for providing a rigorous and challenging doctoral program. In particular, I would like to thank my committee members, Dr. Lee V. Stiff, Dr. Karen F. Hollebrands, Dr. Sarah B. Berenson, and Dr. Thomas Gerig for supporting me throughout this process and for providing me a wonderful education. You helped me to think more critically, understand the importance of conceptual frameworks, and understand research principles.

I would like to thank the North Carolina Department of Public Instruction, under the direction of State Superintendent Dr. June Atkinson, for supporting me in my research and agreeing to participate in the Enhanced Assessment Grant (EAG) that was used to support parts of this research. I would like to thank the North Carolina EAG project team, for their assistance and support in this project. In addition, I would like to thank the schools and districts that agreed to participate in the EAG. The teachers, students, school administrative staff, and district leaders were great to work with throughout my research.
I would like to thank the Council of Chief State School Officers (CCSSO) for selecting me as a member of the CCSSO Formative Assessment Advisory Group, coordinating the EAG, and creating the Formative Assessment for Students and Teachers State Collaborative on Assessments and Student Standards. By working on the formative assessment initiative through these various groups, I had increased motivation and desire to complete my dissertation.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER 1 – INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Need for a Comprehensive Balanced Assessment System</td>
<td>9</td>
</tr>
<tr>
<td>Need for Formative Assessment</td>
<td>13</td>
</tr>
<tr>
<td>Need for Professional Development</td>
<td>16</td>
</tr>
<tr>
<td>CHAPTER 2 – LITERATURE REVIEW</td>
<td>18</td>
</tr>
<tr>
<td>Definitions</td>
<td>18</td>
</tr>
<tr>
<td>Summative Assessment</td>
<td>19</td>
</tr>
<tr>
<td>Formative Assessment</td>
<td>19</td>
</tr>
<tr>
<td>Comparison of Formative and Summative Assessment</td>
<td>22</td>
</tr>
<tr>
<td>Literature Related to Formative Assessment</td>
<td>24</td>
</tr>
<tr>
<td>Clear and Appropriate Learning Targets</td>
<td>24</td>
</tr>
<tr>
<td>Clear Criteria for Success</td>
<td>25</td>
</tr>
<tr>
<td>Rubrics</td>
<td>27</td>
</tr>
<tr>
<td>Self-regulation</td>
<td>28</td>
</tr>
<tr>
<td>Use of Descriptive vs. Evaluative Feedback</td>
<td>30</td>
</tr>
<tr>
<td>Self-Efficacy</td>
<td>33</td>
</tr>
<tr>
<td>Formative Assessment and Achievement on Summative Tests</td>
<td>35</td>
</tr>
<tr>
<td>Teachers as Assessors</td>
<td>36</td>
</tr>
<tr>
<td>Social-Constructivist Perspective</td>
<td>41</td>
</tr>
<tr>
<td>Summary</td>
<td>46</td>
</tr>
<tr>
<td>Research Questions</td>
<td>46</td>
</tr>
<tr>
<td>CHAPTER 3 – METHODOLOGY</td>
<td>48</td>
</tr>
<tr>
<td>Participants</td>
<td>48</td>
</tr>
<tr>
<td>Training Protocol</td>
<td>50</td>
</tr>
<tr>
<td>Module 1</td>
<td>50</td>
</tr>
<tr>
<td>Module 2</td>
<td>50</td>
</tr>
<tr>
<td>Module 3</td>
<td>51</td>
</tr>
<tr>
<td>Module 4</td>
<td>51</td>
</tr>
<tr>
<td>Instruments</td>
<td>51</td>
</tr>
<tr>
<td>Procedure</td>
<td>52</td>
</tr>
<tr>
<td>CHAPTER 4 – DATA ANALYSIS AND RESULTS</td>
<td>54</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>54</td>
</tr>
<tr>
<td>Training Module1: Overview</td>
<td>54</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1: Differences Between Formative and Summative Assessment..........................23

Table 2: Impact of Feedback on Self........................................................................34
LIST OF FIGURES

Figure 1: Items from TIMSS .................................................................7
Figure 2: A Comprehensive Balanced Assessment System ............................10
Figure 3: Stiggins’ Assessment for Learning Illustration .................................13
Figure 4: Bloom’s Learning for Mastery Model ............................................20
Figure 5: Heritage Model of Feedback .....................................................39
Figure 6: Participant Responses to “What is assessment?” ...........................54
Figure 7: Rubric Displayed During Observation 2 .........................................61
Figure 8: Formative Assessment Implemented During Observation 4 ............64
CHAPTER 1

Introduction

The social science version of the Heisenberg Uncertainty Principle states “the more important that any quantitative social indicator becomes in social decision making the more likely it will be to distort or corrupt the social process it is intended to monitor” (Amrein & Berliner, 2002, p. 5). Does this principle apply to the current use of large-scale assessment in K-12 education? Large-scale assessments are those assessments administered under uniform conditions that provide scores that can be used to compare students across classrooms (Stiggins, 2001). These standardized assessments assign students to achievement levels that are subsequently used as mechanisms for conveying information about student progress at set times (Harlen & Winter, 2004). Throughout history, large-scale assessments have been used as objective measures to monitor achievement. However, the United States experienced a proliferation of large scale assessments beginning in the 1980s (Amrein & Berliner, 2002).

In the 1980’s there was growing concern over the level of achievement of American students. *A Nation At Risk*, a report detailing the state of education in the United States, was released in 1983 by the National Commission on Excellence in Education (1983). The report stated that in the American educational system there was a “rise in mediocrity” that threatened the future of both the people and the
nation (p. 1). This report recommended the use of state and local standardized tests to evaluate student progress.

In response to this recommendation, many states implemented or expanded their use of large-scale assessments to monitor the achievement of their students. Policymakers (i.e. legislators, governors, and boards of education) used large-scale assessments to influence and monitor progress in schools and classrooms. Important decisions were based on the results from large-scale assessments in an effort to encourage teachers to use effective instructional practices that would increase test scores.

North Carolina, like many other states, developed policies that provided teacher bonuses based on student test results, labeled schools based on student performances (e.g., School of Excellence, Low-Performing School), and promotion or graduation decisions based on test results. Consequences, or high stakes, were used to hold school districts, schools, teachers, and students accountable for the test results. For example, penalties or sanctions were applied to those schools or school districts that failed to meet established targets.

Having large-scale assessment data was beneficial to states when evaluating curriculum and instructional programs and for identifying those students and schools in the greatest need of improvement (DePascale, 2003; Heubert & Hauser, 1999; McGehee & Griffith, 2001; McMillan, 2001; Popham, 2000; Shepard, 2000; Shepard, Hammerness, Darling-Hammond, & Rust, 2005; Stiggins, 2004; Tindal,
2002). Some found that when used appropriately large-scale assessments led to improved teaching, better student learning, and equality in educational opportunities (Heubert & Hauser, 1999). However, there were concerns that there was too much focus on the use of these standardized measures (Black & Wiliam, 1998; Bright & Joyner, 1998; Brookhart, 1997; McGehee & Griffith, 2001; McMillan, 2001; Popham, 2000; Shepard, 2000; Stigler & Hiebert, 1997) and student achievement was not improving.

Stigler and Hiebert (1997) argued that, “A focus on standards and accountability that ignores the processes of teaching and learning in classrooms will not provide the direction that teachers need in their quest to improve” (p. 19). Black and Wiliam (1998) further expounded on this concern in their analogy of the classroom being treated as a “black box” (p. 140) where inputs — tests with high stakes — are imposed upon it and certain outputs — high test scores — are expected without regards to what happens inside the box. Black and Wiliam studied classroom practices about which they wrote:

….standards can be raised only by changes that are put into direct effect by teachers and pupils in the classrooms. There is a body of firm evidence that formative assessment is an essential component of classroom work and that its development can raise standards of achievement.
We know of no other way of raising standards for which such a strong prima facie case can be made. Our plea is that national and state policy makers will grasp this opportunity and take the lead in this direction (p.12).

Nearly two decades after the release of “A Nation at Risk,” the concerns over the educational system still existed. One of the key concerns was the mathematics performance of students. The mathematical abilities of American students had been compared to those of students in other countries using the Third International Mathematics and Science Study (TIMSS) and Programme for International Student Assessment (PISA). Nationally, the Nation's Report Card compared students' mathematical understanding using the National Assessment of Educational Progress (NAEP). Based on the results of these international and national measures, the United States Department of Education (USED) determined that American schools were not producing the mathematics excellence needed to be successful in the global economy of the 21st Century (USED, 2002).

In January 2002, the No Child Left Behind Act (NCLB) was signed into law. The purpose of NCLB was to improve student learning by ensuring that all students could meet grade-level proficiency standards by the year 2010. In the area of mathematics, this federal law required (1) all states to have statewide mathematics curriculum standards, (2) all students to have a highly quality teacher
certified in the area of mathematics teaching, (3) mathematics teachers to use effective teaching practices to promote learning, (4) student progress on state standards to be measured each year in grades 3 – 8 and at the high school level, and (5) schools and school districts to be held accountable for ensuring that students were making adequate yearly progress.

High stakes were associated with the testing requirements of NCLB. States failing to have approved testing and accountability programs would have funding withheld. Schools and school districts failing to meet adequate yearly progress would face sanctions and penalties such as giving parents the right to transfer their students out of the school at the expense of the school district and entering into district-wide improvement planning. With the passage of NCLB and the proliferation of large-scale assessments across the nation, more intense criticism of the use of these standardized tests followed. Critics argued that large scale assessments were being overused and used inappropriately resulting in negative impacts on education and children (Amrein & Berliner, 2002; DePascale, 2003; Heritage, 2007; Rodriguez, 2004; Shepard; 2000; Stiggins, 2004; Tindal, 2002; Wiliam, 2005; Wilson, 2005).

Although the goal of NCLB was to improve student learning, the high stakes nature of assessment resulted in too much emphasis placed on the product — student test scores — and not enough emphasis placed on the process — effective classroom teaching and assessment practices. Researchers (Heritage, 2007; Shepard...

Researchers began to look at classroom practices and concluded that the high stakes associated with large-scale assessments were having a negative impact on the choice of instructional and assessment strategies used in classrooms (Amrein & Berliner, 2002; Battista, 1999; NCTM, 2000; Popham, 2000; Volante, 2004; Wilson, 2005). In order for large-scale assessments to be valid measures, teachers must use effective classroom instructional strategies to help students learn and avoid “teaching to the test” (Volante, 2004). However, researchers concluded that teachers focused on behaviors that they felt contributed to outcomes for which they were held accountable, that is, testing. (Amrein & Berliner, 2002; Battista, 1999; NCTM, 2000; Popham, 2000; Shepard, 2000; Wilson, 2005).

When high stakes were involved, teachers allocated too much time to teaching test-taking skills at the expense of effective instructional practices, taught only those objectives that were measured on the test (Heubert & Hauser, 1999), and emphasized the teaching of low-level skills. This is captured by the quote, “I’d love to teach deep understanding, but I have to raise my students’ test scores” (Wiliam, 2005, p. 20). When teachers failed to use effective instructional strategies, student learning was negatively impacted (Black & Wiliam, 1998; Shepard 2005; Stiggins, 2004; Volante, 2004; Wiliam, 2005).
Studies showed that students who engaged in rote learning of specific test types did not develop deep understanding and were not able to transfer knowledge from one context to another (Koretz, Linn, Dunbar, & Shepard, 1991). Adding to this concern was the widespread use of multiple-choice items on large-scale assessments. The use of multiple-choice items was the most cost effective and efficient way to gather information on large-scale assessments (Shepard, 2005). However, critics argued that multiple-choice items were ineffective in measuring conceptual understanding, tended to focus on the low-level skills, measured isolated facts rather than higher-order thinking skills, and ignored students’ misconceptions. Wiliam (2005) used the following example to illustrates some concerns.

<table>
<thead>
<tr>
<th>Item 1</th>
<th>What fraction is the smallest?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) 1/6</td>
<td>b) 2/3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item 2</th>
<th>Which fraction is the largest?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) 4/5</td>
<td>b) 3/4</td>
</tr>
</tbody>
</table>

*Figure 1: Items from TIMSS*

If students focus only on the denominator and have the misconception that the fraction with the smallest denominator is the largest fraction and vice-versa, they may answer Item 1 correctly but with the wrong reasoning. When both items were used on the TIMSS the correct response rate for Item 1 was 88% compared to 46% for Item 2 (Wiliam, 2005).
At the turn of the 21st century, policymakers were again urged to focus less on large-scale statewide assessments such as end-of-grade and end-of-course tests and focus more on what was inside the black box (Amrein & Berliner, 2002; Chudowsky & Pellegrino, 2003; DePascale, 2003; Heritage, 2007; Rodriguez, 2004; Stiggins, 2004; Tindal, 2002; Wiliam, 2005; Wilson, 2005). However, as policymakers began turning their attention to the work of Black and Wiliam (1998), another problem emerged. According to Shepard (2005):

Unfortunately, the arrival of formative assessment in America was ill-timed. This potentially powerful classroom-based learning and teaching innovation was overshadowed by the No Child Left Behind Act with its intense pressure to raise test scores on external accountability tests (p. 1).

As a result there was a “burgeoning of so-called formative assessments offered by commercial test publishers to help raise test scores for NCLB” (Shepard, 2005, p. 1). These assessments, henceforth called interim/benchmark assessments, were problematic because they were not consistent with the research of Black and Wiliam which focused on assessment used during instruction to adjust teaching and learning (Shepard, 2005; Perie, Marion, Gong, & Wurtzel, 2007). Perie et al. concluded that the labeling of interim/benchmark assessment as formative would
be misleading to states, school districts, schools, and teachers wanting to implement formative assessment.

Black and Wiliam (2005) argued that the interim/benchmark assessments were merely “early warning summative assessments” that were designed to look like state-wide assessments and used to predict performance on end-of-year or end-of-course assessments. According to Black and Wiliam, interim assessments coupled with large-scale assessments represented substantial and well-entrenched obstacles that would prevent the growth of formative assessment use in the United States.

Recognizing this emerging issue, the Council of Chief State School Officers (CCSSO), an organization devoted to bringing state departments of education together to collaboratively work to improve student achievement, began to promote a comprehensive balanced assessment system (CCSSO, 2006). The concept of a comprehensive balanced assessment system was developed to help policymakers distinguish the difference between formative assessment and interim assessment.

Need for a Comprehensive Balanced Assessment System

The comprehensive balanced assessment system promoted by CCSSO and others includes and values three types of assessments – statewide large-scale assessment, interim/benchmark assessment administered at the district level, and classroom assessment that includes both summative and formative assessment. As
illustrated by Figure 2, all aforementioned types of assessments are included in this system and are aligned to state standards. At the base of the assessment system is classroom assessment. It is represented using the largest rectangle in Figure 2 because not only is it the foundation of the system but, it is also the type of assessment that should be used most often.

**Aligned to State Standards**

![Assessment System Diagram]

*Figure 2: A Comprehensive Balanced Assessment System*

Each type of assessment, its uses, and limitations (Perie et al., 2007; Stiggins, 2006) in this system are important. Statewide tests are summative large-scale assessments used to get a snapshot of where students are at a given point in the year. They are primarily used for accountability and to inform policymakers (Chudowsky & Pellegrino, 2003; Heritage, 2007; Perie et al., 2007; Stiggins, 2006). The data from these assessments can be aggregated across
subgroups, classrooms, and schools. Policymakers can use the information to make decisions about educational programs and policies.

Inherent in their design, statewide assessments are not diagnostic and therefore cannot provide information to teachers and students to assist them on a daily basis (DePascale, 2003; Heritage, 2007; Perie et al., 2007; Stiggins, 2006). Budget constraints limit the types of items that can appear on these assessments (Chudowsky & Pellegrino, 2003). Open-ended, constructed-response type items are much more expensive to hand score than multiple-choice items that can be machine scored. Therefore, in this system they are the least flexible type of assessment (Perie et al., 2007).

Interim/benchmark assessments are summative assessments administered periodically throughout the year. The frequency varies from district to district. Some districts use interim/benchmark assessments quarterly to benchmark progress while others use them more or less frequently. District leaders and school administrators are the main end-users of these assessments for the purpose of program evaluation (CCSSO, 2007; Heritage, 2007; Perie et al., 2007; Stiggins, 2006).

Interim/benchmark tests are often thought of as mini-statewide assessments because they have the same item types and curriculum coverage as statewide assessments. They are used to help districts predict how students will score on end-of-course or end-of-grade assessments. The data can be aggregated across
classrooms. However, they have been found to cover too much information and provide too little detail for teachers to find them useful for improving classroom instruction (Shepard, 2005).

As shown in figure 2, classroom assessments in this system should consist of both summative and formative approaches (Clarke, 2005; Crooks, 1988; NCTM, 2000). The data from these assessments should mainly be used by teachers, students, and parents. Classroom summative assessments should be given periodically to assign grades and to measure learning that has taken place. Formative assessment should occur during the learning process and provide feedback to students and teachers.

According to some researchers, summative assessments should be a small component of all classroom assessment, yet they received the most attention (Clarke, 2005; Crooks 1988). In the age of high stakes accountability testing, summative assessments were found to be the dominant mode of classroom assessment (Bright and Joyner, 1998; Heritage 2007; Stiggins, 2004; Wilson, 2005).

With a focus on summative approaches, classroom assessment was often viewed as an interruption to the instructional and learning process (Clarke, 2005; Heritage, 2007). It was mostly used to determine if learning objectives had been met with the desired outcome being a normal distribution of grades. This traditional approach was used to sort students into winners and losers (Heubert & Hauser,
1999; Stiggins, 2006) in a race that had very few winners. As summative assessments became more and more prevalent in the classroom, the use of formative assessment was disappearing.

Need for Formative Assessment

Researchers argued that formative assessment is integral to effective teaching (Black & Wiliam, 1998; Clarke, 2005; Heritage, 2007; NCTM, 2000; Wilson, 2005). Stiggins (2006) used the following illustration (see Figure 3) to emphasize the need to use assessment to help students while they are learning rather than just categorizing them at the end of the learning process.

©Educational Testing Service/Assessment Training Institute

Figure 3: Stiggins’ Assessment for Learning Illustration
In Figure 3, students on the top met the state standard and the students on the scaffolding were on their way to meeting the standard. Stiggins (2006) stated that, in the past, educators were only concerned with distinguishing between those who met the standard and those who did not. In the old model, many students were left behind.

In the new era of NCLB, Stiggins (2006) argued that educators must be focused on all the children and use formative assessment (assessment for learning) to help all students on their journey to meeting the standards. As Figure 3 illustrated, although students may not be at the top, they are in motion. During the learning process, formative assessment would provide students with opportunities to show what they can do without being graded. Without grades, students are not penalized for their lack of understanding while they are still learning concepts. Because no high stakes are attached, students are more likely be motivated to improve their learning (Clarke, 2005; Crooks, 1988).

Researchers have stated that the lack of formative assessment found in classrooms is a detriment to low-performing students (Wiliam, 2005; Stiggins, 2006). Data from the TIMSS 1995 study suggested that the lowest performing students were exposed to the most summative testing (Beaton, Martin, Gonzalez, Kelly, & Smith, 1997). The overuse of summative assessments was associated with lower self-efficacy, less motivation, and higher drop out rates (Stiggins, 2006).
The National Council of Teachers of Mathematics (NCTM, 1995) stated that the failure to meet the needs of low-performing students is a waste of human potential.

To promote the use of formative assessment in a comprehensive balanced assessment system, CCSSO created a national initiative (CCSSO, 2006). The initiative formally began in January 2006, when CCSSO formed the Formative Assessment (FA) Advisory Group consisting of measurement and education researchers (e.g., Jim Popham, Lorrie Shepard, Rick Stiggins, and Dylan Wiliam) and state agency leaders from across the nation to explore how states could promote the use of formative assessment in the classroom to improve student achievement as they educated students in the 21st century.

CCSSO identified several of its existing State Collaborative on Assessments and Student Standards (SCASS) groups that wanted to form subcommittees to focus on formative assessment. However, due to the importance of the initiative, CCSSO decided to form the Formative Assessment for Students in Teachers (FAST) SCASS (CCSSO, 2006). The purpose of the FAST SCASS was to allow member states (e.g. Utah, Maine, Michigan, and Delaware) and research organizations (e.g., National Center for Research, Evaluation, Standards, and Student Testing) to collaborate on ways to implement the initiatives set forth by the FA Advisory Group. The FAST SCASS identified four key areas of focus – policy, technology, research, and professional development. The biggest area of concern was the area of professional development.
Need for Professional Development

Although it was believed that the use of formative assessment would produce self-directed learners who have the ability to solve problems in a technologically globally competitive society, it was also believed that teachers and administrators lacked the necessary skills and knowledge for effective implementation. Therefore, there was a need for professional development in the area of formative assessment (Perie et al., 2007; Heritage, 2007; Shepard, 2005; Stiggins 2004:2006; Wiliam, 2005). This was captured by Heritage (2007) who stated:

Rather than providing teachers with more tests, leaders at state, district, and school levels should invest in a coordinated effort to establish structures and provide resources that support effective professional development. This investment is a long-term project that should not be shortchanged. The payoff will improve student learning, and that is surely worth it. (p. 145)

To help study states’ capacities for providing professional development in formative assessment the USED granted CCSSO and ten states with an Enhanced Assessment Grant (EAG). The focus of the grant was to determine how to build states’ capacities to support professional development for teachers.
At the time of this study, the author/researcher was actively involved in the CCSSO formative assessment initiative. She was a member of the CCSSO FA Advisory Group as a state leader in her capacity as the Section Chief for Testing Policy and Operations in the Accountability Services Division of the North Carolina Department of Public Instruction. She was a member of the FAST SCASS and was the EAG team leader for North Carolina.

This study was made possible by participation in the EAG grant. However, this study did not focus on how to build states’ capacities to provide professional development. It focused on how teachers and students responded to professional development which called for a change in their classroom assessment practices. This study was situated in high school mathematics classrooms in the age of high stakes testing.
CHAPTER 2

Literature Review

In 2005, at the time the literature review began, there was little research in the area of formative assessment in high school mathematics and various definitions of formative assessment existed. Therefore, a grounded theory approach was used. Grounded theory is action or change oriented research, used to examine a basic social process from the perspective of human interactions (Patton, 2002). Because of the limits of the literature review at the beginning of the study, the author/researcher allowed theory to emerge directly from the data and remain "grounded in" the data. The literature was included as became part of the data. The literature was reviewed continuously throughout data collection and analysis. Presented here is the culmination of that continuous review that ended in 2008.

Definitions

Researchers frequently use different terms for explaining similar concepts and the same terms to describe very different concepts. This phenomenon also occurs when researchers describe classroom assessment. Some researchers take a holistic view of and make no distinctions between formative and summative approaches (Brookhart, 2006) while others make a distinction between the two (Clarke, 2005; Crooks, 1988; Perie et al., 2007). For this literature review, the latter approach was used. A very brief description of summative assessment is presented followed by an extensive look at the literature related to formative assessment.
Summative Assessments

McTighe and O’Conner (2005) defined summative assessment as an instrument that summarizes what students have learned at the end of an instructional unit. The results are reported as a grade, a score, or a level. Arter (2003) defined summative assessment as an instrument that summarizes how much a student has attained at a particular point in time. This definition applies to large-scale assessments used in accountability, interim assessments used to benchmark progress, and classroom assessments that are graded.

Most often, summative assessments are in the form of tests that are evaluated. Tests can include selected-response items such as multiple-choice, true-false, or matching. Tests can also include written-response items that are short (e.g., 1 – 2 sentences) or extended (e.g., one or more paragraphs). Performance assessments, portfolios, and presentations can fall into this category.

Formative Assessment

Formative assessment has been described differently by different researchers. Scriven (1967) has been credited with creating the term formative evaluation which was a precursor to the term formative assessment. Scriven developed the term to describe what he considered the opposite of curriculum evaluation that is summative and occurs at the end of a curriculum program. Scriven described formative evaluation as occurring while a program was being implemented so that improvements could occur prior to the end of the program.
Bloom, Hastings, and Madaus (1971) extended this definition to describe the opposite of summative evaluation tests that were given at the end of a unit for the purpose of grading, certification, evaluation of student progress, or curriculum effectiveness. The opposite of summative evaluation would be that which students, teachers and curriculum makers would find useful in “improving what they wish to do” (Bloom, et al. 1971, p. 117). Formative evaluation was also used in the Bloom’s Learning for Mastery Model (see Figure 4).

**Steps in the Bloom’s Learning for Mastery Model**

1. Determine student prerequisite skills
2. Evaluate components of a learning target (formative evaluation)
3. Determine the level of achievement of the learning target (summative evaluation)

*Figure 4: Bloom’s Learning for Mastery Model*

The first step as shown in Figure 4 was to determine the student’s preexisting knowledge related to the specific learning target. Therefore, it was important for educators to be able to deconstruct learning targets so that they could evaluate each component in step two. This type of evaluation was considered formative because its purpose was to help the student improve the learning of the target.
The final step was summative because its purpose was to determine the level of achievement of the learning target. This model shows a distinction between the formative evaluation and the summative evaluation. However, in practice, this model was often interpreted as providing students with ongoing summative assessments until the optimal level of achievement was reached.

Later, the term “assessment” was used in place of “evaluation.” The term formative assessment was defined as a tool (Kahl, 2005), a test (Shepard, et al. 2005) and as a process (Cowie & Bell, 1999). Others defined it not by what it was but how it was used. According to Arter (2003), Black and Wiliam (1998), Sadler (1989), and Wiliam and Black (1996), a tool/test/process could be called formative assessment only when evidence from the tool/test/process was used to modify teaching. Others argued that the data derived should be used to adjust, enhance, or shape teaching (Cowie & Bell, 1999, Gipps, 1994, Sadler, 1989; Popham, 2005) but if it were not, the tool/test/process could still be called formative assessment.

Agreement regarding the definition of formative assessment centers on feedback. Researchers agree that formative assessment should include feedback (Black & Wiliam, 1998; Cowie & Bell, 1999; Heritage, 2007; Kahl, 2005; McTighe & O’Connor, 2005; Sadler, 1989; Shepard, et al. 2005). Some emphasize that feedback should help students to adjust and modify learning (Arter, 2003; Black & Wiliam, 1998; Cowie & Bell, 1999) and assist students in self-assessment and goal setting (Arter, 2003).
Black and Wiliam (1988) and Stiggins (2004) asserted that formative assessment should involve students as partners and should engage students in assessing their own learning and the learning of their peers. More specifically, it should involve both teacher-student and student-student interactions.

In October 2006, the author/researcher, as part of the CCSSO FA Advisory Group and FAST SCASS, used the evolving literature review found in this chapter to develop and bring consensus around the following CCSSO definition of formative assessment.

Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievement of intended instructional outcomes. (CCSSO, 2008, p. 1)

This CCSSO definition of formative assessment was used for this study.

*Comparison of Formative and Summative Assessment*

The researcher made distinctions between the characteristics of formative and summative assessment (see Table 1). The first and second characteristics were made in regards to time. Formative assessment occurs during instruction while summative assessment occurs at the end of an instructional unit (e.g. a day, a week). Therefore, formative assessment is ongoing, “minute by minute” (Wiliam, 2005) while summative assessment is periodic (e.g., daily, weekly, yearly).
Table 1: Differences between Formative and Summative Assessments

<table>
<thead>
<tr>
<th>Formative</th>
<th>Summative</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Occurs during instruction</td>
<td>1. Occurs at the end of an instructional unit</td>
</tr>
<tr>
<td>2. Ongoing</td>
<td>2. Periodic</td>
</tr>
<tr>
<td>3. Descriptive feedback</td>
<td>3. Evaluative feedback</td>
</tr>
<tr>
<td>5. A Process</td>
<td>5. A Product</td>
</tr>
</tbody>
</table>

The third and fourth characteristics shown in Table 1 relate to feedback. On a summative test, students receive evaluative feedback that judges their work. Evaluative feedback gives the student a summary of how he or she has done but does not provide any information about how to improve the learning (Stiggins, Arter, Chappuis, & Chappuis, 2006). Grading is an example of evaluative feedback. By contrast, descriptive feedback provides students details about how to improve their learning and it does not involve grading.

Formative assessment is a process (see characteristic 5) while summative assessment is a product (e.g., test). Formative assessment involves getting evidence of learning to adjust teaching and/or learning. Sadler’s (1989) theory of formative assessment stated that a student must be able to answer three questions. Where am I going? Where am I now? How do I close the gap? In order for students to know
where they are going they need clear and appropriate learning targets and clear
criteria for success.

Literature Related to Formative Assessment

Clear and Appropriate Learning Targets

Before teachers can provide students with clear and appropriate learning
targets teachers must first understand and master the learning targets they are
teaching (Heritage, 2007). This content knowledge will help teachers’ abilities to
breakdown achievement goals/standards into sub-goals or sub-objectives. Teachers
must then use their pedagogical content knowledge to know how best to help
scaffold students in the learning process.

If students are provided achievement targets that are written at the standard
level rather than being deconstructed at a more detailed level, students most likely
will not understand the targets they must meet (Stiggins, 2006; Heritage, 2007).
Understanding the targets is important because students must be able to answer the
suggested that teachers deconstruct state standards into smaller learning targets that
can serve as scaffolding leading to the standards. Furthermore, these smaller targets
should be translated into language that students can understand. Stiggins et. al
refers to this language as student-friendly language.

Campos and O’Hern (2007) found that when they began putting
mathematics goals and objectives in student-friendly language, 90% of the fifth
graders and 95% of the first graders in their study reported that they knew and could speak about the targets they were expected to meet. This was compared to only 50% of the first and fifth graders prior to the intervention. Campos and O’Hern also found that identifying the criteria for success was also important for student success.

Clear Criteria for Success

It is important that teachers provide students with clear criteria for success by providing examples and non-examples of high quality work (Crooks, 1988; Stiggins 2004; Wiggins, 1993). In other words, teachers must have high expectations and teach those expectations (McIntosh, 1997). Students can later use these criteria to judge their own work (self-assessment) or the work of their peers (peer-assessment). Black and Wiliam (1998) reviewed the research on formative assessment and concluded that learners who understood the goals and the criteria for judging their work were more active participants in the process.

Campos and O’Hern (2007) findings were consistent with this earlier research that found that achievement increased when students became more involved in the assessment process. Student mathematical learning improved when teachers provided students with specific examples of work and students reflected on these examples. The reflections occurred in whole-class discussions, with peer buddies, and as individuals. Specifically, Campos and O’Hern found that all the first graders and 75% of the fifth graders in their study were able to appropriately
self-regulate and monitor their progress in mathematics once they were able to distinguish between strong and weak work.

Clear criteria for success help students to understand where they are in the learning process (Heritage, 2007; Stiggins, et.al, 2006). To accomplish this, teachers should provide students with instructional activities that will elicit evidence of learning. Both teachers and students can use this evidence to compare against the criteria for success.

White and Frederiksen (1998) looked at how reflective assessment would impact learning. The study involved three teachers who taught four classes each of eighth grade science. Each class was given basic skills tests to determine their achievement (i.e., high, middle, or low) prior to the start of a new unit. In two of the classes, the teacher provided the lessons and just asked the students to tell what they liked and disliked about each lesson. In the other two classes, the teacher helped the students reflect on criteria for success and to self-assess themselves using the criteria. As a part of the study, students were assigned a project that was teacher graded. The students that used reflective assessment had higher mean scores than students who did not.

The White and Frederiksen study illustrated that just having criteria is not enough. Students must use the criteria in the context of their own work (Wiliam, 2005). Teachers often develop rubrics that communicate essential features of good work (Shepard, et al. 2005). Studies have shown that students find rubrics helpful
in guiding their work and thinking and to self-assess their completed work (Moon, Callahan, Brighton, & Tomlinson, 2002; Donovan, Larson, Stecheschulte, Taft, 2005).

Rubrics

A rubric is a rating scale consisting of ordered categories containing descriptions and exemplars that are used to sort student-produced responses into levels of achievement (Schafer, Swanson, Bene, & Newberry, 1999). Designed to "yield information about students' strengths and weaknesses relative to the content and processes being assessed" (Moon, et al., 2002, p. 6), rubrics are most often used summatively at the end of a unit.

Rubrics can serve as a reflection of what teachers and students value in regard to the learning experience (Donovan, et al., 2005). Used formatively during instruction, rubrics can help students assess their own work by showing them how levels of their own achievement can be improved (Schafer, et al., 1999). By doing so, the goal of formative assessment can be realized. The goal of formative assessment should be to develop students who self-regulate their own learning by understanding and accepting ownership of the learning goals (Sadler, 1989).

Self-regulation

Self-regulation refers to the extent to which individuals metacognitively participate in their own learning (Travers & Sheckley, 2000). Metacognition refers to both the knowledge of cognition and the regulation of cognition. The knowledge
of cognition involves understanding the limitations of one’s own cognitive ability and the required cognition to be successful on a given task. The regulation of cognition involves checking solutions, planning next steps, or evaluating strategies (Baker and Brown, 1984; Winne, 1996). Garcia and Pintrich (1992) found that metacognition is highly correlated to critical thinking.

Self-assessment is the process in which individuals reflectively think about and apply standards or criteria in the context of their own work and monitor their experiences (Sadler, 1989). When confronted with experiences that do not match their own, students make assimilations and accommodations for them. Research has suggested that high school students are better self-assessors than elementary school students, teacher practices can support the development of students' self-assessment skills, and students' self-assessments of their own academic abilities are important factors in academic achievement.

Blatchford (1997) conducted a longitudinal study to determine how students assessed themselves over time. The study took place in London and consisted of 7-year-olds and 11-year-olds. The 7-year-olds were followed until they reached the age of 16. Each student was interviewed individually to assess their views regarding school and school work. For example, students were shown three pictures of the same group of children and were told one group was good at mathematics, one was not good at mathematics, and one was not particularly good or not good. Students identified the group they belonged to and why.
Each student took an academic self-assessment instrument and a standardized mathematics assessment. These instruments were used to determine the accuracy of the self-assessments. Over all, 7-year-olds were not accurate judges of their academic attainments and typically had high self ratings. The 11- and 16-year-olds more accurately judged themselves which resulted in a drop in self-assessment scores.

Adams, Cooper, Johnson, and Wojtysiak (1996) found similar results when they studied 7-year-olds. These researchers recommended that teachers provide all students with opportunities for goal-setting, self-evaluation, and reflection so that students can self-regulate their learning. In a study involving 139 volunteer mathematics students in England, it was found that the teacher was a critical factor in student self-regulating behavior. Students who were taught using the lecture method were most likely to be teacher-regulated than self-regulated (Travers & Sheckley, 2000).

Researchers (Biggs, 1988) found that self-assessment helped increase students deep approaches to learning and decrease surface approaches. Klenowski (1995) found that students who participated in self-assessment became more interested in the criteria than in grades, that they were more honest about the quality of their work, and realized that they had to be prepared to defend their work.
based on the established criteria. Teachers must teach students to assess their own learning. Students should be asked, “Do you think your answer demonstrates understanding why or why not?

*Use of Descriptive vs. Evaluative Feedback*

In reviewing the literature on formative assessment, a distinction was made between evaluative feedback associated with a summative assessment and the descriptive feedback used with formative assessment. Evaluative feedback gives the student a summary of how he or she has done but does not provide any information about how to improve the learning or what specifically was done well or needs improvement. Examples of evaluative feedback include, “Good Job”, “A”, “87%”, “Try Harder.” By contrast, descriptive feedback provides students details about what was done well, what needs improvement, or how to improve the next time (Stiggins et.al, 2006).

The definition of descriptive feedback is consistent with Ramaprased’s (1983) use of feedback. He stated that feedback provides information about the gap between the actual performance and the desired performance that is used to alter the gap. Emphasis was placed on the use of the information. If the gap was not altered, Ramaprased did not consider the information to meet the definition of feedback. Sadler (1989) used Ramaprased’s view of feedback when developing his theory of formative assessment. Sadler further explained that feedback is
information that is external to the user while information internal is considered self-monitoring data.

Day and Cordon (1993) looked at two variables, the amount of feedback provided when students were unable to accomplish a task and student achievement. Using fourth-grade students working on a set of reasoning tasks, Day and Cordon divided the sample into two groups. One group was provided only enough feedback to get them working on the task again while the other group was given a full explanation of how to solve the task and then given a new problem to solve. Day and Cordon found that the students who received only enough feedback to get them moving forward again achieved more and retained the information much longer.

Butler (1988) compared the use of evaluative feedback, descriptive feedback, and a combination of both. Her sample consisted of 132 Year 7-students in twelve classes in four Israeli schools. Each class was provided the same divergent thinking task that was collected and reviewed.

For the review, the three classes were divided into three groups of 4. The independent reviewers assigned the students in the first group evaluative feedback (i.e., grades) only. They assigned the students in the second group descriptive feedback and evaluative feedback, and assigned the third descriptive feedback only. At the start of the next lesson, the students received their work from the first lesson. They were provided with similar tasks and told that they would get the same type
of feedback after completing the second lesson. The students work was again reviewed by independent reviewers.

Butler found that students who received descriptive feedback only improved their scores, on average, 30%. Those who received evaluative feedback only and a combination had no gains. Butler also observed that students who received descriptive feedback only were highly motivated to complete the second set of tasks. By contrast, only the students who received high grades in the other two groups were motivated to complete the second set of tasks. Based on subsequent interviews with students, Butler concluded that the benefits of descriptive feedback were erased when grades were provided because students tended to focus on the grades more than the descriptive feedback.

Brookhart’s (1997) work supported Butler’s findings. Brookhart found mathematics learning can be negatively impacted when students are given grades. Brookhart concluded that students most likely viewed the grades as judgments rather than as useful feedback.

Meta-analyses have shown that feedback can have a negative impact on achievement (Black & Wiliam, 1998; Kluger & DeNisi, 1996). Although one-third of the studies showed a negative effect, they found that feedback had a positive influence on achievement when it focused on features of the task and emphasized the learning goals (i.e., descriptive), rather than comparison of students.
(i.e., evaluative). Kluger and DeNisi reported an average effect size of .40. The meta-analyses also showed feedback has an impact on student self-efficacy.

**Self-efficacy**

Self-efficacy is how one feels about one’s own abilities (Bandura, 1992). Studies have shown that students see ability as fixed or incremental (Dweck, 1986). Students who see ability as fixed have what Butler (1987) refers to as high ego-involvement. Ego-involvement is related to a student contributing success or failure to his/her own abilities. By contrast, students who see ability as incremental have high task involvement. Task-involvement is related to a student contributing success to the task itself.

Students who see ability as fixed see a new task as a chance to affirm their ability or reaffirm their lack of ability. Those students who are confident in their ability to be successful will readily attempt a difficult task. Those who are not confident in their ability to be successful will be reluctant to attempt a difficult task and would rather be seen as lazy by not trying than to be “shown up” by their peers.

Dweck (1986) concluded that the type of feedback students receive has an impact on how they view their ability. Feedback that was focused on grades (i.e., evaluative) negatively impacted self-efficacy while feedback focused on how to improve learning (i.e., descriptive) had a positive impact. Dweck found that the use of grades as feedback was associated with a focus on competition, judgment,
and comparisons with others while feedback that focused on how to improve learning was associated with a belief in one’s own ability to improve and learn, application of effort, and satisfaction with one’s own performance when doing challenging tasks. Table 2 summarizes these differences.

Table 2: Impact of Feedback on Self

<table>
<thead>
<tr>
<th>Focus on grades (evaluative feedback)</th>
<th>Focus on how to improve learning (descriptive feedback)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability leads to success</td>
<td>Effort leads to success</td>
</tr>
<tr>
<td>Concern about being judged as able</td>
<td>Believes in one’s ability to improve and learn</td>
</tr>
<tr>
<td>Satisfied when does better than others or succeeding with little effort</td>
<td>Gains satisfaction over performance when doing challenging tasks</td>
</tr>
<tr>
<td>Emphasis on interpersonal competition and public evaluation</td>
<td>Emphasis on personal success at different tasks</td>
</tr>
<tr>
<td>When task is difficult evaluates self negatively</td>
<td>When task is difficult applies problem solving skills and engages in task</td>
</tr>
</tbody>
</table>

Dweck recommends that feedback should be designed to help all students come to know that ability—even in mathematics—is incremental.
Butler studied the impact of feedback on ego- and task-involvement. Butler sampled 200 Year 6- and Year 7- Israeli students. All students were provided with a divergent thinking activity. During this study, a fourth of the students received comments (i.e., descriptive feedback), a fourth received grades (i.e., evaluative feedback), a fourth received praise (i.e., evaluative), and a fourth (control group) received no feedback at all.

Butler concluded that the grading and praise increased the students ego-involvement (i.e., focus on contributing success to one’s own ability) without an increase in achievement. In other words, students who were not successful, would contribute failure to their lack of ability rather than to the difficulty of the task and see ability as fixed. By contrast, those who were given comments increased their task-involvement and improved their achievement more than students in the control group. This finding suggested that students who receive descriptive feedback are more likely to contribute success or failure to the difficulty of the task and see ability as incremental.

*Formative Assessment and Achievement on Summative Tests*

A meta-analysis conducted by Black and Wiliam (1998) showed that the use of high quality formative assessment can produce learning gains of $\frac{1}{2} - 1$ standard deviation on summative tests. This is equivalent to raising an average student’s score at the 50th percentile to the 85th percentile (Black & Wiliam, 1998). According to Black and Wiliam, formative assessment strategies are a powerful
way to increase conceptual understanding while raising test scores on mandated large-scale assessments. This finding was consistent with the findings of Black, Harrison, Lee, Marshall, and Wiliam (2004) who also studied the impact of formative assessment on student achievement.

Black et al. selected two local school systems in the United Kingdom. Each school system selected three schools and the schools selected two mathematics and two science teachers. The teachers participated in a series of half-day and one-day workshops where they were introduced to the principles of formative assessment and provided time to develop lesson plans to incorporate formative assessment strategies.

Each teacher was observed in the classroom and provided support as they attempted to implement the formative assessment strategies. The researchers used the results from national large-scale assessments that were given later in the year to measure achievement. Black et al. (2004) concluded that improving formative assessment strategies produces higher achievement on large-scale assessments. In particular, that a school at the 25th percentile of achievement nationally could raise its performance into the 50th percentile.

Teachers as Assessors

Historically, surveys have shown that teachers consider assessment to be an important aspect of teaching and learning (Gullickson, 2001) but are concerned about perceived inadequacies in their ability. Stiggins & Bridgeford (1985) found
that teachers reported being concerned about the lack of formal training in assessment. In addition, Stiggins and Bridgeford found that teachers who reported receiving formal training also reported that the training was of little use or relevance.

Mertler (2003) conducted a study of 67 pre-service and 197 in-service teachers to assess their assessment literacy. She used a sample of convenience based on geographical location to a Midwestern university. The pre-service teachers had recently completed a course in classroom assessment and were a semester away from student teaching. She found that both groups of teachers were weakest in the area of valid grading procedures and communicating assessment results. Mertler concluded that both in-service and pre-service teachers need to have more information about assessment practices. In particular, this information should be provided by someone who understands the important link between assessment and instruction.

Rodriguez (2004) looked at data from the Third International Math and Science Study (TIMSS) to determine if classroom assessment practices had an effect on student performance in the United States. He found that the number of teacher-made multiple-choice tests had an inverse relationship with the mean classroom performance on the TIMSS. In addition there appeared to be an interaction effect between teacher-made multiple-choice tests and student self-efficacy. Self-efficacy alone had a positive effect on student performance.
However, in classrooms where teacher-made tests were more frequently used, the positive effects were diminished. He speculated that this could be caused by the teachers' inability to develop appropriate classroom assessments.

(Black and William (1998) conducted a review of over 250 studies involving classroom assessment. They reported that when teachers regarded the purpose of assessment as just summative, their students were less likely to engage in activities that would improve their work. In this situation, parents and students focused more on interpreting a grade or score than on what was learned and what still needed to be learned. Based on this finding, Black and Wiliam recommended that teachers spend more time on assessments for learning (formative) that are used to support learning and less time on assessments of learning (summative) that simply label or grade students.

There are many models that illustrate how to promote learning in the mathematics classroom. Simon (1995) suggested the Mathematics Teaching Cycle (MTC) which emphasizes the need for pedagogical decision-making. Teachers must know the activities, simulations, investigations, and experiments that will help students understand the content (Shulman, 1986). The MTC consists of learning goals, planned activities, and learning progressions. Learning progressions are the steps a learner must go through to reach a learning goal.

Teachers must have their own mathematical knowledge to interpret the language and actions of their students. Teachers must know more than just facts
and concepts (Shulman, 1986). They must know how mathematics is organized and understand the reasoning behind those facts and concepts.

The Heritage (2006) model based on the work of Sadler (1989) shown in Figure 5 is not specific to mathematics but contains the elements of the MTC.

![Feedback Loop Diagram](image)

**Figure 5:** Heritage (2006) Model of the Feedback Loop
The top left of the model shows the importance of pedagogical knowledge, domain knowledge, and knowledge of learning progressions. It also emphasizes teacher knowledge of formative assessment strategies. For deep learning to take place, teachers must be able to make appropriate decisions about what knowledge the student needs to learn next and by which path (Steffe, 1990). Without the appropriate content knowledge a teacher focuses on surface level understandings, is unable to translate the skills into clear criterion, and is unable to provide related descriptive feedback to students.

Central to this model is the feedback loop. The purpose of the feedback loop (see Figure 5) is to close the learning gap. The arrows show that the evidence is evoked through teacher, peer or self assessment (see bottom right of Figure 5). Once evoked, this evidence is interpreted and the gaps in knowledge are identified. The teacher or student then sets instructional goals and establishes criteria for success. Teachers must recognize individual differences so that instruction can be adjusted based on the needs of the learner and his/her responses to the instruction. The teacher provides feedback or scaffolding that helps extend the knowledge of the student. The arrows help illustrate that the feedback loop is cyclical and may need to be repeated multiple times before the gap is closed.

Regardless of which model is used, researchers consistently state that teachers do not currently have the ability to implement formative assessment and that there are barriers that need to be addressed. Ruiz-Primo and Furtak (2007) did
an exploratory study involving 3 middle school science teachers implementing formative assessment. Teachers had difficulty eliciting evidence from students. Teachers were good at asking students to respond to questions that would transition the teacher from one part of the lesson to the next. However, the questions most often did not elicit evidence of learning that would be used to help the student improve his/her learning. When evidence was elicited, it most often was unused.

This finding was consistent with reports from Wiliam (2005) about formative assessment in the United States. Through his research and observations, Wiliam found that teachers tended to use questioning as a classroom management tool. Questioning kept the lesson moving and the teachers maintained control of their classes. According to Wiliam, this type of questioning didn’t promote the idea of students as partners that the research shows is effective in promoting learning.

Social-Constructivist Perspective

Black and Wiliam (1998) identified two problems with changing to a system of classroom assessment that emphasizes the use of formative assessment. According to Black and Wiliam, the nature of teachers’ beliefs about learning and the beliefs teachers hold about the potential for all students to learn are barriers to success. The NCTM (2000) also recognized these barriers when they called for a reform in assessment compatible with a social-constructivist model of teaching and learning which emphasized that all students can learn mathematics.
Social-constructivism is a theory about knowledge and learning. In reviewing the literature from the last decade, many researchers use the terms constructivism and social constructivism interchangeably to represent the belief that individuals construct their own knowledge schemes, Piaget’s term for cognitive structures, in response to their interaction with both the physical and social world. However, constructivism and social constructivism are not interchangeable ideas and are very different in meaning. Constructivism will be henceforth called cognitive constructivism to help the reader distinguish it from social constructivism.

Historically, the work of Dewey, Bruner, Piaget, and Vygotsky inspired the theory of social-constructivism (Marlowe and Page, 2005). Social constructivism combines the principles of cognitive constructivism and sociocultural theory. Cognitive constructivism refers to the mental construction of knowledge and sense making. A learner’s existing knowledge and beliefs enable or impede the learning when new information is introduced. A learner must rely on self-monitoring and metacognition to be aware of when and how to use skills. To become an expert in a content area means to have a coherent way of thinking and representing problems rather than just having a lot of information.

When an individual is confronted with new information, that individual either assimilates the new information into existing schemes or alters existing
schemes to accommodate the new information creating what Piaget called a sense of equilibrium (Marlowe and Page, 2005).

Sociocultural theory is based on Vygotsky’s zone of proximal development. The zone of proximal development is the region on an imaginary learning continuum between what a child can do independently and what the child can do with support (Vygotsky, 1978). Helping a student get from one point to the next is called scaffolding. Bruner (1975) characterized scaffolding as guidance, coaching or encouragement that adults provide a child so that the learner is able to accomplish that which he or she could not otherwise accomplish.

Vygotsky (1978) highlighted the importance of using tools to mediate learning. Dialogue is a type of discursive tool that can be used as a means of reflecting and communicating in the world of the classroom. Bruner (1975) stated that conceptual learning is a collaborative enterprise involving an adult who alters development within the child by engaging the child in dialogue. The dialogue provides hints and props that allow the child to climb by guiding the child to each step until the child is capable of setting and evaluating his/her own goals and monitoring his/her own progress.

In the nineteenth century, John Dewey believed that students should be active participants in the learning process and that learning should involve reflective thinking and problem solving (Marlowe and Page, 2005). He believed problems should be solved in groups where students could work together to
organize, digest, and process information. Bruner further postulated that discovery increases intellectual ability and leads to new insights and new inquiry. For example, when students make errors, the errors should not be immediately corrected. Errors can cause what Piaget refers to as disequilibrium. When disequilibrium occurs students should be encouraged to develop and test hypotheses.

Although social constructivism is not a theory of teaching, it guides classroom and instructional practices (Fosnot, 1996). A social-constructivist classroom is an environment where students are responsible for communicating and justifying their ideas. Students are provided opportunities to clarify meanings and understandings of mathematical concepts and to practice the social norms and use of mathematical language (Cobb, Wood, and Yakel, 1993). Lave and Wenger (1991) found that students who are novices become increasingly adept at participating in a community. As a result, they become more expert in that community.

In the community of the mathematics classroom, the teacher must guide instruction and students must be actively involved in the learning process (NCTM, 2000). In this type of environment, the students collaborate with the teacher and with one another to form a sociocultural learning system. Through sociocultural interactions, learners construct new meaning and gain deeper understanding by
being able to dialogue with other students. Dialogue is important in helping students make connections with abstract ideas (Shepard, 2000).

von Glasersfeld (1996) reminds teachers that students may view the world differently than teachers expect. Teachers must look at students as individuals and try to find out how they view the world. Teachers must understand that students construct their own knowledge and are not passive receivers of knowledge. Educators must provide students opportunities and incentives that will allow them to construct knowledge.

Bjorkqvist (1996) recommends that educators not only reflect on how they are teaching students but also on how they are assessing students in the classroom. He states that teachers must develop assessments that are aligned with the philosophy of social constructivism in order for social-constructivist classrooms to continue to exist. In keeping with this belief, this research is guided by the following questions:

1) If students actively construct their own knowledge, shouldn't they also be involved in constructing the assessment of that knowledge?

2) What happens to teacher and student behaviors when a traditional view of assessment transforms to a social-constructivist one?
Summary

The goal of formative assessment should be to develop students who self-regulate their own learning by understanding and accepting ownership of the learning goals (Sadler, 1989). This is consistent with Bruner’s (1975) philosophy of scaffolding – to foster internalization and the assumption of responsibility by the learner. It is also consistent with the goal of 21st century learning.

Formative assessment is an integral part of teaching. Attributes of effective formative assessment include collaborative partnerships between teachers and students, students actively engaged in self-assessment as well as peer-assessment, and evidence eliciting instructional activities that often include classroom discourse. These attributes are compatible with a social-constructivist model of teaching and learning because it is the process that recognizes that all students can construct knowledge with the help of teachers and peers (Black, 2001; Klenowski, 2004; Clarke, 2005) and that “knowledge is situated in the discourse and practice of the community” (Black, 2001, p. 79).

Research Questions

Based on the literature review, there is concern that the emphasis on standardized tests and teachers’ emulations of them thorough summative classroom assessments are barriers to implementation of formative assessment. This study was
conducted to answer the following research questions.

1. What themes emerge as teachers implement formative assessment in the age of high stakes testing?

2. What new behaviors and/or attitudes do students exhibit as a result of the implementation of formative assessment?
CHAPTER 3
Methodology

This chapter describes the purpose of this study, the methodology used, the treatment, the instruments used to collect evidence, and the procedures. The purpose of the study was to focus on the use of formative assessment as a means for improving student learning. A grounded theory approach with a restricted timeframe was used to emphasize the generation of theory from the data. Teachers were subjected to a formative assessment training protocol developed based on the research found in Chapter 2. Subsequently, data was collected to determine what themes emerged as teachers responded to the training protocol and what behaviors, and/or attitudes emerged as students adjusted to being formatively assessed.

Participants

In grounded theory research, samples must include people who are in the process of experiencing the social process being studied. In keeping with this requirement, three sites were selected as Formative Assessment High Schools (FA schools). FA schools were asked to select two teachers who taught at least two classes of Algebra I to participate in the study. The researcher specifically asked for Algebra I teachers because these teachers were responsible for preparing their students to take an end-of-course (EOC) assessment.

EOC assessments are large-scale statewide assessments with high stakes associated with them. At the time of the study, the stakes associated with taking the
Algebra I EOC were the highest, in the area of mathematics, for North Carolina (NC) students. In order to receive a NC high school diploma, students were required to earn credit for Algebra I (SBE, 2006). Credit for the course was awarded based on the final course grade. An EOC assessment was administered at the end of each Algebra I course. The grade on the EOC counted at least 25% of the final course grade. Beginning with students who entered the ninth grade in 2006-07, students were required to achieve a level III or above on the Algebra I EOC as part of the exit criteria for getting a NC high school diploma.

Although the focus of the study was on Algebra I teachers and their classes, the researcher felt it was important to have district and school administrators on the team so that they could also understand the importance of formative assessment and its use. The FA schools were asked to form a team consisting of two Algebra I teachers (FA teachers), an administrator, and a district mathematics coordinator to participate in the study. All three FA schools agreed to participate in the study. The study ran for 16 consecutive weeks (see appendix E for a schedule of activities).

Six teachers – three females and three males were selected to participate in the study. The range of teaching experience was 1-30 years. Twelve classrooms (two per teacher) were selected to participate in the study. However, only four teachers and 8 classrooms remained at the end of the study.
Training Protocol

According to Wiliam (2005) formative assessment must be integrated into daily instruction and it will look slightly different in each classroom. Teachers must learn how to do this in the context of their own classrooms. Therefore, the teachers in this study were not provided a script or recipe to carry out. Instead, participants were provided formative assessment training using the training protocol. Materials from the text Classroom Assessment for Student Learning: Doing It Right – Using It Well (Stiggins et al., 2006) were adapted for use with the permission of the authors. The training was provided in four parts.

Module 1. The initial training module included the rationale for why a comprehensive balanced assessment system was needed. Focus was placed on the need for both formative and summative assessment at the classroom level. This provided a context for the study. The teachers were then introduced to the CCSSO definition of formative assessment and the research regarding the use of formative assessment. Strategies were shared with participants on how to build a collaborative environment where students are partners in assessment, how to communicate clear learning targets that are in student-friendly language, and how to provide descriptive feedback to students.

Module 2. The second module of the Training Protocol focused on the use of rubrics to help students understand the learning targets and the criteria for success.
Strategies were shared to help include students in the development of the rubrics. Sample rubrics were provided.

*Module 3.* The importance of self-assessment was the focus of this module. Samples of self-assessment instruments were shared with participants.

*Module 4.* The last module included discussions about the importance of effective questioning skills during the assessment conversations or classroom discourse. The Revised Bloom’s Taxonomy was shared in the context of formative assessment.

*Instruments*

An observation protocol was developed for use during site visits. The protocol consisted of an Observation Form (see Appendix A) and a Recognizing Formative Assessment Form (see Appendix B). The Recognizing Formative Assessment Form was adapted from the framework Ruiz-Primo & Furtak (2007) developed for identifying when assessment conversations have taken place in the classroom.

Assessment conversations embed formative assessment into any activity (Ruiz-Primo & Furtak, 2007). They allow teachers to recognize misconceptions, problem-solving strategies, language use, and communication skills. During the assessment conversation, evidence is elicited and the teacher determines where the student is in his/her learning. This information is used almost instantaneously to help promote learning.
Grounded theory research is conducted using interviews with open-ended questions and through skilled observations (Patton, 2002). In this study an interview guide approach was used. With this type of approach topics and issues are specified in advance and the interviewer decides the sequence and wording of questions during the interview (Patton, 2002). Interview Protocols based on the interview guide approach were developed for use during student and teacher interviews (see appendix C). The researcher used professional educators to assist with the camera during the interviews.

Procedure

The study was conducted over a 16-week period (see appendix B for week by week activities). Upon completion of the first module of the training protocol, the FA teachers were required to begin implementing formative assessment in the classroom. The FA teachers were required to collect artifacts related to their implementation of formative assessment.

The FA classrooms were observed using the observation protocol to verify that formative assessment was being used in the classroom during day-to-day instructional activities. During the site visits hand written notes were taken and/or classes were videotaped. Students and teachers were interviewed using the interview protocols. The students were asked to share their thoughts on how the classroom environment changed once formative assessment was introduced. They were also asked about their involvement with the assessment process, their use of
rubrics and their motivation to learn. The teachers were asked to reflect on how they thought students changed once formative assessment was introduced, the challenges associated with implementing formative assessment and the successes.
CHAPTER 4
Data Analysis and Results

This chapter is divided into two parts. The Data Analysis section focuses on the data collected during the four training modules, two rounds of site visits, artifact reviews, and interviews of select teachers and students. All data was analyzed to determine instances of formative assessment use. Interview data was compared continuously with artifact data and observations to detect emerging and themes and to direct further data collection processes. Thematic categories included direct quotes from teachers and students to capture their personal perspectives and experiences. The findings are presented in the Results section.

Data Analysis

Training Module 1: Overview

At the start of Module 1, the researcher asked participants, “What is assessment? Figure 6 shows the recorded responses.

“Evaluation of what someone knows and does not know,”
“Test scores such as Pass or Fail,”
“Tool used by stakeholders.”

Figure 6: Participant Responses to “What is assessment?”

None of the responses in Figure 6 indicated that the participants viewed assessment as something that could be used to promote learning. Terms such as “tool” and “test scores” focused on assessment as a product rather than as a process. Phrases such
as “Pass or Fail” and “knows and does not know” focused on a summative view of assessment rather than a formative view.

There was 100% agreement on the following concerns about assessment:

- Distrust of teacher judgments
- Too much time spent on measuring learning rather than helping learning to occur
- Summative assessments look too much like mini-statewide assessments
- Teachers are forced to “teach to the test”
- Students are not motivated

All of the teachers reported using multiple-choice assessments the majority of time. One participant offered this rationale, “There is so much focus on the EOC that I mostly gave multiple-choice assessments.” An analysis of teacher comments indicated that 100% believed that there was too much focus on tests at the end of the year and “staying on a pacing guide” and not enough time to “really focus on learning.” All the teachers reported feeling pressured to teach to the EOC test. This is what Heubert and Hauser (1999) referred to as “teaching to the test.” This is one of the negative consequences of high stakes assessment use (Amrein & Berliner, 2002, Battista, 1999; NCTM, 2000; Popham, 2000; Wilson, 2005).

As part of the training protocol for Module 1, the researcher presented information about the need for more formative assessment, the research on
formative assessment, the CCSSO definition of formative assessment, and the features that should be included in the formative assessment process. The researcher provided the Module 1 participants with activities to elicit evidence of understanding of formative assessment. For example, participants were asked to read *Emily’s Story* to determine how one student was able transform from a “poor writer” to a good one. As the participants described Emily’s journey they began to relate it to the three fundamental questions in Sadler’s (1989) theory of formative assessment.

Together the group determined that Emily was provided with clear and appropriate learning targets to help her answer the question, “Where am I going?” Next the group determined that Emily’s teacher shared examples of good and poor writing so Emily could “see” what was expected. This helped Emily to answer the question, “Where am I now?” Emily was given ongoing descriptive feedback to help her to answer the question, “How do I close the gap?” Participants also noted that Emily received feedback from her peers and support from her parents.

During Module 1, participants were trained on with the differences between of descriptive and evaluative feedback. Participants were then able to determine the difference between the two types of feedback when provided examples.

The FA teachers were also trained on how to breakdown learning targets into sub-targets. The FA teachers worked in groups to deconstruct an Algebra I goal and objective into smaller sub-objectives. There was much discussion about
what the targets meant and what sub-goals and sub-skills were needed to be successful on the identified targets.

The researcher asked the participants to reflect on their current practice. The participants were asked to respond to a 9-item questionnaire. The data was displayed in graphs for the group to analyze. Interestingly, the three questions that ranked the lowest were those involving student behaviors. The researcher noted that the teachers answered more favorably about teacher behaviors rather than student behaviors. Although the responses were anonymous, teachers may not have felt comfortable reporting their weaknesses in a group of peers for fear that they may be seen as nonprofessional.

In regards to teacher behaviors, the group concluded that as teachers, they guided instruction well, provided students with descriptive feedback and motivated students. In regards to student behaviors, the teachers reported that students were often not involved in the assessment process, could not describe the learning targets, or communicate about the assessments.

At the conclusion of Module 1, the researcher concluded that teachers were able to identify features of the formative assessment process in illustrative examples and correctly identify the difference between evaluative and descriptive feedback. Teachers were also able to identify the importance of communicating targets in student-friendly language and the need for involving students in the assessment process.
Training Module 2: Rubrics

During Module 2, the four FA teachers were provided information about involving students in both formative and summative assessment. Teachers were reminded that formative assessment should be part of a balanced assessment approach that includes a variety of summative assessments types.

The FA teachers were provided sample rubrics and were asked to pick one of their classes and have the students help them develop a rubric that would be used for judging their success. The other class would be used for comparison purposes. The teachers were asked to guide the students in the development and to keep the students focused on the quality of the content of the assignment. Three out of the four teachers reported that they would be willing to involve students in developing rubrics along with implementing other formative assessment strategies. One teacher reported uncertainty with implementation. The school administrator offered to assist the teacher in finding ways to implement the rubrics.

Initial Site Visits

The purpose of the site visits was to observe formative assessment in practice. The researcher trained a team of professional educators to assist with conducting the site visits. All visits were scheduled at the start of the study and confirmed a week prior to the visit.

There was no evidence of formative assessment in 3 out of the 4 classrooms during the first site visits. The researcher and trained observers witnessed teachers
engaging students in activities that could have been used to gather evidence of
learning. Although there was student involvement in class activities, there was no
evidence that teachers elicited information during instruction from the activities
used, or that instruction was adjusted based on information about student learning.
The researcher nor the trained observers found instances where teachers analyzed
thinking or provided descriptive feedback to students.

There was limited dialogue between teachers and students. Teachers
provided students with information and asked recall and procedural type questions.
Teachers provided students with reminders to avoid possible misconceptions.
When teachers asked if students understood how to do a particular problem, the
teachers redirected the students back to the steps to be performed. Teachers did not
ask students to explain their thinking. Therefore, teachers could not be certain why
students had taken alternate approaches.

In one of the classes, students showed the researcher a sheet listing the
learning targets for the unit of study. The purpose of the sheet was to help students
monitor progress. Students were required to highlight each target as it was
mastered. Students reported that this helped them keep track of what they knew and
did not know.

Training Module 3: Self-Assessment

During Module 3, the FA teachers were introduced to strategies for helping
students to assess themselves. The participants were provided with several
examples of self-assessment instruments. All four teachers committed to using self-assessment instruments in their classrooms.

Training Module 4: Effective Questioning to Elicit Understanding

All four teachers reported using self-assessment strategies in the classroom. Each teacher reported that the self-assessment instrument used was helpful to students. The teachers submitted examples of the self-assessment instruments. These artifacts were reviewed by the researcher.

During Module 4, the teachers were provided information about differentiated instruction as it related to the Revised Bloom’s Taxonomy and the use of effective questioning to elicit evidence of learning in an Algebra class. Participants were given a sample lesson on Linear Regression and asked to think of possible questions to elicit evidence of learning and how to respond to evidence that showed gaps in understanding.

Final site visits

During the final site visits, each teacher was observed and videotaped. The researcher conducted one unannounced site visit which was not videotaped. The researcher reviewed the tapes and the classroom observation notes to determine how the formative assessment process was being used and what behaviors may have manifested as a result.
Observation 1. The site visit was videotaped. The teacher used discourse as a tool to help mediate understanding. This was consistent with Vygotsky’s theory of learning. The teacher involved students in assessment conversations by first providing them with an activity to compare graphs to determine commonalities and differences. Students were asked to first reflect and then to share their analyzes with other classmates near them.

Students shared their analyses with the class. The teacher displayed both correct and incorrect responses and asked students to analyze responses based on previous mathematical knowledge. Counterexamples to hypotheses were provided by both students and the teacher. The teacher used scaffolding to help students further their understanding (see Appendix C for a transcription of the lesson and the analysis).

Observation 2. The teacher provided students with the learning target and an essential question prior to the start of the lesson. The teacher also provided the rubric (see Figure 7) that would be used to assess the students at the end of the lesson.

<table>
<thead>
<tr>
<th>Scoring Rubric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphs = 10 pts</td>
</tr>
<tr>
<td>Equations = 10 pts</td>
</tr>
<tr>
<td>Questions = 25 pts</td>
</tr>
<tr>
<td>Presentation = 5 pts</td>
</tr>
</tbody>
</table>

Figure 7: Rubric Displayed During Observation 2.
This was an example of the teacher providing students with information about how they would be assessed. However, the teacher did not provide examples or non-examples of quality work.

Students worked in cooperative groups conducting experiments and graphing data. Students were engaged in the activity. The teacher moved around the room reminding students of procedures that should be followed. The researcher did not observe any examples of the teacher pointing out learning gaps or the teacher attempting to close learning gaps related to mathematics. The majority of the teacher and student interactions were the procedures that needed to be performed to complete the assignment (e.g., making sure all materials were available, clarifying how many graphs were needed).

Part of the assignment was to do a presentation to the class. Each group was expected to answer questions regarding their findings from the experiment. The teacher explained that during the presentation, students would be asked to justify their answers.

Observation 3. The teacher assigned three problems to students and asked select students to put their solutions on the board. This activity was used to elicit evidence. Students put problems on the board while the teacher circulated around the room. Once students finished, the teacher stated, ”Let’s look at number 1. Is there anyone who has questions about number 1? [The solution] is correct. Is there anyone who did not get this?”
According to Fisher and Frey (2007), these types of general questions are not sufficient for eliciting evidence of learning. Most learners do not respond or respond in the affirmative to general questions such as the ones the teacher asked. When learners are silent, the teacher does not know if the students understand or are too confused to answer. In addition, students who do respond in the affirmative may not be aware of what they do or do not understand.

The teacher did point out a common mistake in the second problem based on observations of student work. This was done to provide advice to the students. However, there was no discussion or probing to check for the students’ understandings.

Later students worked on individualized lessons on the computer. Each lesson consisted of a lecture, guided practice, and a multiple-choice quiz. The students wore headphones while working on the computer. The researcher noticed that the students had a self-assessment checklist available to them to use during the quiz. For each question on the quiz, the students were instructed to check a box indicating whether they understood the question or not prior to the correct response being displayed on the computer screen. When applicable, students were to circle why they responded incorrectly to an item. The observer randomly asked a few students about the checklist, each student said the sheet was helpful in helping to understand whether or not the information was understood. The students also stated that the teacher would review the checklist and assist as needed.
Observation 4. Vignettes are a powerful way to render field observations (Miles & Huberman, 1994). The formative assessment process described in Figure 8 included the use of pair-share thinking. This was consistent with John Dewey’s belief that students should be active participants who work together to organize, digest, and process information (Marlowe and Page, 2005).

Vignette: Structured Pair-Work

Each student was given an appointment clock and was required to make an appointment with three other students for discussion later in the lesson. Once all the appointments have been made the teacher begins the lesson, providing information and posing questions that require higher-order thinking about the information. The students are asked to reflect on the information and to answer specific questions. Then the students go to their first appointment and spend approximately 15 minutes sharing their thinking as it relates to one or two of the posed questions. They analyze each other’s responses and come to consensus. As the students work with their partners, the teacher walks around and notes common misunderstandings and gaps in understanding. At the conclusion of the first appointment, the teacher uses the information gained during the informal observations to help redirect thinking, to reinforce ideas, and to provide cues that would help advance student thinking. The students then go to their next appointment and class continues in this manner until all appointments have been met and all questions have been discussed.

This is an example of formative assessment where the posed questions and the peer conversations are used to elicit evidence of the students' understandings. In this context, the formative assessment process is embedded into the learning activity itself due to the teacher’s careful engineering of the activity. The students are able to self-reflect and get feedback from their peers. The teacher is able to listen to the conversations between students to note the current level of understanding for the class and for individual students. The teacher uses the information immediately to assist students in their learning by redirecting thinking, reinforcing ideas, or providing cues.

Figure 8: Formative Assessment Implemented During Observation 4.
The vignette was included in the *Formative Assessment: Examples of Practice: A work product initiated and led by Carolyn Wyle, ETS for the CCSSO Formative Assessment for Students and Teachers (FAST) SCASS* published in April 2008 (see Appendix H) that the author/researcher helped to edit and coordinate.

*Observation 5.* When the researcher arrived for the unannounced visit, the teacher had a multiple-choice question on the overhead. Students were working out the problem and submitting their answers via an electronic device. Once the answer choices were submitted, the results were displayed. The teacher then proceeded to ask the students why they chose one answer over another or why a student might choose a particular answer. The students were asked to listen to all explanations and discuss the pros and cons based on the mathematical concepts they had learned. The students were then asked to reach consensus.

When interviewed students responded favorably to this type of process as evidenced by the following statement.

Student: Everyone was going against each other. It was like a race. I mean, it was like I got this answer, but he got this answer. And I am like, “I know mine has to be right.” And, he’s like, “I know mine is right.” And as soon as [the teacher] gives the answer, you are like “yeah, I got it!”
Emerging Themes, Behaviors, and Attitudes

The researcher reviewed written notes, artifacts, and observation tapes to identify emerging themes, behaviors, and attitudes. The teacher and student interviews were used to verify those themes, behaviors and attitudes. The literature was continuously reviewed as data was analyzed.

Each theme is identified followed by supporting interview statements and literature references. Statements are marked either “teacher” or “student.” No attempts were made to associate each comment with a particular teacher or student. Embedded within each theme are the new behaviors and attitudes exhibited by the teachers and students resulting from the implementation of formative assessment.

Theme 1: Dialogic vs. Univocal Discourse. Evidence showed that 100% of the participants were able to implement the use of discourse to elicit evidence of learning. During the initial observations, there was limited discourse between teachers and students. The discourse that was observed was univocal. In other words, it was used passively as a way of conveying information between the teacher and student (Wertsch and Toma, 1995).

The researcher noted that teachers used questioning to determine if the information they had provided students was received. If there was a discrepancy, the teacher attempted to adjust this breakdown in communication. Therefore, questions that arose as a result of a evidence tended to align to the teacher’s original instructions or comments.
In contrast, during the final visits, the researcher noted a move toward using dialogic discourse as a mechanism for gathering evidence about thinking and learning. In essence, discourse is dialogic when it becomes a starting point for making sense of an idea or constructing new ideas (Wertsch & Toma, 1995). Dialogic discourse has been shown to help teachers improve instructional practice (Blanton, Berenson, & Norwood, 2001).

Dialogic discourse was observed during final site visits in observations 1, 4, and 5. During these observations, participants actively interpreted information by questioning, validating, or even rejecting it. Rather than just checking to determine if the information transmitted was received, the teacher’s questions and feedback served to generate new meaning for the students. Overall, evidence showed that 2 out of 4 of the teachers moved to using discourse dialogically to elicit evidence of student learning.

During the interviews, teachers had the following to say in regards to discourse and questioning.

Teacher: I probably ask why a little bit more than I did before. I always try to lead [students] from step to step. What would we do next and then why did you go this way. It gives me a better opportunity to allow for different methods for approaching the same problem.
Teacher: I structure my instruction around where ever
[the students] are at the time. Because I do
use the questioning method when I teach, I
usually ask the questions and I formulate what
I am going to have to respond based on
whatever it is that they answer.

During interviews with students the use of questioning in this manner was
confirmed.

Student: When we did not understand a problem, [the
teacher] asked where our exact mistakes were.

According to Bruner (1975), when teachers engage students in dialogue as
described above, development is altered. The dialogue provides hints and props that
allow the child to climb the scaffolding that leads to new insights and new inquiry.

Theme 2: Collaboration vs. Power. During the initial observations, the
classes were teacher-centered. The teachers had control of the information and the
classroom. The students were well behaved and willing to participate. However,
teachers reported that using formative assessment resulted in students being more
involved in the learning process as evidenced by the following statements.

Teacher: They are absolutely more involved in the
classroom when you do it that
way…formatively and when they are
responsible for the learning. They do become more involved and it is harder work for them and it is harder work for you.

This involvement lead to students being more engaged in the learning and taking more ownership for the class. With this ownership came more work for both the teacher and student. However, it also resulted in a changes in student behaviors as evidenced by the following statements.

Teacher: I think that in my situation my students became a lot more comfortable discussing in class because it became more their class. They challenged me a lot. You know, I didn’t care… I did not mind them challenging. It made them think more about the concepts we were doing at the time.

Students also noticed changes in their teacher as a result of formative assessment being implemented in the classroom.

Student: From the beginning (the teacher) hardly ever worked with us but now she goes over [the concepts] more.

The evidence suggested that the teachers were developing a more collaborative, open classroom where students felt comfortable asking and answering questions.
The following statement supported this finding.

Student: You felt like [the teacher] was open to going

over what you didn’t understand and going

over the wrong answers that you got.

Theme 3: Comments vs. Right/Wrong. During the implementation of

formative assessment students found that the comments they received were more

helpful than just being told they got the answer right or wrong. One teacher

described how this was implemented.

Teacher: I try to get [my students] to see those things

that they are doing wrong. [I have them]

explain why they followed whatever process

they followed. If it is a written assessment, I

try to point out exactly

what it is they did wrong in the process….I do

write notes on papers.

Students supported this when interviewed.

Student: During traditional, when you looked at a

check you don’t know what you did wrong.

With this one, it is easier so that you know

what you did. When we miss it [the teacher]

puts information on there to help us
understand what we did wrong so that we
could do better on the next test.

The comments described here were consistent with Butler’s (1987) research which found descriptive feedback (comments) was more effective than evaluative feedback (right/wrong). Three out of four of the teachers reported giving students descriptive feedback. This could only be verified for one of the four teachers using additional data from the observations, artifacts, and student interviews.

**Theme 4: Trying something new vs. Keeping with the status quo.** All the teachers were willing to try something new. One of the teachers using formative assessment sparingly while the other three attempted to use it daily. The teacher explained why formative assessment was used sparingly.

Teacher: It was too late in the year for me…. There might be some things that I will change next year. But everything was set in stone so much that it would have been too much for the kids to change over. [The students] were so used to doing it one way that when you get them used to doing it one way it is hard to switch over.

As noted under Themes 1 and 2, teachers who attempted to use formative assessment daily, began to experienced the power of students’ ideas through dialogic discourse and the student roles changed. Students took more ownership
and became partners in the learning process. As the classroom environment changed, teachers had to be willing to adapt their pedagogy and beliefs to this new environment. Teachers had to be better prepared for class in regards to the content. The importance of teacher content knowledge was essential in being successful in this new environment.

Teacher: When you give [students] control, you are going to have to be on top of your game.

Teacher: [Teachers] need to make sure that they are as well versed on the subject matter as they can be. Because you do not know which way a student is going to come at you. You have to be prepared for a lot. If you were in control, you could much better know what you are going to do from step to step.

These acknowledgements are consistent with Heritage’s (2007) model of formative assessment where content knowledge is an important factor in being able to effectively implement formative assessment.

Teachers also reported using less multiple-choice items to assess students. One student expressed appreciation for the teacher using assessment methods that went beyond multiple-choice assessments in the following statement about constructed response items.
Student: The thing about not using multiple choice is that it is better. It is better because that way you know how to work out the problem. If you have multiple choice you can say, “Well, I got close to this [answer choice] so I can guess this [answer choice].” If you work [the problem] out you have to get the exact right answer so that will help you on the [EOC] test.”

Theme 5: Student-developed vs. Teacher-developed criteria. In a comprehensive balanced assessment system, a variety of techniques should be used. Rubrics can be used to clearly identify learning targets and criteria for success. In addition, they can also provide guidance as the teacher formatively assesses students prior to using the rubric as a summative measure. Three out of the four teachers implemented student-developed rubrics in their classroom. The teacher who did not use student-developed rubrics felt that class structure did not permit rubric use. 100% of the teachers who implemented the student-developed rubrics reported that their use was successful. Some of the comments were related to the criteria chosen.
Teacher: They were a lot more strict on some items
then I would have been and there were some
things that were more loose.

Teacher: They picked up on things I wouldn’t have
such as if an assignment was tardy, neatness
(even though it may seem small, neatness
was a priority), showing work rather than
just giving the answer. Showing work is
always a good thing.

The benefits to students was also discussed.

Teacher: They owned their own learning and they got
to see what was important. They got to decide
what was important. They really did think
about it. They did not slack off on the job.

When teachers were asked if they would use student-developed rubrics again, here
were some of the responses.

Teacher: I definitely would. When they finally gave me
the project they gave it to me with a little
more respect and a little more insight on what
would be graded.
Teacher: Definitely. I found it really boosted their moral as far as grading goes. Some of my students were down in the dumps as far as, “here comes a test.” “I don’t want to see that.” Now, it’s like, “what did I make?” Or, “I’ll answer that one.”

Students reported the following about developing and using the rubrics that they helped develop.

Student: On a recent project we made up our own rubric. We still had to do the work but it was more laid back. Instead of getting points off for what you did wrong, you got points for what you tried to do, what you attempted to do. [The teacher] came back and let you know what you did wrong and everything.

Evidence suggested that the rubric helped inform students about the expectations so that grading was more transparent.

Student: We got to make up our own rubric so we made it so we knew what we were looking for unlike traditional grading so we knew exactly what to do so we could get a better grade.
Theme 6: Peer- and self-assessment vs. teacher assessment. Peer-assessment was used by 3 out of 4 of the teachers and self-assessment was implemented in 100% of the classrooms. One teacher reported that although not all students used the self-assessment instrument implemented, those students who did use it did much better on the EOC assessment. This was self-reported and there was no empirical evidence to support this statement.

Students did report that peer-assessment was beneficial for helping them compare and analyze solutions as described below.

Student: All of us may have different answers or two of us may have the same answer. We see what they did in their work and they see what we did in our work and we compare answers and we figure out what went wrong between the two answers.

This evidence illustrated how Vygotsky’s (1978) Zone of Proximal Development does not only apply to teachers working with students but also students helping each other. The following evidence supported the finding that students took responsibility for their own learning and learned how to use their peers as resources.

Student: It’s like, if we are doing something new.

Some people might not get it and [the teacher]
doesn’t have time to come to everyone at the same time. So, the people who do get it come to the people who don’t understand it. And, the person who is explaining it gets a better understanding by explaining it to the other person.

Results

Although past research has shown that students who are formatively assessed achieve better on standardized assessments, this was not explored in this study. Instead, the focus was on emerging themes, behaviors, and views about assessment. Based on the data collected there was enough evidence to support the finding that new views about assessment emerged and new student and teacher behaviors were exhibited.

The evidence showed that 100% of the four FA teachers reported implementing formative assessment in the classroom. All teachers were observed doing something different from what was done in the beginning of the study. However, the data could only confirm that 3 out of 4 of the teachers attempted to implement formative assessment on a daily basis.

The three teachers who attempted implementation of formative assessment on a daily basis, reported changes in their perception of roles and pedagogy. They moved from a view of teacher in control to one of collaboration. These teachers
looked for ways to include students in assessment and learning processes. They also engaged students in developing criteria for success through the use of student-developed rubrics, used self-assessment instruments to help students reflect on learning targets, allowed students to monitor their own progress, and used discourse to elicit evidence of learning.

Students and teachers reported that their classrooms became more open and collaborative. The data suggested that a sociocultural environment emerged as a result of the use of formative assessment. In this new environment, evidence supported the finding that students became more involved in the learning process and took more ownership of their learning. According to Bandura (1992), people with high self-efficacy develop deep interest in the activities in which they participate and form a strong sense of commitment to their interests and activities. Students became more engaged in the process by justifying answers, and analyzing solutions. They also participated in more peer and self-assessment.

The students enjoyed being involved in developing rubrics and appreciated having more transparency in the grading process. Teachers reported that students had more positive attitudes in their classes. This was exhibited by students being more respectful when providing assignments to teachers, doing more homework, and making more positive comments about assignments.

The researcher concluded that teacher’s views about assessment were changed to become more inclusive of students as partners in the process. Students’
self-efficacy increased as evidenced by their increased commitments to the learning process and levels of engagement.

During the culmination of this research, the author also coordinated with members of the CCSSO FAST SCASS to review the literature to identify five attributes of effective formative assessment. The five attributes were published in the document, *Attributes of Effective Formative Assessment: A work product coordinated by Sarah McManus for the Formative Assessment for Teachers and Students (FAST) Collaborative* (see Appendix G). The five attributes are

1. Learning progressions,  
2. Learning goals and criteria for success,  
3. Descriptive feedback,  
4. Peer-and self-assessment, and  
5. Collaboration.

These attributes are consistent with the themes, behaviors, and attitudes that emerged from the data. The findings from this study provide additional support for these attributes.
In the age of a global economy, education leaders are looking for ways to improve student achievement. The goal is to produce self-directed learners who can compete in the 21st Century. For example, in 2006, the North Carolina State Board of Education adopted a guiding mission to have every public school student graduate globally competitive for work or postsecondary education and prepared for life in the 21st century. To achieve this mission, students must be measured with an assessment system that “informs instruction and evaluates knowledge, skills, performance, and dispositions needed in the 21st century” (SBE, 2007).

It is the author’s opinion that formative assessment must be a part of a 21st century assessment system. Formative assessment, as in this study, should be defined as a process. Based on the results of this study, the researcher identified four steps that must exist in the process.

*Step 1. Identification of learning targets and criteria for success.* This identification can be done by the teacher or the teacher and students together. At first thought, teachers may reject the idea of involving students in developing rubrics or having students decide how they should be graded. Some may argue that students do not have the knowledge and skills to understand what they are supposed to learn. Therefore, students would not be able to effectively assist in the assessment process. This research challenges that notion. At the beginning of this
study, teachers were skeptical about having students involved in the assessment process. However, 100% of the teachers who had students develop rubrics found that it was successful. This was captured by the comment, “It was hard for me to give up that power. But, they did a good job.”

**Step 2. Elicitation of Evidence of Learning.** It is important that students are taught something before evidence of learning can be elicited. Therefore this step occurs after a teaching episode but during the learning process. Some may argue that formative assessment and teaching are one in the same. However, the data showed that there were observations of teaching without elicitation of learning. Discourse was one of the main ways teachers in this study elicited evidence of learning.

**Step 3. Recognition of a Gap in Understanding.** It is important that teachers and students recognize when there is a gap in understanding the learning target. This is different from recognizing gaps in understanding procedures or instructions for an activity that is done in class as described in observation 2. Although recognizing these gaps in communication may be important, they will not close gaps in understanding of mathematical content.

**Step 4. Implementation of Action(s) to Close the Gap.** When reviewing the literature there was not consensus around whether a tool/test/process could be called formative if action was not taken to close the gap in understanding. As a result of the data, the author argues that action must be taken. Without action, gaps
in understanding will not be closed and student learning is not impacted. Action is most often exhibited in the use of descriptive feedback to students. As students explained in the previous section, action such as comments help them to understand how to do better the next time. More importantly, it showed them that the teacher cared about their learning.

In addition to these four steps in the process of formative assessment, the author believes the data supports the ascertain that there are three essential elements that must be present for effective implementation of formative assessment in high school mathematics classrooms. The four steps of formative assessment and three elements are inter- and intra-related.

*Element 1: Teachers must have a high level of content and pedagogical content knowledge.* Without appropriate content knowledge, teachers will not be able to effectively provide learning targets and criteria for success (Step 1) or recognize gaps in understanding (Step 3). Without appropriate pedagogical content knowledge, teachers will not be able to provide activities that elicit evidence of learning (Step 2) or know what actions are needed to close learning gaps (Step 4). More importantly, the teacher must know the depth of knowledge needed to achieve the learning targets and how best to help students to reorganize when misconceptions occur.
Element 2. The classroom environment must be an open, trusting environment where students are partners in learning process. Students must believe that the teacher is open to working with them and helping them to learn. This type of open environment is difficult to develop when students are constantly being graded. Many low achievers are reluctant to answer questions or answer in the affirmative when asked if they understand a concept (Fisher & Frey, 2007). Many are too afraid they will be embarrassed by their peers if they ask a question or admit they do not understand a concept. When students are formatively assessed they are not afraid to show their lack of understanding. They take risks and break their silence by becoming actively engaged in the dialogue of the content.

Element 3. Discourse about mathematical content must be dialogic. By being active participants in the language of mathematics, students can move from being novice learners to being experts. They will be able to provide counterexamples, analyze the veracity of mathematical arguments, and justify solutions. Teachers must have appropriate content knowledge to recognize different ways of thinking (Element 1) and must be willing to allow students to be partners in the learning process (Element 2).

If students are to be self-directed learners, perhaps it is time for more teachers to give up some of their power and control and transfer this to students. This may just be what is needed to turn around some of the students who are at risk of failure and are seen as unmotivated and low performing. As one teacher
commented, “…they are learning more. They are doing their homework. They are participating more. It is kind of a turnaround in the classroom.”

Recommendations

This study was conducted in schools that had low performance as measured by high stakes end-of-course assessments. The low performance and demands for dramatic changes from the state created an environment where there was tremendous pressure for teachers to have high test scores. Conducting research in this type of an environment can pose challenging. For example, it is not the ideal environment for conducting empirical research because if all classes need an intervention it is difficult to limit the intervention to one class. In addition, one must be careful to ensure teachers do not expect dramatic gains in a short period of time. Because it takes time to gain the skill set for successful implementation of formative assessment, caution should be taken in this area. For example, this study did not look at end-of-course test data. Instead, the focus was on emerging themes and new behaviors and attitudes.

Regardless of the school environment where formative assessment research is conducted, the researcher recommends getting the support of the school and district administration. School policies (e.g., grading policies, pacing guides) may exist that constrain the use of formative assessment. The school administrators and district coordinators in this study attended the professional development meetings and supported the teachers.
As Wiliam (2005) suggests and this research supports, teachers need ongoing professional development when implementing formative assessment. Teachers should work in learning communities where they can share strategies, successes, and challenges. For this study, the four FA teachers, administrators and district mathematics coordinator served as the learning team with the researcher as the facilitator.

In conducting research in the area of formative assessment, one must keep in mind that teachers beliefs may also need to be challenged. As one participant in this study stated, “When you get up to 9, 10 years of teaching, you have your way and you don’t want to do it any other way.” The researcher recommends providing teachers with a summary of the rationale for the need for formative assessment as presented in Chapter 1, and research similar to that presented in Chapter 2.

Once teachers understand the need and importance of formative assessment, teachers have to be convinced that formative assessment is worth the effort and not just extra work. With all the pressure of high stakes testing, this may be difficult. However, if they take on the challenge hopefully they will find as one participant stated, “[Formative assessment] does take longer. It takes longer because you are allowing [students] more of an opportunity to talk to you about what they are doing. But, in the long run, it is time well spent.”
Conclusion

This research suggests that the use of formative assessment is one way to improve student self-efficacy and teacher views about assessment in the mathematics classroom. Formative assessment is not a magic bullet. It will not compensate for inadequate curriculums, unsound instructional practices, or students’ lack of prerequisite skills. However, based on this study, it is the author’s belief that if the three essential elements are present and the four identified steps are followed during the process, formative assessment will increase student motivation and attitudes. In addition, students will accept more ownership for learning. This, in turn, will provide the students with the self-efficacy they need to be globally competitive in the 21st century.
REFERENCES


Kahl, S. (2005, October 26). Where in the world are formative tests? Right under your nose! *Education Week*.


Appendix A. Classroom Observation Form

Date ____________________

School _______________________           Teacher ________________________

Class Period ___________________          Subject ________________________

Learning Targets:
__________________________________________________________________

Are the learning targets presented in student friendly language? Why do you think this way?
__________________________________________________________________

__________________________________________________________________

Do students understand the learning targets? Why do you feel this way?
__________________________________________________________________

__________________________________________________________________

__________________________________________________________________

Describe the formative assessment practices that were used during instruction (Give a description that includes who it was used by (teacher/student/both), how it was used and how effective it seemed to be:
__________________________________________________________________

__________________________________________________________________

__________________________________________________________________

__________________________________________________________________

__________________________________________________________________

__________________________________________________________________

__________________________________________________________________

__________________________________________________________________
Describe formative assessment practices that could have been used during instruction that would help the teacher and students understand where they are in the learning process. (Give a description that includes who could have been used by (teacher/student/both), how it could have been used and why you think it would have been effective):

___________________________________________________________________
___________________________________________________________________
___________________________________________________________________
___________________________________________________________________
___________________________________________________________________

Please meet with the teacher and give him/her an opportunity to reflect on the lesson first and then share your observations with the teacher.

**Additional Information Needed During the Site Visit**

1. When will the current unit/chapter be tested? ______________

2. What is the next unit/chapter of study? ______________

3. When will the next unit/chapter of study be started? ______________

4. When will the next unit/chapter be tested? ______________

5. Describe how the learning team has been implemented and the teacher’s role.
   ___________________________________________________________________
   ___________________________________________________________________
6. Has the teacher had an opportunity to involve students in rubric development? If so, how and have him/her describe impact on the students thus far.

____________________________________________________________________
____________________________________________________________________

7. If not, when will this take place?

____________________________________________________________________
____________________________________________________________________

8. What can the Formative Assessment Project team do to assist the teacher in the future?

____________________________________________________________________
____________________________________________________________________
### Recognizing Formative Assessment
Adapted from Ruiz-Primo and Furtak (2007)

<table>
<thead>
<tr>
<th>Eliciting Evidence</th>
<th>Identifying the Gap</th>
<th>Attempting to Close the Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students share observations</td>
<td>Compare/contrast responses with mathematical rules, postulates, theorems</td>
<td>Promotes argumentation/debate among ideas</td>
</tr>
<tr>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
</tr>
<tr>
<td>Evidence:</td>
<td>Evidence:</td>
<td>Evidence:</td>
</tr>
<tr>
<td>Students elaborate on their responses (why, how)</td>
<td>Explores students ideas</td>
<td>Provides descriptive feedback</td>
</tr>
<tr>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
</tr>
<tr>
<td>Evidence:</td>
<td>Evidence:</td>
<td>Evidence:</td>
</tr>
<tr>
<td>Students present problems to the class</td>
<td>Captures/displays students ideas</td>
<td>Have students look back at definitions, rules, etc</td>
</tr>
<tr>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
</tr>
<tr>
<td>Evidence:</td>
<td>Evidence:</td>
<td>Evidence:</td>
</tr>
<tr>
<td>Students provide representations</td>
<td>Students explain representations</td>
<td>Have groups of students to compare and contrast ideas</td>
</tr>
<tr>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
</tr>
<tr>
<td>Evidence:</td>
<td>Evidence:</td>
<td>Evidence:</td>
</tr>
<tr>
<td>Interpret data or patterns</td>
<td>Compare results with others</td>
<td>Justify each step</td>
</tr>
<tr>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
</tr>
<tr>
<td>Evidence:</td>
<td>Evidence:</td>
<td>Evidence:</td>
</tr>
<tr>
<td>Evaluate the quality of a solution</td>
<td>Look for misconceptions, analyze use of language</td>
<td>Review criteria and self-assess</td>
</tr>
<tr>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
</tr>
<tr>
<td>Evidence:</td>
<td>Evidence:</td>
<td>Evidence:</td>
</tr>
<tr>
<td>Takes a vote on which solution is the best</td>
<td>Have students explain why they voted a particular way</td>
<td>Promote consensus</td>
</tr>
<tr>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
<td>Observed: yes/no</td>
</tr>
<tr>
<td>Evidence:</td>
<td>Evidence:</td>
<td>Evidence:</td>
</tr>
</tbody>
</table>
Appendix C. Transcription and Analysis of Final Site Visit - Observation 1.

Throughout this description, discourse is provided regarding the formative assessment observed. Line numbers are used to distinguish each line of recorded discourse. A “T” is used to represent the teacher talk and an “S” is used to represent the student talk. There was no attempt to attribute the talk to any particular student. If the two consecutive lines of student talk were recorded, the student talk is numbered (e.g., S1, S2).

The teacher provided students with equations to graph. The students were asked compare their graphs to determine commonalities and difference among the equations. Students were asked to first reflect and then to share their analyses with other classmates near them. Students worked independently while the teacher circulated. When one student stated that she was confused, the teacher tried to elicit additional information.

1 T: Why are you confused?
2 T: What does increasing mean to you? Start looking to see which ones are increasing.
3 T: Discuss what you are seeing.
4 S: It goes up but it is negative.
5 T: What does that mean?
6 S: It is increasing.
In this exchange, the teacher tried to pull additional information from the student (see lines 2, 3, 7, and 8). In line 2, the teacher’s feedback to the student was to look at the vocabulary (i.e., to think about the meaning of the word “increasing”). The teacher provided the student with advice to try to determine which graphs were increasing. In line 6 when the student states that the graph is increasing, the teacher again asked the student to explain the response. This was done to elicit evidence about whether or not the student understood what increasing meant in the context of the graphs.

The teacher focused on getting the student to achieve the learning target and not just providing a correct answer. This was illustrated in line 8 when the teacher again provides the student with feedback that would help the student get closer to achieving the learning target. The teacher advised the student to graph the equations and to apply vocabulary to determine which graphs were increasing or decreasing.

During a whole group discussion, the teacher asked the students what they noticed about the graphs. The following is some of the dialogue recorded.

9 S: All cross the y-axis at a positive number.

10 T: Why?

11 S: Because they have an exponent.
107

12 T: What do you think about his answer? Is it true?
13 T: What about \( y = x^2 \)? Does that have an exponent?
14 S1: It is exponential?
15 S2: No, it is a quadratic.
16 T: How is \( y = x^2 \) different from these?
17 S: Is it because the ones we have, have a number as the base
and the exponent as the variable?
18 Teacher writes the equation, \( Y = 3^x \) on the board.
19 T: Is this true?
20 T: Think about it and share your thinking with someone else
and then we will discuss.
21 T: How are the curves different?

In this exchange, the teacher tried to get the students to describe the characteristics of the graphs of the equations provided to them. As the students offered explanations, the teacher asked the class to analyze the responses to determine their correctness (lines 12 and 19). The teacher asked probing questions to elicit evidence of understanding and provided a counterexample (line 13). When students were silent, the teacher continued probing and suggested the students reflect and partner with someone to share thinking (line 20).

Later in the class, a student noticed a similarity with some of the graphs. The following discourse occurred.
S: All fractions decrease.

T: What do you mean?

S: \((\frac{1}{2})^x\) and \(.7^x\)

One of the student’s peer offered a counterexample.

S2: No, because \((\frac{5}{3})^x\) increases

According to von Glasersfeld, providing this information should result in a perturbation in which students must assimilate and accommodate the new information into their existing cognitive structures.

T: So, what is going on?

T: Look at the patterns at it applies to the equation.

The teacher provided feedback or advice to the students without providing evaluative information. This allowed students to extend their thinking.

S1: They are not whole numbers.

S2: I know…but I can’t say it.

S3: The number is not 6 or 7.

S4: The numerator is larger than the denominator.

T: What does that mean?

S: It is an improper fraction.

T: What does that mean?

S: You can turn it into a mixed number.

T: What does that mean? What about the amount?
As students began to focus on how improper fractions could be written as mixed numbers (lines 36), the teacher redirected thinking. The teacher suggests that the students look at the amount of the fraction represents.

38 S: You can…Smaller than a whole number? Anything lower than 1 decreases…anything greater increases!

39 T: Let’s try it.

The teacher does not provide an evaluative comment (line 39) but instead asks the class to investigate the hypothesis in line 38. The class seems satisfied with the student’s explanation that “Anything lower than 1 decreases…anything greater increases!”

40 S: I think we need a pop quiz.

At least one student feels very confident that the class has successfully found one meaningful characteristic of the graphs and asks for a pop quiz. The teacher continued to probe for more understandings about characteristics of graphs.

Overall, the teacher was observed using formative assessment by eliciting evidence of learning through the classroom discourse and a learning activity. The teacher and students offered counterexamples to help extend understandings based on the evidence. Finally, the discourse was used to help students close gaps in understanding.
Appendix D. Interview Protocols

Protocols are based on an interview guide approach.

Teacher Interview Protocol

Get the teachers' consent to be interviewed.

Ask the teacher a question about how he/she is implementing formative assessment. Probe to find out about feedback, use of questioning, student-developed rubrics, and other types of assessment techniques.

Ask the teacher to describe the impact the techniques had on their students and on their teaching.

Student Interview Protocol

Only pull students who have a consent form on file. Get the students’ verbal consent to be interviewed.

Ask students if they have noticed changes in the classroom over the course of the semester. Probe about the types of questioning, feedback, grading, and assessment. Ask the students what changes the teacher has made and the impact on them.
Teacher Interview 1

You have been implementing formative assessment in the classroom this semester. Tell me how that has been going for you.

I think it has been going pretty well. What I have noticed is that my students…I give them an opportunity or I structure my instruction around whatever it is that they…where ever they are at the time. Because I do use the questioning method when I teach I usually ask the questions and I formulate what I am going to have to respond based on whatever it is that they answer.

Can you describe the type of feedback that you give your students and their response to that.

Yes. I try to get them to see those things that they are doing wrong. Explain why they followed whatever process they followed. If it is a written assessment I try to point out exactly what it is they did wrong in the process. If they go back to the first step, where did they leave out a negative sign and I do write notes on papers.

Have you implemented any rubrics in the classroom?

I have used a rubric. In fact, I allowed my Algebra II students to come up with a rubric for their project that they did recently. What I found out from them was what things they valued. I think there were some things they thought were more important than those that I thought.

How did it help them with doing the assignment if they helped you develop the rubric?

They developed a rubric that they were comfortable with. They paid more attention to what was on the rubric so the assignment looked more like what they would have wanted it to look like. In the end it was very good.
You talked about questioning a little bit earlier. How has your class
discussions changed since you have been using formative assessment?

I probably ask why a little bit more than I did before. I always try to lead them from
step to step. What would we do next and then why did you go this way. It gives me
a better opportunity to allow for different methods for approaching the same
problem.

What would you say is the advantage and disadvantage of using formative
assessment?

I think that in my situation my students became a lot more comfortable discussing
in class because it became more their class. They challenged me a lot. You know, I
didn’t care… I did not mind them challenging. It made them think more about
whatever it was, the concepts we were doing at the time.

Did you see any disadvantages of using formative assessment? Any challenges?

Well I’m not sure if it was a disadvantage. Of course, it does take longer. It takes
longer because you are allowing them more of an opportunity to talk to you about
what they are doing. But In the long run it is time well spent.

What advice would you have for a teacher who is just going to start out using
formative assessment in the classroom? What would you tell the teacher?

They need to make sure that they are as well versed on the subject matter as they
can be. Because you do not know which way a student is going to come at you.
You have to be prepared for a lot. If you were in control, you could much better
know what you are going to do from step to step.

When you give them that control, you are going to have to be on top of your game.

You mentioned control. Do you think students are more involved in the
classroom?

They are absolutely more involved in the classroom when you do it that
way…formatively and when they are responsible for the learning. They do become
more involved and it is harder work for them and it is harder work for you.
We talk about 21st century learning and trying to make sure we have students who are, self directed learners, lifelong learners. Do you think formative assessment will help your students go in that direction?

Yes, it would because it would give them an opportunity now to start looking at what they need. As a teacher do I get to direct my staff development

As students they should be in the same position. They know what it is that they need to learn. So, they are going to take more ownership of that and be successful.

Can you describe your students from before you started implementing formative assessment and afterwards?

They were more successful from the grade perspective afterwards. Their level of understanding was more than it had been.

Teacher Interview 2

What kind of changes have you seen since you have been implementing formative assessment in the classroom?

Mainly just with me doing different things with the computer. keeping them together. If I don’t like what the computer is doing then I can do it myself and get the kids interested that way.

Have you tried to put any feedback into your instruction?

Mainly just the forms I gave you all. We have done a lot of the overhead and things like that with the review sessions. They have had hardly no time on the computer these last few weeks.

How has the interaction changed between you and your students?

Some of them like it and some of them dislike it

One of the kids asked me today, when are we going to get back on the computer? I am tired of hearing you talk. For the most part, a lot of them like it.
What do they like about it?

I am focusing on now especially now on what it is specifically what exactly they need for the EOC work. We are working on that with the company because sometimes that put something that they don’t need in there.

How do you know it is what they need? How do you know that what you are working on right now is what they need?

That is according to the standard course of study. The workbook that we use really does a good job of breaking it down into sections. The computer is nationally done and some times it gives them things that they do not really need.

So it is not aligned to the standard course of study?

It will be next year because I gave them a copy of the standard course of study and I told them exactly what I need especially with the new test. It was really done well with the old test but with the new test I have had to tell them this is what I want here, here, and here. Put this in this order, put that in that order so that it can be more aligned

Have you done anything that has improved your instruction or has furthered the understanding of mathematics with your students?

I don’t know about myself because pretty much what I did that day was what I used to do when I used to get up in front of the classroom. I don’t see anything in myself that has changed unless I do it more but I have not done enough of it to get in front of the class for it to effect my work

In one of our previous meetings, you mentioned that your county used to focus on formative assessment in the past and then moved away from it. Do you know why they moved away from it?

I have no idea. I know that when I was in methods class they always aligned stuff with what the county was doing. I happened to stumble onto some of my stuff from college and there was a whole lot of formative assessment stuff and I had forgotten about that stuff. There was some rubrics and different stuff they were telling us to focus on but when I came out of school we did not worry about that anymore. It was all about the test.
Have you started to use any other forms of assessment or grading?

No, like I said it was too late in the year for me to do that. There might be some things that I will change next year. But everything was set in stone so much that it would have been too much for the kids to change over. They were so used to doing it one way that when you get them used to doing it one way it is hard to switch over.

Next year, we want to bring in the entire math faculty and train them. What kind of challenges do you think they will face in implementing formative assessment?

Being set in their ways. You have some young teachers like me. But when you get up to 9, 10 years of teaching, you have your way and you don’t want to do it any other way.

Have you shared any of the information with them?

We have only mentioned it and shown them the brochure and that is it.

Teacher Interview 3

What is formative assessment?

That would allow them to have an opportunity to have break points to check reassessments and to redo.

How did you implement formative assessment?

Not necessarily written work. I used match cards to match things and I go around to check.

I had a class that did a rubric and I graded it. Some of the things I they did.

What impact did this have on your students?

More insight into what will be graded so they knew what to do.

They were a lot more stricter on some items then I would have been and there were some things that were more loose. It was hard for me to give up that power. But they did a good job.
They owned their own learning and they got to see what was important. They got to decide what was important. They really did think about it. They did not slack off on the job.

**Would you do it again?**

I definitely would. When they finally gave me the project they gave it to me with a little more respect and a little more insight on what would be graded.

---

**Teacher Interview 4**

**What is formative assessment?**

Summative is a test, quiz, free response.

Formative is involved in the learning more, asking questions, getting feedback

Students involved in their learning and getting feedback.

**How did you implement formative assessment?**

Rubrics. Students formed their own rubrics.

**What impact did this have on your students?**

They picked up on things I wouldn’t have such as if an assignment was tardy, neatness. Even though it may seem small, neatness was a priority showing work rather than just giving the answer. Showing work is always a good thing.

**Would you do it again?**

Definitely, I found it really boosted their moral as far as grading goes. Some of my students were down in the dumps as far as, here comes a test. I don’t want to see that. Now its like, what did I make? Or, I’ll answer that one.

It seems like they are learning more. They are doing their homework. They are participating more. It is kind of a turnaround in the classroom.
Student Group Interview

At the beginning of the year, we asked your teacher to introduce you to formative assessment. Perhaps you noticed that she was doing things a little bit different and maybe you didn’t. Did you notice anything different about how she taught at the beginning of the year and how she taught towards the end of the year?

Yes, from the beginning she hardly ever worked with us but now she goes over it more. People can hardly get it so she goes over it more

Does that help you learn it better?

Yes

How has it helped you?

Cause a lot of problems I couldn’t get so we went over it again and again and again until it like clicked.

So [the teacher] went over the problems more and more and you felt like she was more open to going over what you didn’t understand and going over the wrong answers that you got.

Yes Ma’m

Did [the teacher] do that by questioning you?

Yes

So [the teacher] used questioning. What type of questioning did [the teacher] use that was effective for you as a student?

When we did not understand a problem, She asked where our exact mistakes were.

When it comes to learning, what is your role and what is the teacher’s role?

[The teacher] is responsible to teach and we are responsible to listening.
Anything else?

We are responsible for taking what we learn here and doing it at home.

I think about how I am going to use this in the field that I want to go into. Right now, I am thinking about going into the military.

[The teacher] might do a problem on the board and do it wrong on purpose and then she ask us if it is right and we say yes and then she erases it and says no here is the right way to do it

[The teacher] has us write down the question and how to do the problem. And, then at the end if we don’t think it is right, we go over to her and she shows up what we got wrong.

[The teacher] explains how like what we did wrong and how to redo that. Then she gives us the paper and asks us to rewrite the whole thing.

Is that a good thing?

Yes, because when you rewrite it over again you can when you are taking a test remember all the things you did wrong and all the things you did right so on the test you remember what you did wrong so you can get it right.

Does [the teacher] give you all the information or pieces of information?

Usually she give us pieces but some times she might give us the whole thing.

Do you ever get to assess your peers?

Yes all the time.

Describe how that works.

All of us may have different answers or two of us may have the same answer. We see what they did in their work and they see what we did in our work and we compare answers and we figure out what went wrong between the two answers.
What were you going to say?

It's like, if we are doing something new, some people might not get it and she doesn't have time to come to everyone at the same time so the people who do get it come to the people who don't understand it and the person who is explaining it gets a better understanding by explaining it to the other person.

Can you tell me what kind of assessments you like to see in this class?

I can't choose.

Does she always use multiple choice questions?

Hardly ever.

The thing about not using multiple choice is a better think that way you know how to work out the problem. If you have multiple choice you can say well, I got close to this so I can guess this. If you work it out you have to get the exact right answer so that will help you on the test.

Did you use rubrics this year?

At first I didn't get it?

What did you do?

At first I went to him (points to another student in the group) and then I went to [the teacher].

[The teacher] broke it down so I could get it.

What about it did you not understand?

Where a plus is supposed to be, I put a negative. She grades like this is right put you wrote the wrong equation.

When we miss it she puts information on their to help us understand what we did wrong so that we could do better on the next test.

During traditional when you looked a check you don't know what you did wrong. With this one it is easier so that you know what you did.
Did you ever help her develop a rubric? Did you ever come up with criteria that you wanted to be graded on?

On a recent project we made up our own rubric. We still had to do the work but it was more laid back. Instead of getting points off for what you did wrong, you got points for what you tried to do, what you attempted to do. She came back and let you know what you did wrong and everything.

Did you get a chance to improve it after you received the feedback?

On my own I went back and improved it later.

We got to make up our own rubric so we made it so we knew what we were looking for unlike traditional grading so we knew exactly what to do so we could get a better grade.

When we walked in you were doing a lesson. Tell me about it.
Everyone was going against each other it was like a race I mean it was like I got this answer but he got this answer and I am like I know mine has to be right and he’s like I know mine is right and as soon as she gives the answer you are like yeah I got it.

We used the responders and it came up on the computer.

Do you think it is important that it is anonymous?

It is important because some people are bashful about their answers they don’t want everyone to know they got something wrong.
Appendix F. Study Timeline

Week 1  Module 1 (Face-to-face Meeting of FA School Teams)

Week 2  Module 2 (Web meeting of FA School Teams)

Week 4  Site visits to FA schools

Week 5  Site visits to FA schools

Week 6  Telephone Correspondence (as needed)

Week 7  Module 3 (Web meeting)

Week 8  Module 4 (Face –to Face Meeting of FA schools)

Week 11 Site visits to FA schools (videotaped)

Week 12 Site visits to FA schools (videotaped)

Week 14 Teacher and student interviews

Week 15 Additional interviews and site visits

Week 16 Students took end-of-course assessments; All data turned in to researcher
Appendix G

*Attributes of Effective Formative Assessment*

A work product coordinated by Sarah McManus

NC Department of Public Instruction

for the Formative Assessment for Students and Teachers (FAST) Collaborative

Council of Chief State School Officers: Washington, DC 2008
ATTRIBUTES OF
EFFECTIVE FORMATIVE ASSESSMENT

A WORK PRODUCT COORDINATED¹ BY SARAH McMANUS
NC DEPARTMENT OF PUBLIC INSTRUCTION

Paper prepared for the Formative Assessment for Teachers and Students (FAST) State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers (CCSSO)

¹ Grateful thanks go to the various members of the FAST SCASS who contributed text and edits, and provided feedback on various iterations of this document.
THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS

The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

Formative Assessment for Students and Teachers State Collaborative on Assessment and Student Standards

The Council’s State Collaborative on Assessment and Student Standards (SCASS) strives to provide leadership, advocacy and service in creating and supporting effective collaborative partnerships through the collective experience and knowledge of state education personnel to develop and implement high standards and valid assessment systems that maximize educational achievement for all children.

COUNCIL OF CHIEF STATE SCHOOL OFFICERS

Rick Melmer (South Dakota), President
Elizabeth Burmaster (Wisconsin), Past President
T. Kenneth James (Arkansas), President-Elect

Gene Willhoit, Executive Director

John Turner, Director Center for Innovative Measures
Douglas Rindone and Duncan MacQuarrie, Co-Coordinators, FAST-SCASS

Council of Chief State School Officers
One Massachusetts Avenue, N.W., Suite 700
Washington, DC 20036-1431
Phone (202) 336-7000
Fax (202) 408-8672
www.cosso.org

Copyright © 2008 by the Council of Chief State School Officers, Washington, DC. All rights reserved.
ATTRIBUTES OF EFFECTIVE FORMATIVE ASSESSMENT
A work product coordinated by Sarah McManus, NC Department of Public Instruction, for the Formative Assessment for Students and Teachers (FAST) Collaborative

Background of a Definition

There has been substantial interest in formative assessment among U.S. educators during recent years. Increasing numbers of educators regard formative assessment as a way not only to improve student learning, but also to increase student scores on significant achievement examinations. To promote the use of formative assessment, the Council of Chief State School Officers (CCSSO) created a national initiative. The initiative formally began in January 2006, when CCSSO formed the Formative Assessment (FA) Advisory Group consisting of measurement and education researchers including Jim Popham, Lorrie Shepard, Rick Stiggins, and Dylan Williams and state agency leaders from across the nation. (A complete list of FA Advisory Group members is at end of document.)

CCSSO also formed a new State Collaborative on Assessment and Student Standards (SCASS) to implement the vision of the FA Advisory Group. The first challenge for the Formative Assessment for Students and Teachers (FAST) SCASS was to work with the FA Advisory Group to review the various definitions of formative assessment and related research. The FA Advisory Group and FAST SCASS devoted substantial effort to clarify the meaning of “formative assessment,” based on current literature, and determine how formative assessment may best be used by the nation’s educators.

In October 2006, FAST SCASS educators representing approximately 25 states agreed on the definition of formative assessment presented in this document and it was subsequently approved by the FA Advisory Group. In the year following, the FAST SCASS and FA Advisory Group isolated the attributes that, based on the research and current literature, would render formative assessment most effective. This document presents the definition of formative assessment and identifies and explains the five attributes of effective formative assessment.

The Definition of Formative Assessment

During the October 2006, inaugural FAST SCASS meeting in Austin, Texas, the following definition of formative assessment was adopted, without dissent:

Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes.

A Closer Look at the Definition

The primary purpose of the formative assessment process, as conceived in this definition, is to provide evidence that is used by teachers and students to inform instruction and learning during the teaching/learning process. Effective formative assessment involves collecting evidence about how student learning is progressing during the course of instruction so that necessary instructional adjustments can be made to close the gap between students' current understanding and the desired goals. Formative assessment is not an add-on to teaching but, rather, integrated into instruction and learning with teachers and students receiving frequent feedback.

One key feature of this definition is its requirement that formative assessment be regarded as a process rather than a particular kind of assessment. In other words, there is no such thing as a “formative test.” Instead, there are a number of formative assessment strategies that can be implemented during classroom instruction. These range from informal observations and conversations to purposefully planned instructional techniques designed to elicit evidence of student learning to inform and adjust instruction.

A second important part of the definition is its unequivocal requirement that the formative assessment process involve both teachers and students. The students must be actively involved in the systematic process intended to improve their learning. The process requires the teacher to share learning goals with students and provide opportunities for students to monitor their ongoing progress.

THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS
THE FAST SCASS • FORMATIVE ASSESSMENT FOR TEACHERS AND LEARNERS

125
Attributes

There are five attributes that have been identified from the literature as critical features of effective formative assessment. No one of the following attributes should be regarded as a sine qua non, that is, an attribute without which the assessment would not be formative.

1. Learning Progressions: Learning progressions should clearly articulate the sub-goals of the ultimate learning goal.

Learning progressions describe how concepts and skills build in a domain, and show the trajectory of learning along which students are expected to progress. From a learning progression teachers have the big picture of what students need to learn, as well as sufficient detail for planning instruction to meet short-term goals. They are able to connect formative assessment opportunities to the short-term goals to keep track of how well their students' learning is moving forward.

For example, at the earliest stages of a progression for historical inquiry students must learn how to investigate the past from a range of sources of information, (e.g., stories, eyewitness accounts, pictures, photographs, artifacts, historic buildings, museums, galleries, and technology-based sources). Students build on this learning in later stages of the progression to develop an understanding that people represent and interpret the past in different ways (e.g., through pictures, plays, films, reconstructions, museum displays, and fiction and nonfiction accounts), and that the interpretations reflect the intentions of those who make them (e.g., writers, archaeologists, historians, and filmmakers). A goal for students at each level of the progression would be to investigate a set of artifacts in increasingly sophisticated ways to extract information about a particular period or event in history. Not only would such investigations support the students' development of historical reasoning, they would also provide evidence of the students' ability to reason in increasingly complex ways. This involves moving from the early stages of reasoning based on simple observation to the more complex stages based on indirect observation and the synthesis of multiple sources of information. Using the evidence elicited from such tasks connected to the goals of the progression, a teacher could identify the "just right gap" — a growth point in learning that involves a step that is neither too large nor too small — and make adjustments to instruction accordingly.

2. Learning Goals and Criteria for Success: Learning goals and criteria for success should be clearly identified and communicated to students.

Because the formative assessment process helps students achieve intended learning outcomes based on explicit learning progressions, teachers must first identify and then communicate the instructional goal to students. In addition to communicating the nature of the instructional goal, teachers must provide the criteria by which learning will be assessed so that students will know whether they are successfully progressing toward the goal. This information should be communicated using language readily understood by students, and may be accompanied by realistic examples of those that meet and do not meet the criteria.

For example, suppose the goal of a social studies instructional unit was to have students "prepare a written critique of the quality of arguments in political essays in a local newspaper's editorial pages." The teacher might first offer students a paraphrased version of that goal such as, "You will be able to judge the strengths and weaknesses of arguments in the editorials you find in our daily newspapers." The teacher would discuss the criteria for evaluating arguments and then provide several examples of critiques of political essays. This will provide students with a reasonably clear idea of the analytic skills they are to develop and also provide them with the tools required to assess their own written analyses.

3. Descriptive Feedback: Students should be provided with evidence-based feedback that is linked to the intended instructional outcomes and criteria for success.

Descriptive feedback should be about the particular qualities of student learning with discussion of suggestions about what the student can do to improve. It should avoid comparisons with other pupils. Specific, timely feedback should be based on the learning goal and criteria for success. It should help the student answer three basic questions: Where am I going? Where am I now? How can I close the gap?

For example, in an eighth grade writing class the students are learning how to construct an argument. They are focusing specifically on speech writing and have examined several effective speeches, both from prominent speech-makers in history and from previous year's eighth grade students. In this particular lesson, students have been asked to write an opening paragraph to their speech with the success criteria of introducing their topic in a way that engages the audience. The feedback the teacher gives to one student is, "The opening paragraph does not capture the audience's attention because it does not clearly state what the speech is about. However, the opening sentence of the second paragraph states your position with an effective contrast. What can you do to improve or strengthen your opening paragraph?" With this kind of descriptive feedback and collaboration, the teacher clarifies the goal for the student, provides specific information about where the student is in relation to the criteria and offers enough substantive
information to allow the student an opportunity to identify ways to move learning forward.

Similarly, in a sixth grade math class, students working in groups have been asked to review an example of the steps a student from a previous year took to solve a problem. They must decide if the work is correct or incorrect and provide an explanation for their view. The success criterion that the teacher gives them is, “Include any properties or rules that may apply in your explanation.” When the groups report back after their discussion, the teacher listens for rules or properties in the explanations, and this becomes the focus of her feedback. To one group she says, “Your explanation shows me that you understand that the steps the student took to solve the problem were incorrect. Remember the success criterion. You must also relate your explanation to one of the properties we have been discussing in class to indicate the reason the steps were incorrect.” Again, the students knew the goal, where their response differed from the criteria, and how they can improve their explanations.

4. Self- and Peer-Assessment: Both self- and peer-assessment are important for providing students an opportunity to think metacognitively about their learning.

Formative assessment is a process that directly engages both teachers and students. In addition to teacher feedback, when students and their peers are involved, there are many more opportunities to share and receive feedback. Helping students think metacognitively about their own learning fosters the idea that learning is their responsibility and that they can take an active role in planning, monitoring, and evaluating their own progress. To support both self- and peer-assessment, the teacher must provide structure and support so students learn to be reflective of their own work and that of their peers, allowing them to provide meaningful and constructive feedback.

In self-assessment, students reflect on and monitor their learning using clearly articulated criteria for success. In peer-assessment, students analyze each other’s work using guidelines or rubrics and provide descriptive feedback that supports continued improvement. For example, students can work in pairs to review each other’s work to give feedback. A teacher needs to have modeled good feedback with students and talked about what acceptable and unacceptable comments look like in order to have created a safe learning environment. Students can use a rubric to provide feedback to a peer by articulating reasons why a piece of work is at one level and discussing how it could be improved to move it to the next level. Alternatively, feedback could be given using a format such as “two stars and a wish,” which provides a structure for a student to identify two aspects of the work that are particularly strong (stars) and one aspect the peer might improve (a wish). Students then need time to reflect on the feedback they have received to make changes or improvement. In addition, students can be encouraged to be self-reflective by thinking about their own work based on what they learned from giving feedback to others. A further benefit of providing feedback to a peer is that it can help deepen the student’s own learning. However, student- and peer-assessment should not be used in the formal grading process.

5. Collaboration: A classroom culture in which teachers and students are partners in learning should be established.

Sharing learning goals and criteria for success with students, supporting students as they monitor and take responsibility for their own learning, helping students to provide constructive feedback to each other, and involving students in decisions about how to move learning forward are illustrations of students and teachers working together in the teaching and learning process.

However, for students to be actively and successfully involved in their own learning, they must feel that they are bona fide partners in the learning process. This feeling is dependent on a classroom culture characterized by a sense of trust between and among students and their teachers, by norms of respect, transparency, and appreciation of differences, and by a non-threatening environment. Creating such a culture requires teachers to model these behaviors during interactions with students, to actively teach the classroom norms, and to build the students’ skills in constructive self- and peer-assessment. In this type of classroom culture, students will more likely feel that they are collaborators with their teacher and peers in the learning process.

While evidence exists in varying degrees to support the five attributes presented, there is clearly no one best way to carry out formative assessment. The way these attributes are implemented depends on the particular instructional context, the individual teacher, and—perhaps most importantly—the individual student.

For examples on how to incorporate the five attributes into practice refer to the document Formative Assessment: Examples of Practice.²

Suggested Readings


CCSSO Formative Assessment Advisory Group

Annette Bohling, AZ, Sr. Vice President of Accreditation, Advanced, Arizona State University
William Bushaw, IN, Executive Director, Phi Delta Kappa
Doug Christensen, NE, Commissioner of Education, Nebraska Department of Education
Angela Faherty, ME Deputy Associate Superintendent, Standards and Assessment Section, Maine Department of Education
Margaret Heritage, CA, Assistant Director for Professional Development, National Center for Research on Evaluation, Standards and Student Testing, University of California
Gerunda Hughes, DC, Associate Professor, Curriculum and Instruction Program Coordinator, Secondary Education, Howard University
Henry Johnson, Former Assistant Secretary for Elementary and Secondary Education, U.S. Department of Education
Stuart Kahl, NH, President and CEO, Measured Progress
Ken Kay, AZ, President, Partnership for 21st Century Skills
Sarah McManus, NC, Section Chief, Testing Policy and Operations, North Carolina Department of Public Instruction
Bob Nielsen, IL, Superintendent, Bloomington, Illinois Public Schools
Scott Norton, LA, Director, Student Standards & Assessments, Louisiana Department of Education
Jim Popham, HI, Emeritus Professor, University of California
Doris Redfield, WV, President/CEO, Edvantia
Wendy Roberts, DE, Assessment Director, Delaware Department of Education
Lorrie Shepard, CO, Dean, School of Education, Professor of Education, University of Colorado at Boulder
Rick Stiggins, OR, CEO, Assessment Training Institute (AT), ETS
Martha Thurlow, MN, Professor, National Center on Education Outcomes, University of Minnesota
Dylan Willam, UK, Deputy Director, Institute of Education, University of London

3 The CCSSO Formative Assessment Advisory Group was formed in March 2006. This is a list of the original members who were responsible for developing and approving the definition and attributes of effective formative assessment.
Appendix H

Formative Assessment: Examples of Practice

A work product initiated and led by E. Caroline Wylie

ETS

for the Formative Assessment for Students and Teachers (FAST) Collaborative

Council of Chief State School Officers: Washington, DC 2008
FORMATIVE ASSESSMENT:
EXAMPLES OF PRACTICE

A WORK PRODUCT INITIATED AND LED BY
E. CAROLINE WYLIE, ETS

Paper prepared for the Formative Assessment for Teachers and Students (FAST)
State Collaborative on Assessment and Student Standards (SCASS) of the
Council of Chief State School Officers (CCSSO)
THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS

The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

Formative Assessment for Students and Teachers State Collaborative on Assessment and Student Standards

The Council’s State Collaborative on Assessment and Student Standards (SCASS) strives to provide leadership, advocacy, and service in creating and supporting effective collaborative partnerships through the collective experience and knowledge of state education personnel to develop and implement high standards and valid assessment systems that maximize educational achievement for all children.

COUNCIL OF CHIEF STATE SCHOOL OFFICERS

Rick Melmer (South Dakota), President
Elizabeth Burmaster (Wisconsin), Past President
T. Kenneth James (Arkansas), President-Elect

Gene Wilhoit, Executive Director

John Tanner, Director Center for Innovative Measures
Douglas Rindone and Duncan MacQuarrie, Co-Coordinates, FAST SCASS

Copyright © 2008 by the Council of Chief State School Officers, Washington, DC
All rights reserved

131
FORMATIVE ASSESSMENT: EXAMPLES OF PRACTICE
A WORK PRODUCT INITIATED AND LED BY E. CAROLINE WYLIE, ETS, FOR
THE CCSSO FORMATIVE ASSESSMENT FOR STUDENTS AND TEACHERS (FAST) SCASS

The purpose of this document is to share some examples of the Council of Chief State School Officers (CCSSO) definition of formative assessment in practice. The CCSSO definition of formative assessment developed and approved by the CCSSO Formative Assessment Advisory Group and Formative Assessment for Teachers and Students (FAST) SCASS is presented below:

Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes.

The following are five attributes based on current literature that render formative assessment most effective.

Learning Progressions
Learning progressions should clearly articulate the sub-goals of the ultimate learning goal.

Learning Goals and Criteria for Success
Learning goals and criteria for success should be clearly identified and communicated to students.

Descriptive Feedback
Students should be provided with evidence-based feedback that is linked to the intended instructional outcomes and criteria for success.

Self- and Peer-Assessment
Both self- and peer-assessment are important for providing students an opportunity to think meta-cognitively about their learning.

Collaboration
A classroom culture in which teachers and students are partners in learning should be established.

It is important that the reader first recognize formative assessment and what it is not before developing a more nuanced understanding of formative assessment. This is akin to learning to appreciate jazz. The first step is to be able to distinguish jazz from blue grass or funk. Recognizing the broad genre is an important prerequisite before moving on to learn how the various aspects of jazz music such as blue notes, call-and-response, improvisation, and syncopation all work together to create a musical performance.

Therefore, there are two sets of vignettes. The first set provides very brief examples and counter-examples of formative assessment. The second set illustrates extended examples of formative assessment practices and the interconnectedness of the various attributes. The vignettes are taken from teacher observations conducted in a variety of schools across the U.S. These vignettes illustrate formative assessment practice across a range of grade levels and content areas. Each vignette provides a description of the classroom activities, followed by a brief analysis that relates the actions of the teacher and students to one or more of the five attributes of effective formative assessment. Note that a particular formative assessment practice may not exemplify all five attributes.

1 Grateful thanks go to the various members of the FAST SCASS who contributed examples of formative practice and provided feedback on various iterations of this document.
VIGNETTE SET A:
IS IT OR ISN'T IT FORMATIVE ASSESSMENT?

Vignette 1: Thumbs Up and Thumbs Down

A high-school biology teacher frequently reads aloud a prepared biology-related statement, then asks students to hold up their thumbs under their chins and signify whether the statement is true or false by showing a “thumbs-up” for true or a “thumbs-down” for false. Depending on the number of students who respond incorrectly, the teacher may have students present arguments for both sides, he may pair students and ask them to discuss the concept further, or he may decide that he needs to present the same concept using a different representation or instructional approach.

This teacher is using a formative assessment approach to collect evidence to adjust instruction. This is, therefore, an instance of formative assessment.

Vignette 2: Structured Pair-Work

Each student is given an appointment clock and is required to make an appointment with three other students for discussion later in the lesson. Once all the appointments have been made the teacher begins the lesson, providing information and posing questions that require higher-order thinking about the information. The students are asked to reflect on the information and to answer specific questions. Then the students go to their first appointment and spend approximately 5 minutes sharing their thinking as it relates to one or two of the posed questions. They analyze each other’s responses and come to consensus. As the students work with their partners, the teacher walks around and notes common misunderstandings and gaps in understanding. At the conclusion of the first appointment, the teacher uses the information gained during the informal observations to help redirect thinking, to reinforce ideas, and to provide cues that will help advance their learning. The students then go to their next appointment and class continues in this manner until all appointments have been met and all questions have been discussed.

This is an example of formative assessment where the posed questions and the peer conversations are used to elicit evidence of the students’ understandings. In this context, the formative assessment process is embedded into the learning activity itself due to the teacher’s careful engineering of the activity. The students are able to self-reflect and get feedback from their peers. The teacher is able to listen to the conversations between students to note the current level of understanding for the class and for individual students. The teacher uses the information immediately to assist students in their learning by redirecting thinking, reinforcing ideas, or providing cues.

Vignette 3: Collective Definitions of Success Criteria

The teacher provides students with an open-ended question related to a concept they are studying and asks the students to identify the information or details necessary for a response to demonstrate full understanding of the concept. A list of these details is recorded on the board. The teacher then provides students with examples of several student responses that were given by students in previous years. The students are asked to analyze the responses and to determine if the responses show full understanding, partial understanding, or no understanding of the concept. Students must justify their answers. As this thinking is shared, the list of details or supports necessary for a response to the question is further refined until a set of criteria emerges that students can use to self-assess and peer-assess their responses to the question.

In this example of formative assessment the teacher is provided with information about student learning and the process used to gather that information also requires students to reflect on their own learning. This activity provides the teacher with information about how well the students understand the concept and how best to demonstrate that understanding. To fully participate in the activity, students must reflect on their own level of understanding as they analyze the work of others and provide reasons why they think there are gaps in understanding.
Vignette 4: District-Developed Assessments

District-developed monthly exams are to be administered to all students at the end of each of the school year's first eight months. The exams are based on state-authorized curricular goals for the grade and subjects involved. Because district administrators insist that teachers send results of these tests home to parents, all teachers do so. Yet, because the content covered by the monthly tests typically doesn't coincide with what is being taught at the time the tests are administered, teachers rarely alter their instruction based on students' performances on the monthly exams.

In this example, we see neither teachers' adjustment of their instruction nor students' adjustment of their learning tactics. Thus, this probably well-intentioned distribution of the monthly exams' results to parents would constitute a counter-example of formative assessment.

Vignette 5: Classroom Quizzes

During a unit on photosynthesis, the teacher administers a weekly quiz that addresses all of the material covered for the week. The quizzes are supposed to motivate students to study for the summative unit as well as provide students with a sample of the types of questions they may encounter on the unit test.

This is not an example of formative assessment because the teacher does not use the evidence from the quizzes to adjust instruction, nor does the teacher provide direction to students for them to think metacognitively about their own learning. The only information the students receive is a score for the number of correct answers. This is an example of ongoing summative assessment, not formative assessment.

VIGNETTE SET B:
EXAMPLES OF FORMATIVE ASSESSMENT IN PRACTICE

In this set the analysis relates each vignette back to the five attributes of effective formative assessment. The attributes are characteristics or features of effective formative assessment that the literature suggests are important. As the vignettes illustrate, some instantiations of formative assessment practice do not incorporate all of the attributes. For example, a particular vignette might not involve self- and peer-assessment, but it could still represent formative assessment practice. However, a teacher with a well-developed repertoire of formative assessment practices incorporates self- and peer-assessment, as appropriate.

These vignettes should not be viewed as complete descriptions of how particular teachers operationalize the concept of formative assessment in their classrooms, but rather illustrative of aspects of that practice. One way to consider the vignettes is to focus on the ways that one attribute appears across multiple vignettes. For example, the vignettes can inform the reader about the many ways in which feedback can be used, or provide insight into variations in teacher and student collaborations.

Vignette 1: Language Arts, Upper Elementary

An upper elementary language arts teacher began the lesson by asking a series of planned questions about a story that students had just finished reading. The teacher first reminded the students about their reading learning goals for this week that focused on identifying the main idea and supporting details within a story. Her questions required careful analysis by the students, so the teacher structured her approach by asking students first to think about their answers as individuals and, then discuss their answers in small groups. Each group was to reach consensus on a single answer and that group answer was then shared with the rest of the class using Whiteboards that designated students held up. With this questioning and group work approach, the teacher was able to identify several groups of students who were having difficulty understanding the concept. Summaries of the main idea of the story varied widely in accuracy and clarity. As the lesson was nearing the end, she asked the students to look at the various groups' answers about the main idea, to select the one that they thought was the best answer,
and to write down why they made the choice they did. She had students answer using an Exit Ticket—index cards on which students wrote their individual answers and then handed to her as they left the classroom. This approach provided her with a quick way to review student thinking at the individual level, thus providing information that she could use to shape the next day’s lesson.

Several attributes of effective formative assessment are illustrated in this vignette. As the teacher planned for this lesson to address the learning goal about the identification of the main idea, she developed a series of questions to ask her students. In addition, she planned a systematic way to allow students to think deeply about the questions, and to share their thinking with her and with members of the class. For this to be accomplished, the teacher had, over time, established a learning environment that emphasized collaboration: students were used to working in small groups and using whiteboards was part of the routine so that the teacher could use these evidence-gathering approaches with little explanation required. Finally, to support the impression developed during the lesson about group difficulties, the teacher was able to elicit evidence of student learning using Exit Tickets, an approach that required little time to review. With this student-level evidence of understanding she was able to tailor her lesson the following day to capitalize on those students who had a deeper understanding of the learning goal. By incorporating various formative assessment strategies during a single lesson, the teacher deepened her understanding of where her students were in their learning, and gathered evidence to inform her instructional decision making.

Vignette 2: Mathematics, Upper Elementary

A fifth grade mathematics teacher had been working with his students in the area of data analysis, and had recently introduced the class to the concept of using measures of central tendency to summarize data. He was aware of several of the typical misconceptions that students had about the concept of “median.” In particular, he knew that students often did not think that ordering the numbers in a data set was a necessary first step, and that students often did not understand how to handle data sets with an even number of elements. He wrote two multiple-choice questions to address these common misconceptions.

At the start of the lesson the teacher reviewed what they had covered in regards to measures of central tendency. He also wrote the learning goal on the board: “Today we will learn how to select appropriate measures of central tendency.”

Students had been using electronic clickers for the opening questions in mathematics class each morning. As a quick review of previous lessons, the teacher presented both multiple choice questions to the students. Almost all students answered the two questions correctly. He was about to begin to address the goal for that day’s lesson, when a student asked, “But there could be two answers, couldn’t there?” He asked the student to explain his reasoning to the class. The student explained that the problem could be solved in two ways—either select the middle number in the set, or put the numbers in order and then select the appropriate number. The teacher decided to poll the class and asked how many agreed with the student’s explanation of two possible answers. Just over half the class agreed that the problem could have two different answers.

The teacher, on-the-fly, wrote up two identical data sets on the board, each with five items, except one set was ordered and the other was not. He asked students to think on their own and then discuss with a partner to decide whether the two sets had the same median value. As students discussed this with a partner, the teacher circulated around the groups, making some notes of what he heard in the conversations. After about 10 minutes, he polled the class a second time, and now much fewer than half the students thought that the two sets had different medians. From the notes that the teacher took as he listened to students, he was able to identify several students who had very clear explanations for why the two sets had the same median value. He called on those students first to share their thinking with the class, and then asked students who disagreed to give their explanations. One student who had not previously been convinced by her partner shared her new understanding with the class. The teacher decided that the class was now ready to move on to the planned part of the lesson, but made a note to return to this problem for the class warm-up in a couple of days.

Several attributes of effective formative assessment are illustrated in this vignette. The teacher had a clear understanding of the learning progression that he wished to move students through on their way to learning the larger learning goal of applying and interpreting measures of central tendency. He was also aware of common misconceptions that students have or develop in this particular area, and so created two questions to get at these misconceptions. Although students answered his questions correctly, he was still sensitive to the additional evidence provided by the one student who asked the question, and supplemented that evidence by a quick class poll. Realizing that although students
were able to complete the procedural steps to find the median value, they clearly did not understand the concept yet. He adjusted his instruction by creating an additional question for them to discuss in pairs. He continued to collect evidence systematically, both through listening to students’ conversations and through another poll. After facilitating and guiding students’ conversations, the teacher decided that the class was ready to move on. However, he planned to return to this concept in a couple of days to check that the students had internalized the concept. This teacher used multiple sources of evidence in an effective way: planned questions at the start of the class; class polling to get a sense of the students’ thinking; and listening to student conversations.

Vignette 3: Mathematics, Middle School

A middle school mathematics teacher established a start-of-class routine with her students that she called “Homework Help Board.” Twice a week the teacher would assign problems for students to complete as homework. As students came into class the following day, they would review their homework and identify any problems with which they had difficulty. They would write that problem number on the board. A student who had been able to successfully complete the question would write the solution on the board. This process was followed until all identified problems had correct solutions. If another student had a different approach to solving the same problem or a different answer, that student would post the alternative solution.

If there were alternative correct solutions to any one problem, the teacher would point that out and ask students to discuss the different approaches. If something was incorrect, the teacher would ask the class if someone could correct the problem. If a problem had been noted on the board and no solution provided, the teacher would then review that problem, suggesting perhaps a first step or something to think about to see if a student, with support, could complete the solution. If all problems had a correct solution on the board, the teacher would either begin the lesson directly or perhaps ask one or two questions as a check to be certain all students understood the concepts.

Usually the teacher was able to begin the lesson quite quickly as students had been able to resolve all the homework questions themselves. When students struggled to do this, she modified her instruction as necessary for that day or the following day in order to review or revisit the topic.

This formative assessment example focuses primarily on the collaboration and self- and peer-assessment attributes of effective formative assessment. Students had to self-assess their homework efforts, and the classroom culture emphasized collaboration so that students were willing to take a risk and share their approach to a problem that another student had encountered. This approach was a form of systematic data collection; the teacher was able to identify across the entire class which homework problems had caused difficulties. The teacher was able to decide whether to react to this information immediately or whether a more detailed response in a subsequent lesson might be necessary. Either way, the information collected by the teacher would cause her to adjust instruction to improve student learning. The students’ solutions on the board also provided some teaching opportunities to help students think about multiple solution strategies, a particularly valuable approach when students use different representations to display their thinking.

Vignette 4: Language Arts, Lower Elementary

A lower elementary teacher had been working with his students on how to use the writing process to improve their own writing and to assist their peers. They were all working on writing personal stories around the topic of animals to publish in a class magazine. He introduced the idea of Two Stars and a Wish as a way of providing feedback which requires the reader to identify two positive aspects of the piece of work (the stars), and one specific suggestion for improvement (the wish). He began by explaining that the feedback he had given on their writing adhered to this very same principle. Now students would use the Two Stars and a Wish approach as they worked with each other in their writing groups. They discussed as an entire class the types of comments they might write when addressing positive aspects as well as comments to write when suggesting improvements. They also discussed the types of comments that would not be appropriate to write to a peer.

The teacher then reminded students about their current writing project. He returned the writings to each student, and asked them to reread their essays and to review his comments. He then directed them to discuss with their partner the feedback given and specific ways they might improve on their next draft. During the student discussions, the teacher circulated the room, offering suggestions to those who had difficulties planning their next steps.
During the next lesson, the students revised their work, and then began peer editing using the Two Stars and a Wish approach. He reminded them of the conversation they had about appropriate and inappropriate comments, of the type of comments he had written, and of the purpose of writing these stories to share in the class magazine. As students peer edited and provided feedback to their partners, the teacher again circulated the room and made suggestions, as necessary.

The sharing of the feedback process was repeated and students then had an opportunity to make final revisions before submitting their final work for inclusion in the magazine.

The example focuses on the following attributes of effective formative assessment: collaboration, the use of descriptive feedback and self and peer assessment. The teacher introduced a new structure for the students to give feedback to each other. The collaboration between and among the teacher and the student was evident in the way that he modeled the feedback approach and also gave students an opportunity to think about appropriate and inappropriate comments to write. The feedback provided by the teacher not only modeled the approach, but also was part of the learning process.

An important aspect of this instruction was that the teacher provided a structure for the students to review the feedback, and decide what they would do in the light of the feedback, using a peer as a sounding board, in order to stimulate action. Giving feedback without the time to react to it is of little value. When it was time for the students to analyze the writing of their peers, the teacher revisited the earlier discussion about the structure of Two Stars and a Wish and the types of appropriate feedback. The peer assessment was done in quite an informal way, without descriptions of performance levels, but the students had a clear purpose and audience for the writing and the Two Stars and a Wish approach provided structure.

Vignette 5: Science, Middle School

A middle school science teacher decided that her eighth grade students were not benefiting as much as they could from the science inquiry experiments around which she structured her units. Her goal was that each week students would complete a lab report and, as part of the report, connect what they learned from the experiment to the “big ideas” that she was presenting throughout the unit. However, she realized that students were struggling with the lab report content, and she was spending much of her grading time commenting on earlier sections of the report rather than focusing on the connections that they were making to the big ideas. She had been using a “criteria-for-evaluation form” that described her expectations for the reports, but decided that it needed to be revised so that it was more easily understood by the students. Because the form was to be used by the students as they prepared their lab reports, she decided to allow the students help her in designing the new instrument.

Before the start of the new school year, she pulled four student lab reports from the previous year, removing all identifying information. During the first class, she reviewed the criteria for evaluation form, and then handed out the four student reports. Working together in small groups, she asked the students to rank the reports, using the evaluation form, and to justify their ranking. A member of each group was randomly selected to explain his or her analysis of one of the reviewed reports. Other students were then allowed to add their comments. Once all four reports had been ranked and discussed by the class, she presented the rank order based on her grading of the reports and tied it back to the criteria-for-evaluation form. She responded to student questions about her scoring of the reports. She then reviewed again the purpose of the experiments and the write-ups, to help them better understand and make connections between their results and the big ideas of the course. She challenged the students to improve the criteria-for-evaluation form by creating their own check-lists that would help them do a better job on the lab reports.

Students then went back to their small groups, and reviewed again the two higher scoring lab reports, against the criteria-for-evaluation form, discussed how to describe the important aspects of the reports in their own words and from there developed their own list of important criteria. A second class period was spent compiling the ideas from each group, creating a final criteria list, and ensuring common understanding.

For each lab report that students completed during the year, students were reminded to compare their own work against the criteria list as a first check. Then the teacher allowed 15 minutes of class-time each week, for students to exchange draft reports with another student and to provide feedback. Students used the criteria list as they peer-evaluated the reports and followed a 3 = 2 feedback strategy (two positives and two suggestions for improvement). Students then made any final revisions that evening before submitting final draft.

The teacher discovered that students were better able to understand her expectations using the student-developed criteria list, and that the quality of the lab reports increased significantly, along with greater depth of student thinking. Furthermore she was able
to spend more time focused on the connections that students were making to the big ideas which in turn was helping her plan her instruction better having a clearer sense of what they were still struggling with and what they understood.

This formative assessment example focuses on several attributes including collaboration between and among teacher and students, and self- and peer- assessment. The teacher structured her instruction around a learning progression, using a sequence of inquiry experiments that were intended to help students develop understanding of several "big ideas". Recognizing a weakness in her process, she engaged her students collaboratively to develop lab report criteria list that they clearly understood and that would help them improve the quality of their reports. Students were then encouraged to engage both in self-assessment, to monitor their own work against these criteria, and then to engage in peer-assessment by using the criteria to review the lab report of another student. Time was built into the regular classroom schedule allowing students time to incorporate the descriptive feedback before handing in their reports. In addition, as the quality of the reports improved, the teacher was able to attend more to the content that was most critical to the overall learning. As a result she was able to use that information to adjust teaching in the light of student learning.

Vignette 6: Economics, High School

An important concept in a high school economics curriculum is that of supply and demand. The high school economics teacher knew that his students needed to have a firm grasp of the five determinants of demand and the six determinants of supply in order to understand how all the variables interacted to determine pricing.

After introducing the five determinants of demand, the teacher showed a movie clip from Hudsucker Proxy that shows a sudden change in price of a product due to a change in one of the five determinants. The teacher then asked students to identify which of the five determinants it was and to justify their response. From the student responses, the teacher was able to identify those who had understood the concept and those who had not. He had an already-prepared worksheet for those who did not yet grasp the concept and he worked with this small group of students, using the worksheet to identify and address areas of misconception. Students who had answered the question correctly were able to begin the homework assignment which reinforced the concept covered in class while the teacher worked with the other group.

The teacher repeated this process for the six determinants of supply, using a clip from Forest Gump, and again asking students to identify which determinant was involved. Again, he worked with the students who did not understand and allowed those who did to begin the reinforcement activity.

The key attributes of effective formative assessment in this vignette is that of the learning progression. The teacher had a clear understanding of the learning progression that she wished to move students through in addition to the subsidiary learning goals within the topic. He anticipated key junctures in the unit on supply and demand and embedded an assessment point into his instruction that allowed him to quickly determine students' understanding of the determinants. Importantly, he had a plan of action prepared ahead of time of what he would do to support student learning both for those individuals who did not yet understand the content, and for those who did. If students did not grasp this concept, moving on to interpreting supply and demand graphs would be pointless.

Vignette 7: History, High School

A history teacher created an end-of-year research project for students to work on over a period of several weeks. The goals of the project were to enable students to gain a broader perspective on domestic policy across 60 years of history and to develop their research skills. Each group was assigned a different decade between the 1950's and 2000. Students were given guidelines regarding the structure of their 15-minute presentation as well as specific criteria for the study notes that they had to create. Each student within the group had to be responsible for a piece of the presentation.

The first step was for students to identify important policy issues for their assigned decade and to develop a research plan for deepening their understanding. Each group met with the teacher for feedback on their list of identified issues and their plans for tackling the project. Since the teacher wanted students to learn from each others' presentations and notes, he met with each group on a regular basis to monitor progress, to give them an opportunity to reflect on their group's progress and next steps, and to provide feedback on their presentations and draft notes. A few days before the final presentations were to be made, the groups were paired so that they could listen to each other's dress rehearsal presentations and give each other feedback on the notes, the content of the presentation and the delivery using the 2+2 format to note two positive aspects and two aspects that could be improved. The teacher then met with each group to review the feedback they had
received and to discuss possible solutions to issues or concerns.

The teacher created an end-of-project multiple-choice assessment that was based on the content across all the presentations. In addition, the year-end essay required students to analyze themes and patterns in domestic policies across the decades.

This is an example of a complex project that required students to become experts in a particular decade so that they could teach their peers. This example illustrates how formative assessment is part of the balanced assessment approach. Because of the structure of the final summative assessment (multiple-choice quiz and essay), students had to use the presentations of their peers and the study notes to become familiar with the major events, influences, and policies of the other decades. Within the extended project, there are multiple aspects of effective formative assessment practice. The teacher monitored and guided student progress and gave students descriptive feedback along the way. He encouraged them to be self-reflective of their own progress and to use peer assessment strategies to give feedback to another group.

Vignette 8: Band/Orchestra, High School

The band teacher worked through a section of a musical piece with the class that was selected because it contained several challenging sections that would require students to apply some of the recent lessons. As the band played through a particular measure of music, he noticed a sour note. He stopped the group and asked whether anyone has noticed anything. Several students indicated that they noticed some off-notes towards the end of the section. He then asked everyone to look at the second to last measure in the section. He asked everyone to play that measure and hold each note until directed to play the next one. By doing this, he was able to help the students isolate the instruments that were playing the wrong pitch.

On occasion, this approach results in the students discovering for themselves that they are using an incorrect fingering. In this case, students were still unsure about what exactly needed to be corrected. However, a couple of students volunteered to the class that they heard the wrong pitch in the trumpet section in the second beat of the measure. The teacher then asked the trumpets to hold up their instruments so he could see the valves and then hold down the fingering for the note(s) in the second beat of the measure. He then had the students observe the fingerings they had used. It was obvious where incorrect fingerings were applied when students looked at each other's instruments and the correction was made immediately. A final check of understanding required the students to play the measure again by holding each note and checking for correct pitches.

This is an example of how small the grain-size of a formative assessment can be. There is only a fine line between the instruction and the assessment process so that the assessment is part of the instruction. In addition to working on playing the particular piece of music, the teacher is also developing students' musical ears so that they can identify problems, and then correct them. He encourages students to listen to each other and to observe each other in order to self-assess themselves.