Eye on Evaluation
DATA AND ACCOUNTABILITY DEPARTMENT

D&A Report No.13.05 **May 2013**

Formative Assessment with Technology 2011-12: Second Year of Implementation

Talbot Troy and Dina Bulgakov-Cooke

Summary

Background

In 2010-11, Wake County Public School System (WCPSS), North Carolina Department of Public Instruction (NCDPI), and SMART Technologies began a two-year collaboration to promote the use of formative assessment in mathematics classrooms at three Wake County public schools: Millbrook Elementary, East Millbrook Middle, and Millbrook High school. The Formative Assessment with Technology project was unique in its design: participating mathematics teachers at an elementary, middle, and a high school received face-to-face training on formative assessment and the use of SMART Boards, additional training on formative assessment components using online North Carolina's Formative Assessment Learning Community Online Network (NC FALCON) modules, and two years of follow-up ongoing support similar to coaching on a monthly basis. The NCDPI support included training, observations, follow-up feedback submitted to teachers, and monthly discussions of the implementation challenges. SMART Boards were made available to all project participants to give an opportunity for teachers and students to utilize the instructional technology to enhance formative assessment, learning, and student engagement. Project coordinators from NCDPI expected that 2011-12, the second year of implementation, would help improve student achievement if the formative assessment process was implemented with fidelity.

Abstract

The second year (2011-12) of the Formative Assessment with Technology Project in three WCPSS schools was studied to determine the fidelity of implementation of the formative assessment practices and the project's impact on student achievement. Classroom observations and surveys of teachers and students showed that implementation was moderate, and it varied widely across teachers. Certain targeted practices were reported or observed much more often than others. High turnover in the teachers participating in the project likely contributed to the uneven implementation. High school staff had the highest level of implementation of training concepts as well as the highest use of the electronic response tools for assessment. Full implementation and impact on state achievement test scores was expected to take two years to be evident. Since only nine teachers participated and had data for two years, impact could not be reliably assessed. This pilot training model would be too expensive to roll out district-wide in WCPSS. More cost effective models may be used to encourage appropriate use of formative assessment.

Table of Contents

Formative assessment was defined in this project as "a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes." The definition was adopted from the Council of Chief State School Officers (2008) which described formative assessment as a process that teachers use while providing instruction to gain feedback on students' understanding of the content.

Implementation

Data for this report were collected through participation records, teacher and student surveys, and 20-minute monthly classroom observations by NCDPI staff. Overall fidelity of implementation of the project in 2011-12 was moderate. As a group, high school teachers showed consistently higher levels of implementation, followed by the elementary school, and finally the middle school. The electronic response systems were most frequently used in high school, while elementary school students most frequently had an opportunity to talk about mathematics.

Almost all teachers who responded to the teacher survey (17 of 18), considered all components of formative assessment to be "very" or "moderately valuable". However, a few key concerns exist.

Participation: As shown in Table 1, only nine of the 25 teachers (36%) participating in 2011-12 had also participated in 2010-11. Seven of the 16 first-year participants left the program and 16 new teachers joined the program during the second year of implementation.

Table 1

Project Participants by Year

	2010-11 Participants	2011-12 Participants	Both Years Participants
Elementary	5	5	3
Middle	6	12	2
High	5	8	4
Total	**16**	**25**	**9**

Data Source: NCDPI.

Training: While the expectation that all teachers would go through at least four of the five FALCON modules, only about 83% of the participants stated that they did. Additionally, the face-to-face introductory trainings conducted by NCDPI were considered at least "slightly effective" by all respondents. The monthly post-observation meetings were rated as only "slightly" or "not effective" by seven of the 15 respondents. It is noteworthy that fewer than 40% of teacher survey respondents reported that they felt fully proficient with any of the components of formative assessment.

Application in Classrooms: Inconsistent levels of implementation were a third issue (by teacher and by component). While most teachers reported that all of the components of the formative assessment training were valuable, few responded that they felt proficient in applying all of the

components, especially in the areas of using descriptive feedback and facilitating student self-assessment.

Twenty-minute monthly classroom observations focused on formative assessment components that were expected to be seen frequently. Learning targets and criteria for success, in particular, were expected to be posted, easily visible (or provided to the observer) nearly all of the time. The other items might not be observed in every 20-minute observation, but were expected to be seen often. A wide range of implementation of formative assessment components was evident, with all items seen less often than desired for full implementation.

- Half of observed classrooms provided evidence of posted learning targets and criteria for success.
- About one-third of classrooms had learning targets and criteria for success that were aligned and that were written in a student-friendly language.
- Teachers in one-fourth of classrooms were observed communicating about learning targets and criteria for success and providing descriptive feedback on student work.
- Few instances of peer review were observed.
- Students were almost never observed using the criteria for success to explain where they were in their learning.

Student surveys reported greater use of some components than observations suggested. Most students reported teachers "always" or "frequently" talked to them to see if they understood the mathematics and provided helpful feedback. Over half of the students at all three schools reported that feedback they received from their teachers "always or frequently" helped them improve their mathematics skills and gave them confidence to do mathematics. The vast majority of high school students (79%) and elementary school students (84%) felt that teachers made sure they understood the feedback, with middle school students being less positive (62%). Thus, students saw teachers' feedback as more frequent and more helpful than teachers did. It could be that students received more frequent traditional feedback, and teachers were holding themselves to the higher standards offered in the training.

Differences between the observation and survey results likely relate to two factors:

1) Teachers were not observed for full duration of the lessons and therefore use of some components could have been missed; and
2) Observers may have been more stringent in their ratings based on expected behaviors covered in the training than teachers and students were in their responses.

Teachers and students were generally positive about the usefulness of SMART Boards in mathematics instruction. A large majority of students (83%) reported that SMART Boards "always" or "frequently" made mathematics class more interesting. All teachers "agreed" or "strongly agreed" that students were more engaged in learning with SMART Boards than without, and that the SMART Board allowed to quickly give a real-time picture of student learning. While over three-fourths of students reporting that teachers "always or frequently" invited them to use the SMART Boards during class, observations recorded student use of SMART Boards in less than 50% of elementary and middle school classrooms.

Impact: Schools did not show consistent patterns of significant improvement in student outcomes compared to demographically similar schools. Given the assumption that two years of implementation were necessary before student outcomes would be evident, the value of two years of training could not be assessed. Only nine teachers participated both years and had suitable data. On the other hand, the cost-benefit of this expensive training model, as implemented, was not validated.

Discussion and Recommendations

District leadership should provide consistent and focused support in accessing the formative assessment training with NC FALCON as one of the sources of such training. Some confusion still exists on the difference between benchmark and formative assessment. The online NC FALCON formative assessment modules used in this initiative rely on well-established research and explain the characteristics that distinguish formative assessment from other types of assessment. On the other hand, the DPI model of in-person training and coaching would be too expensive to replicate in a large school district. Results also suggest that a one-year model of support would be more practical than a two-year model (given the turnover encountered), and that school-wide training would be more efficient than selecting a subset of teachers.

- School-level administrators should promote PLTs as venues in which teachers support each other in continually improving their formative assessment practices.
- The potential for using new technologies, including SMART Boards and related instructional software, should be developed by central and school staff for learning and assessment. However, formative assessment can be carried out in multiple ways. Formative assessment is as old as the art of teaching and does not depend on new devices.

The full recommendations are offered at the end of the report.

Introduction

The 25 mathematics teachers from three WCPSS schools (Millbrook Elementary, East Millbrook Middle, and Millbrook High school) participated in the Formative Assessment with Technology project. They received face-to-face professional development on the use of formative assessment from the NCDPI team and training online using a set of training modules developed by NCDPI called North Carolina Formative Assessment Learning Community Online Network (NC FALCON). The goal of the use of formative assessment is to tailor daily instruction to student learning and thus ultimately enhance student learning.

As a condition of their participation, teachers were provided with SMART Boards and training on the use of SMART software specifically designed to facilitate formative assessment activities. It was hoped that, with this ongoing training and additional technology, the teachers would improve their use of formative assessment. During the first year of the project, evaluators surveyed students and teachers and observed classrooms to document the teachers' understanding and implementation of formative assessment. Details of the program and findings about the teachers' implementation of formative assessment are presented in *Status of Formative Assessment with Technology Project, 2010-11* (Bulgakov-Cooke, 2011).

The current report continues the study of teachers' understanding and implementation of formative assessment in the classrooms conducted in 2010-11. Furthermore, it explores the changes in student achievement outcomes that span the two years of implementation, and compares the participating schools' achievement to those of demographically similar schools that did not participate in the program.

Ideally, for a comparison of implementation and performance over a period of years, the participating teachers would be the same from the start of the project to the end. In this case, however, there were substantial changes in personnel, as seven of the 16 first-year participants left the program and 16 new teachers joined the program during the second year. Nonetheless, certain observations were made about how the program may have impacted formative assessment practice and student achievement.

The 2010-11 report was based on a student survey, a teacher survey, a teacher focus group, and direct observations in the classroom. The report indicated that during the first year of the program, participating teachers implemented at least some components of formative assessment that were emphasized in the training they had received. For example, they consistently posted learning targets and criteria for success, and students reported this to be helpful in their learning. Teachers were cognizant of the value of frequently gathering evidence of student learning in a variety of ways and adjusting instruction accordingly, but they expressed and exhibited that following through on these practices was sometimes challenging. Likewise, teachers understood the need to provide descriptive feedback to help students make decisions about advancing their learning; to the extent that teachers acted on this, a survey indicated that students considered it beneficial.

This report showed similar findings about the participants' dispositions towards an implementation of formative assessment. The report also presented student outcomes data for project participants including growth and proficiency percentages. Data were presented by school in comparison to demographically similar schools and for three years (before and after participation) for all teachers who had these data available.

Implementation

In studying the implementation of the project, the evaluators sought to examine the fidelity with which the expectations for professional development were met and the degree to which the participating teachers understood and applied formative assessment. The study documented participation as one of the implementation issues. While project coordinators from NCDPI hoped that the mathematics teachers would continue in the initiative for two years (Millbrook Elementary school kept participation at the same grade level), about 36% of the staff continued in the project in 2011-12. Because of the teacher turnover, about 64% of teachers were new and had to be introduced to the major concepts and strategies of formative assessment again in 2011-12.

Similar to 2010-11, data collection was conducted through a teacher survey, a student survey, and classroom observations. The survey and observation instruments were only slightly modified from those used in 2010-11. Unlike spring 2010-11, the observations by the NCDPI coordinator rather than the WCPSS evaluator were examined. Observations were conducted monthly, and lasted approximately 20 minutes per classroom.

Teacher Survey

During the first year (2010-11) of implementation, 10 out of the 16 participating teachers (63%) responded to a survey about how participation in the program influenced their use of formative assessment. The results of that survey are discussed in the 2011 evaluation report (Bulgakov-Cooke, 2011). A similar survey was given at the end of the 2011-12 school year with 18 of the 26 participating teachers (69%) responding. Table 2 shows the number of respondents by school**.**

Table 2

2011-12 Participants who Responded to the Survey

School	Number of Participants	Number of Respondents
Millbrook Elementary	5	4
East Millbrook Middle	12	6
Millbrook High	8	8
Totals	25	18

The 2011-12 survey included questions about the participants' teaching experience. As shown in Table 3, teachers most commonly had 6-10 years of mathematics teaching experience (8 of 18). Four of 18 respondents had taught mathematics for over 10 years, and four were still in their first two years of teaching mathematics.

Table 3

Number of Years Teaching Mathematics

Number of Respondents	Years Teaching Math
4	1-2
2	3-5
8	6-10
4	11 or more

As a condition of joining the program, all participants were required to attend face-to-face trainings provided by NCDPI and SMART Technologies. These trainings explained techniques of formative assessment and taught participants to use the SMART Boards and the accompanying instructional software. The survey sought to establish what other formative assessment training, including completed FALCON modules, participants had received during the previous three years. The results are summarized in Table 4. Two of the respondents reported having completed no additional training during that time span. All but three of the 18 participating teachers reported having received NC FALCON training.

Table 4
Formative Assessment Training (n=18)

Professional Development	Respondents
RttT Day 5	8
Formative assessment training created by your school's staff	6
Modules I through IV of NC FALCON	12
Modules I through III of NC FALCON only	3
No training other than face-to-face training provided by DPI and SMART Technology	2

In addition to training sessions, teachers were given the opportunities to explore formative assessment through Professional Learning Teams (PLTs) and through the post-observation meetings with NCDPI and SMART Technology staff. Participation in these activities was high to moderate: All but two respondents attended the post observation meetings fairly regularly, and 10 participated regularly in PLTs where formative assessment was a topic of focus (see Table 5). Another activity available to teachers was NC FALCON's online Professional Learning Community (PLC), but only three respondents indicated any substantial participation.

Respondents were asked to describe the effectiveness of the online NC FALCON modules; and only a few were able to offer comments. One respondent explained that the modules were effective but could not elaborate because "it was a while ago when I did them." Another respondent explained that NC FALCON's effectiveness was limited because the modules "were not geared towards specific grade levels or content areas." A third respondent considered the modules effective but that "doing them alone without having a group to discuss with makes it difficult to process." Due to the low participation rate in the online PLC component of NC FALCON, participants' perceptions about its effectiveness were not reported. The participants' views of the effectiveness of the face-to-face trainings, post-observation meetings, and PLTs are summarized in Table 5.

Table 5
Teacher Participation in Other Professional Development (n=18)

All five post-observation meetings with DPI and SMART Technology staff members (during 2011-12)	8
Three or four of the post-observation meetings with DPI and SMART Technology staff members (during 2011-12)	8
Three or more PLT's per month in which formative assessment was a focus (during 2011-12)	10
NC FALCON online PLC	3

The required face-to-face trainings were viewed as either "very" or "moderately effective" by 12 of the 18 respondents (see Table 6). The school-based PLTs also had favorable ratings, with 12 of 14 respondents rated those as "very" or "moderately effective."

The post observation meetings were intended to allow the teachers and observers to share with each other about what had been observed and how instruction has been influenced since the previous month's meeting. In addition, the observers followed up by emailing more detailed feedback to each teacher. However, only two of 15 respondents rated the meetings as "very effective," and four rated them as "not effective." The comments provided by participants seemed to indicate that these lower ratings were due largely to the fact that the emailed feedback from the previous month's observations had not been made available quickly enough to allow teachers time to reflect and apply it to the currently observed lessons.

Table 6
Training Effectiveness Ratings

	Very Effective	Moderately Effective	Slightly Effective	Not Effective
Face-to-face trainings provided by DPI and SMART Technologies	7	5	6	0
Post-observation meetings	2	6	3	4
PLTs	5	7	2	0

Table 7 shows participants' ratings of six different components of formative assessment used to improve student achievement on a scale from "very valuable" to "not valuable." For every component, "very valuable" was the most common response. Overwhelmingly, 17 of 18 (94%) participants considered all components of formative assessment to have at least moderate value. The most highly valued component was adjusting instruction in response to evidence about learning, which 17 (94%) of the responding teachers rated as "very valuable." Least often participants rated collecting evidence of learning as "very valuable", with 11 (61%) responses.

Table 7
Six Components of Formative Assessment

Component of Formative Assessment	Very Valuable	Moderately Valuable	Not Valuable
Helping students use clear targets*	13	4	1
Helping students use criteria for success*	13	4	1
Collecting evidence of learning	11	7	0
Using descriptive feedback between teacher and student	14	4	0
Facilitating student self-assessment	12	6	0
Adjusting instruction in response to evidence about student learning	17	1	0

Participating teachers were further asked about their proficiency in applying the six components of formative assessment. The component in which respondents most often assessed themselves as having "high" proficiency (seven respondents) was adjusting instruction in response to evidence about learning (see Table 8). Using descriptive feedback was rated as "very valuable" by 14 respondents (Table 7). However, seven of 18 respondents indicated "substantial room for improvement" in their use of that same component (Table 8). The component in which the most respondents (nine) felt their proficiency had "substantial room for improvement" was facilitating student self-assessment. It is noteworthy that fewer than 40% of respondents felt they were fully proficient with any of the components.

Table 8
Respondents' Self-Rating of Proficiency (n=18)

Components of Formative Assessment	High: Little or no room for improvement	Moderate: Some room for improvement	Low: Substantial room for improvement
Helping students use clear targets	4	12	2
Helping students use criteria for success	0	16	2
Collecting evidence of learning	3	12	3
Using descriptive feedback between teacher and student	4	7	7
Facilitating student self-assessment	1	8	9
Adjusting instruction in response to evidence about student learning	7	10	1

According to the survey, teachers used various techniques for collecting evidence of learning. For example, as shown in Table 9, all teachers reported using oral questioning multiple times per lesson. However, it is worth noting that the survey did not specify the type or quality of questioning. A frequently used technique of collecting evidence of learning from a group of students was the use of hand-held whiteboards, thumbs-up, or similar low-tech strategies used once or multiple times per lesson by 15 participants (83%). All four of the elementary teachers reported using this technique multiple times in every lesson, while none of them reported using the electronic response clickers that are part of SMART Board technology any more than once every few lessons. Response clickers were used about as often as short written quizzes.

Table 9
Collecting Evidence of Learning
(n=18 unless noted otherwise)

Techniques Used	Multiple Times per Lesson	Once per Lesson	Once Every Few Lessons
Oral questioning	18	0	0
Hand-held whiteboards, thumbs-up, or similar low-tech strategy*	11	4	3
Electronic response system** (n=16)	5	0	11
Think-pair-share (or similar reporting out activity) (n=16)	5	5	6
Short written quiz (non-graded) (n=14)	1	3	10

*All four elementary teachers report using this technique multiple times per lesson.
** None of the elementary teachers reported using this technique.

SMART Board technology was made available to all project participants as a tool for assessment and learning. Because at times teachers tend to limit the use of SMART Boards to a traditional whiteboard due to a lack of training, participants were asked about their SMART Board training. Almost all participants (16) indicated that they felt "well trained" or "somewhat well trained" to use a SMART Board. It is worth noting that 14 participants still indicated that they used it in the same way they would use a white board or projector; however, none used it only in that fashion. As Table 10 shows, the two most popular ways to use a SMART board, beyond its basic white board and projector capabilities, were in conjunction with the response clickers (16 participants) and accessing the lesson formatting capabilities from the SMART Exchange (15 participants). Lessons from the Gallery (software) and the lesson template provided by the project team were used by half of the respondents (9 of 18). Several respondents mentioned other ways they used their SMART Board, including creating lessons through SMART Notebook and using one's own templates. All of these resources and techniques were covered in the face-to-face trainings and the monthly meetings.

Table 10
Number of Teachers Responding about Training and Uses of SMART Board (n=18)

Well or somewhat well trained in its uses	16
Regular white board and projector capabilities	14
Electronic response system (clickers)	16
Lessons from SMART Exchange	15
Lessons from the Gallery	9
Lesson templates from project team	9

The respondents were generally positive about the usefulness of a SMART Board in mathematics instruction. All "agreed" or "strongly agreed" that students were more engaged with SMART Boards than without, and that the SMART Board quickly gave a real-time picture of student learning. It was also unanimously agreed that "the SMART Board offers many options for collecting evidence of learning" and that training students to use it is easy. At the same time, a few teachers felt the SMART Board was not an easy device to use for collecting and storing electronic response results (3 teachers) or for storing and re-accessing demonstration work (2 teachers).

Student Survey

A 20-question survey was given to students to measure their experiences with participating teachers' uses of formative assessment and the SMART Board technology. Two of the participating classes at the elementary school and three at each of the middle and high schools were randomly selected for the survey (approximately 200 students); 162 students (81%) responded, as shown in Table 11.

Table 11
Students who Responded by School (n=162)

School	Number of Respondents
Millbrook Elementary	37
East Millbrook Middle	73
Millbrook High	52
Total	**162**

The questions on the student survey covered five areas of interest to the sponsors and participants of the program:

- the use of learning targets and criteria for success;
- collection of evidence of student learning;
- feedback;
- self- and peer-assessment; and
- the use of the SMART Boards.

As reported in Table 12, 91% of the responding students at all three schools reported that their teachers "always" or "frequently" made certain that students knew the learning targets and criteria for success during mathematics instruction. However, being aware of the targets and criteria for success did not always translate into the students setting their own goals. About half (53%) of the respondents reported doing so "frequently" or "always." Middle school students set goals at the lowest rate (38%) compared to elementary or high school students. The ability to tell good math work from bad math work was at about 71% overall.

Table 12
The Use of Learning Targets and Criteria for Success

Students Responding "Always" or "Frequently"				
	Elem (n=37)	**Middle** (n=73)	**High** (n=52)	**All** (n=162)
Teacher made certain students knew the learning targets and criteria for success during instruction.	97.3%	83.6%	98.1%	91.4%
Students used learning targets and criteria for success to set their own goals.	64.9%	38.0%	65.4%	53.1%
In math class, students knew how to tell good math work from bad math work.	75.7%	67.1%	75.0%	71.6%

The student survey included questions about two of the ways teachers collected evidence of student learning: the use of electronic clickers for selected response items and speaking directly with students. As Table 13 shows:

- In the elementary school, only one-fourth of the students (24%) described the use of the clickers as "always or frequently," and more than a fifth (22%) reported that this method occurred "rarely or never" in their classrooms.
- The middle school students indicated substantially more frequent use, with approximately half (51%) indicating use "always or frequently". Even so, more than a fourth of them (28%) reported that clickers were used "rarely or never."
- The high school students reported the most frequent use of the clickers, with 77% reporting use "always or frequently."

At all three schools, three-fourths of the students felt comfortable using the clickers.

Table 13
The Use of Clickers—Student Responses

	Always or Frequently	Sometimes	Rarely or Never
The teacher had students use response clickers to find out if they understand the math.			
Elementary	24.3%	54.1%	21.6%
Middle	51.4%	20.8%	27.8%
High	76.9%	23.1%	0.0%
Students felt comfortable using response clickers to answer questions about math.			
Elementary	73.6%	13.5%	10.8%
Middle	73.6%	11.1%	15.3%
High	76.5%	19.6%	3.0%

Another aspect of a successful formative assessment cycle is teachers checking student understanding of the work (see Table 14). Students at the elementary and high school levels had more positive responses to these items than the middle school students. (The middle school also had the most new participants during 2011-12.)

- Over three-fourths of the respondents at the elementary and high school indicated teachers checked for understanding, knew when students were confused, and made it comfortable to raise their hands to ask questions "always or frequently." In contrast, 52.8 to 58.9% of the middle school responded that this was the case.

- Eighty percent or more of the students said teachers "always or frequently" checked on why students were confused.

Table 14
Student Responses: Teacher Gave me Feedback on Math Work

	Always or Frequently	Sometimes	Rarely or Never
Teachers talked to students to check for understanding of math.			
Elementary (n=37)	83.8%	13.5%	2.7%
Middle (n=73)	52.8%	25.0%	22.2%
High (n=52)	78.8%	17.3%	3.8%
Teachers know when students were confused about the math.			
Elementary (n=37)	86.5%	13.5%	0.0%
Middle (n=73)	58.9%	30.1%	11.0%
High (n=52)	74.5%	13.7%	11.8%
When students were confused, teachers tried to find out why.			
Elementary (n=37)	89.2%	5.4%	5.4%
Middle (n=73)	80.8%	9.6%	9.6%
High (n=52)	84.6%	15.4%	0.0%
Students were comfortable raising hands to answer questions about math.			
Elementary (n=37)	75.0%	19.4%	5.6%
Middle (n=73)	57.7%	21.1%	21.1%
High (n=52)	78.8%	9.6%	11.5%

One of the elements of formative assessment as presented in the NC FALCON modules is an expectation that teachers provide an extended, descriptive feedback to students when incomplete understanding occurs. Thus, students were asked about the feedback teachers gave them in mathematics class. Eighty-nine percent of elementary students reported that, when they do mathematics work or ask questions, their teachers always or frequently give them helpful feedback. See Table 15. They were also very positive on the question of whether teachers made sure students understood the feedback when it was given, with only two of the 37 respondents reporting they felt this happened "rarely or never," and 84% of elementary students reporting that it occurred "always or frequently."

Table 15 also shows that over 60% of all students "always or frequently" experienced beneficial practices related to descriptive feedback. At the elementary and high school levels, the responses indicated that teachers "always or frequently" provided helpful feedback (79% and 89%, respectively) and that they made sure students understood the feedback. There was less strong indication that the feedback helped students improve their mathematics skills or gave them more confidence to do mathematics; nonetheless, these two benefits were reported as occurring "always or frequently" by 62.2% to 74.5% of the students at all three levels.

Table 15

Student Responses: The Use of Descriptive Feedback

	Always or Frequently	Sometimes	Rarely or Never
Students felt they got helpful feedback from their teachers.			
Elementary (n=37)	88.9%	11.1%	0.0%
Middle (n=73)	61.1%	29.2%	9.7%
High (n=52)	78.8%	15.4%	5.8%
Students felt that teacher made sure they understood the feedback.			
Elementary (n=37)	83.8%	10.8%	5.4%
Middle (n=73)	61.6%	21.9%	16.4%
High (n=52)	78.8%	15.4%	5.8%
Students felt that teacher feedback helped them improve their math skills.			
Elementary (n=37)	72.2%	22.2%	5.6%
Middle (n=73)	63.4%	28.2%	8.5%
High (n=52)	72.5%	19.6%	7.8%
Students felt that teacher feedback gave them more confidence to do math.			
Elementary (n=37)	62.2%	32.4%	5.4%
Middle (n=73)	63.4%	25.4%	11.3%
High (n=52)	74.5%	21.5%	3.9%

It should be noted that the definition of "helpful feedback" in the survey is open to interpretation. Helpful feedback should be descriptive in terms of learning targets and should cause a student to think more deeply about a concept, a next step, or a solution. However, it is not completely clear from the survey if a student might consider feedback to be helpful if it includes too much guiding information that undercuts the challenge, or if it is simply evaluative feedback such as grades, checkmarks, or brief comments celebrating success and effort.

Although not presented in a table, student responses noticeably indicated low occurrence of two key behaviors associated with self- and peer-assessment. Less than a third of the middle school students (31%) said they were "frequently or always" given time to talk to each other about the

mathematics they were learning. At the high school and elementary school levels responses were significantly more positive, with 66% and 88% respectively selecting "always or frequently." Responses indicated even lower occurrence of students providing comments on each other's work, with fewer than half (46%) of all students reporting that this indicator of peer-assessment occurred "frequently or always."

Questions about the use of the SMART Boards generated very positive responses among students. The vast majority (almost 87%) of all the students reported that SMART Boards always or frequently made mathematics class more interesting, and over three-fourths (79%) reported that teachers frequently or always invited students to use the SMART Boards during class.

Classroom Observations

Ninety-one classroom observations occurred during five visitations to each of the three schools from October 2011 through April 2012 as shown below.

Table 16
Classroom Observations

Millbrook Elementary School	17
East Millbrook Middle School	46
Millbrook High School	28
Total	**91**

Each visitation was arranged to occur during mathematics instruction and lasted approximately 20 minutes. During each visitation, a DPI observer used a short checklist to tabulate teacher and student behaviors associated with implementation of formative assessment; for example, posting learning targets and criteria for success or writing those in student-friendly language. Table 17 and the observation checklist (see Appendix) list these behaviors along with the percentage of visitations in which each behavior was observed at least once. It is important to note that some behaviors were not easily detected without more thorough observations, or when the observations started in the middle or closer to the end of the class period. Because the visitations were relatively short, some behaviors related to formative assessment that may have occurred in class before the visit were not observed or reflected in the summary. Therefore, Table 17 should not be interpreted to indicate all of the mathematics lessons in which teachers and students engage in formative assessment or its associated behaviors.

Nonetheless, some teacher or student behaviors related to implementation of formative assessment were absent at times when they may have been expected. For example, only 4% of students were observed referring to the learning targets during visitations and only 1% referenced criteria for success while giving their explanations (see Table 17). A higher rate of reference to learning targets was observed in high schools (11%) than in middle or elementary schools. Student use of technology (SMART Board) was observed in less than 50% of the

elementary and middle school visitations. Use of feedback to advance learning was rated low, which may also be related to how frequently students actually received such feedback. On the positive note, nearly 41% of students were observed explaining their processes and/or methods of work, with the highest rate of observation in 59% of elementary classes.

Table 17
Observations of Student Behavior *(n=91)*

Student Behaviors	All Schools	Elem (n=17)	Middle (n=46)	High (n=28)
Use technology to demonstrate learning	47.3%	29.4%	45.7%	61%
Explain process/method used in work	40.7%	58.8%	30.4%	46%
Use feedback to advance learning	13.2%	5.9%	13.0%	18%
Provide feedback to others (peer review) to advance learning	12.1%	5.9%	6.5%	25%
Reference Target in explanation	4.4%	5.9%	0.0%	11%
Use criteria to explain "where he/she is" in learning	1.1%	0.0%	2.0%	0%

The expectation of the project designers was that some teacher behaviors related to formative assessment implementation should be observed in all classrooms. Such behaviors include:

- providing targets and criteria for success that are aligned,
- providing targets and criteria for success that are written in student-friendly language,
- providing instruction and formative assessment strategies that are aligned with criteria for success, and
- providing feedback on student work.

The actual observations showed a wide range of implementation (see Table 18). The targeted teacher behaviors that were most frequently observed were the use of a variety of strategies to collect evidence of learning: the use of student whiteboards, thumbs-up/thumbs-down, and exit tickets (56%). Collecting evidence of learning most frequently occurred in the elementary classrooms (almost 71%). The high school had a lower rate of observations of such strategies (64%), with middle schools observations of collecting evidence of learning at the lowest rate (near 46%).

The posting of learning targets and criteria for success by mathematics teachers was observed the second most often, but only 44% of the time—despite its being an expectation at or near the start of each mathematics lesson (see Table 18). The use of student-friendly language for the learning targets and criteria for success was also relatively low, occurring in only 34% of the observations**.** Thus, teachers appear to continue having problems stating the learning targets and criteria for success so that students can internalize them.

At the elementary and middle schools, the following key teacher behaviors were logged in 20% to 40% of the observations:

- Gathering evidence of learning as students use technology;
- Analyzing evidence of student learning;
- Providing targets and criteria for success that are aligned;
- Providing targets and criteria for success that are written in student-friendly language;
- Providing instruction and formative assessment strategies that are aligned with criteria for success.

During visits to the high school, the behaviors listed above were observed at a substantially greater rate (during 39% to 75% of observations). Additionally, more frequent occurrences of the teacher behaviors observed at the elementary and middle schools at lower rates were recorded in the high school:

- Providing feedback on student work based on criteria for success;
- Communicating about targets and criteria for success;
- Adjusting instruction based of evidence collected.

Overall, implementation of the formative assessment was below 50% on all except one behavior based on the observations alone. Less than half of classrooms provided evidence of learning targets and criteria for success being posted; and less than half of teachers gathered evidence of learning as students used technology. Only one-third provided learning targets and criteria for success that were aligned and were written in student-friendly languages. Providing feedback on student work was observed in only one-fourth of classrooms. Teachers communicated about learning targets and criteria for success in one-fourth of classrooms.

Student and teacher survey responses suggest some instructional and assessment strategies occurred which were not captured in these short observations. However, these relatively low implementation levels still make it difficult to expect higher student outcomes in the classrooms where formative assessment project participants taught mathematics.

Table 18
Observations of Teacher and Student Behaviors *(n=91)*

Teacher Behaviors	All Schools	Elem (n=17)	Middle (n=46)	High (n=28)
Collects evidence of learning using a variety of strategies (thumbs up, observation, questioning, math talk, exit tickets, etc.)	56.0%	70.6%	45.7%	64%
Posts Learning Target and Criteria for Success	44.0%	41.2%	47.8%	39%
Gathers evidence of learning as students use technology	39.6%	29.4%	37.0%	50%
Analyzes evidence of student learning	38.5%	29.4%	19.6%	75%
Provides Targets and Criteria for Success that are aligned	34.1%	29.4%	32.6%	39%
Provides Targets and Criteria for Success that are written in student-friendly language	34.1%	29.4%	30.4%	43%
Provides instruction and formative assessment strategies that are aligned with Criteria for Success	28.6%	35.3%	19.6%	39%
Provides feedback on student work based on Criteria for Success	24.2%	11.8%	17.4%	43%
Communicates about/shares Targets and Criteria for Success	23.1%	29.4%	15.2%	32%
Adjusts instruction based on evidence collected	23%	6%	13%	50%
Documents evidence of student learning (other than mental notes)	14.3%	17.6%	4.3%	29%
Checks for impact of feedback on student learning by providing other learning experiences	7.7%	0.0%	0.0%	25%
Identifies new Target and Criteria for Success	1.1%	0.0%	0.0%	4%

Student Outcomes

To examine student outcomes for participants of the Formative Assessment with Technology project, the evaluator examined the 2011-12 academic change scores for participating schools and demographically similar schools, looked for longitudinal change, (Tables 19 and 20) and at student outcomes for each participating teacher (Table 21). The North Carolina State ABCs academic change score and percentages of students who made growth were examined first. The academic change score is a measure based on state ABCs formulas used to show student growth.

The academic change score for a school or a teacher showed whether his/her students' growth as a group was more or less than the target projection for that year. A growth score of zero generally means the growth target was met exactly, with students showing one year of growth. Any positive growth shows outcomes better than expected, with a maximum value of 1.0; any negative score, with a minimum value of -1.0, shows that the teacher's students performed less well on average than predicted. Another ABCs indicator used – percent of students who made growth – showed the percentage of students taught by each project participant who met or exceeded predicted growth targets based on state ABCs formulas.

First, comparisons were made between student outcomes – examining proficiency, average academic change scores, and percentages of students who made growth –for the schools participating in the project and comparison schools (see Table 19). Comparison schools were the schools with comparable previous year ABCs performance composites and similar demographic characteristics. At the elementary school level only grade 5 teachers participated in the project. Thus, results for grade 5 at two comparison schools were presented.

The academic change score and the percentages of students who made growth were higher for Millbrook Elementary School than for comparison schools; however, the difference was not significant when all participating teachers (including four new teachers) were included into the analysis. A z-statistic was used for the analysis with the significance level at $p = .05$. Table 19 shows that the percentage of students who made growth was higher for all three elementary schools, including comparison schools, than for the middle and high schools. The fact that many elementary schools had a WCPSS mathematics coach in 2011-12 may have contributed to higher growth at elementary schools.

At the participating middle school, the outcomes for all grades in mathematics were higher, but not significantly higher, than for one of the two comparison schools, and lower than the other comparison school. The same was true for Algebra I. This pattern was accompanied by a significant rate of change in project participation membership among mathematics teachers, where only two of the teachers were the two-year participants. The high school did not show better outcomes than the comparison schools.

Table 19
2011-12 School Level Comparisons of Academic Change in Mathematics and Algebra I for Formative Assessment with Technology Project Participants and Comparison Schools

School	Percent Proficient	Academic Change	Percent of Students Who Made Growth
5 Grade Mathematics Elementary			
Millbrook Elementary	79.5	.37	79.4
Barwell Road Elementary	76.9	.32	74.6
Fox Road Elementary	80.6	.30	73.0
WCPSS: Grade 5	*87.1*	*.25*	*64.4*
All Grades Mathematics Middle			
East Millbrook Middle	68.9	-.01	49.8
East Garner Middle	75.0	.09	56.8
Carroll Middle	62.9	-.17	37.5
WCPSS: Middle Schools All Grades	*84.9*	*.10*	*62.3*
Algebra I Middle			
East Millbrook Middle	84.3	-.18	39.1
East Garner Middle	88.5	-.12	45.7
Carroll Middle	86.7	-.24	34.9
WCPSS: Middle Schools Algebra I	*95.3*	*.04*	*68.8*
Algebra I High			
Millbrook High	82.7	.05	54.4
Wakefield High	89.1	.15	63.7
Sanderson High	85.3	.40	78.1
WCPSS: High Schools Algebra I	84.9	.12	*62.5*

Proficiency percentages, academic change scores, and percentage of students who made growth at the three schools were also examined longitudinally, before the start of the initiative and after two years of implementation. Outcomes for the district were provided for comparison purposes. See Table 20.

- The proportion of students at Millbrook Elementary who made growth in 2011-12 compared to 2009-10 increased by about 28 percentage points and was higher than the district and any other participating schools.
- There was a five percentage point increase in 2011-12 in the proportion of students who made growth in mathematics (not including Algebra I) at East Millbrook Middle School.

This increase was not statistically significant, but did compare favorably with the district. At the same time, Algebra I results were lower in 2011-12 than in 2009-10, and decreases were larger than for the district for both proficiency and growth. Decreases in general may be related to the substantial increase in the enrollment in Algebra I classes due to a WCPSS policy change.

- At Millbrook High School, the proportion of students proficient in Algebra I increased from 2009-10 to 2011-12 by about 13 percentage points—more than the district. On the other hand, the percentage of students reaching growth targets did not improve at the school or in the district.

Table 20
2009-10 and 2011-12 Comparisons of Academic Change in Mathematics and Algebra I for Formative Assessment with Technology Project Participants

School	Percent Proficient		Academic Change		Percent of Students Who Made Growth	
	2009-10	2011-12	2009-10	2011-12	2009-10	2011-12
5 Grade Mathematics						
Millbrook Elementary	88.2	79.5	.07	.37	51.7	79.4
WCPSS: Elementary Schools	85.1	87.1	.17	.25	64.4	69.6
All Grades Mathematics						
East Millbrook Middle	67.6	68.9	-.09	-.01	43.6	49.8
WCPSS Middle Schools	83.6	84.9	.13	.10	62.3	59.9
Algebra I						
East Millbrook Middle	98.2	84.3	.03	-.18	56.4	39.1*
WCPSS Middle Schools	99.1	95.3	.25	.04	68.8	64.4
Algebra I						
Millbrook High	70.1	82.7*	.03	.05	56.1	54.4*
WCPSS High Schools	83.2	84.9	.15	.12	62.5	60.9

Note: * marks statistically significant positive or negative change.

Proficiency percentages and academic change scores for individual project participants were also examined longitudinally. As shown in Table 1 at the beginning of this report, a high proportion of teachers who were new at the school participated in the initiative in the second year of implementation (16/25). This was especially true in East Millbrook Middle School, where previous year's scores were unavailable for many new participants (see Table 21). At the elementary school, three of the five teachers remained during the second year of the project, at the middle school only two of the 12 teachers were participating in the second year, and four of eight participants in the high school were part of the initiative for the second year of the project.

Among individual teachers, EOG impact was not expected after one year. For teachers in the project two years, the three grade 5 teachers at Millbrook Elementary School showed considerable increases in the percentages of students who made growth and improved academic change scores. (Again, support of the mathematics coaches at the school may have contributed to this change.) At the middle school, one of the two teachers showed an increase in the percentage of students reaching growth targets, while the other showed a decrease. High school outcomes for the four two-year participants fluctuated and were inconclusive.

Table 21

A Longitudinal Look at Student Outcomes for Participants in the Formative Assessment with Technology Project

	Percent of Students Who Made Growth			Average Academic Change		
	2009-10	2010-11	2011-12	2009-10	2010-11	2011-12
Millbrook Elementary School						
Teacher 1	**na**	**30.4%**	**68.2%**	**na**	**0.045**	**0.174**
Teacher 2	**60.9%**	**61.9%**	**95.2%**	**0.077**	**0.196**	**0.548**
Teacher 3	na	47.6%	np	na	-0.072	np
Teacher 4	na	50.0%	np	na	0.005	np
Teacher 5	**45.8%**	**60.9%**	**86.4%**	**0.024**	**0.112**	**0.535**
Teacher 6	np	na	78.9%	np	na	0.347
Teacher 7	np	na	66.7%	np	na	0.244
East Millbrook Middle School						
Teacher 1	**69.3%**	**63.7%**	**81.0%**	**0.203**	**0.093**	**0.35**
Teacher 2	42.2%	32.1%	np	-0.119	-0.154	np
Teacher 3	30.4%	29.1%	np	-0.325	-0.149	np
Teacher 4	26.0%	23.3%	np	-0.299	-0.289	np
Teacher 5	**70.5%**	**54.8%**	**54.7%**	**0.188**	**.085**	**0.05**
Teacher 6	na	48.8%	na	na	-0.005	na
Teacher 7	na	na	49.1%	na	na	-0.04
Teacher 8	np	54.2%	30.0%	na	-.063	-0.22
Teacher 9	np	49.0%	56.5%	na	-.021	0.03
Teacher 10	na	na	no score	na	na	no score
Teacher 11	na	na	57.1%	na	na	0.03
Teacher 12	na	na	52.2%	na	na	0.07
Teacher 13*	na	na	na	na	na	na
Teacher 14	na	na	32.7%	na	na	-0.02
Teacher 15	na	na	40.2%	na	na	-0.06
Teacher 16	na	na	29.2%	na	na	-0.22
Millbrook High School						
Teacher 1	**na**	**53.3%**	**58.7%**	**na**	**0.117**	**0.00**
Teacher 2	**62.2%**	**46.9%**	**47.6%**	**0.094**	**0.013**	**-0.04**
Teacher 3**	30.0%	46.6%	na	-0.328	0.057	na
Teacher 4	**50.0%**	**72.4%**	**67.6%**	**-0.045**	**0.265**	**0.30**
Teacher 5	**44.4%**	**na**	**55.2%**	**-0.016**	**na**	**0.10**
Teacher 6	np	70.7%	63.4%	na	0.210	0.21
Teacher 7**	na	68.4%	na	na	0.104	na
Teacher 8	na	46.1%	55.8%	na	-0.074	0.01

Note: 1. na - no scores available (mostly new teachers or those who were not teaching at the school during the year).
np - not a participant (either not participating or left the school system).
2. Bold indicates teachers who were in the project for two years.
3. East Millbrook Middle: * a teacher had only a few scores available in 2011-12. ** Millbrook High: two teachers had no academic change scores in 2011-12.

Discussion and Recommendations

As an instructional process shown to improve student learning (Black & Wiliam, 1998), formative assessment should be promoted by educational leaders and practiced in all classrooms. The Formative Assessment with Technology Project was designed to provide teachers with training in the definition, value, and use of formative assessment. By including SMART Boards and related software to foster formative assessment practices, the project sought to expand the tools with which students and teachers could engage in this fundamental component of good instruction.

At the same time, the training model piloted here is not feasible to implement system-wide, in that it is too expensive in personnel time. School-wide models were previously found to be more cost effective with other district approaches or initiatives (SIOP® and the mathematics coaches). Also, if train-the-trainer models are found to be necessary, materials should be turnkey—ready for trainers to use. WCPSS is exploring creating teams of teachers at each school to serve as trainers for new initiatives.

A one-year model may be easier to implement than a two-year model if teacher turnover exists. While this report shows that over 90% of participants considered all components of formative assessment to have at least moderate value, implementation was uneven. The initial sample size of teachers was too small, and teacher turnover was too high for a solid experiment. Other training model issues existed, such as somewhat delayed observation feedback and limited structure at discussion meetings.

The following recommendations are offered to ensure the sustainability or expansion of similar professional development efforts:

- District leadership should provide consistent and focused support in providing access to training in the use of formative assessment. One source of such training is the online NC FALCON program, which provides solid content in terms of what formative assessment is. The more difficult part is determining the best way to help teachers apply the skills. In this pilot, implementation of different facets was still uneven at the end of two years and teachers did not feel competent with all the skills.
- School-level administrators should promote PLTs as venues in which teachers support each other in continually improving their formative assessment practices. The NC FALCON is a platform that PLTs can use in this effort, and it provides its own online forum that allows teachers and PLTs to connect with others making similar efforts across North Carolina.
- At the classroom level, teachers should go beyond the use of learning targets and criteria for success when planning the lessons and during instruction. They should follow through with all components of formative assessment, including variety of methods of collecting evidence of learning, intentional use of descriptive feedback, and deeper involvement of students in assessing and planning for their own learning.

- Students must be given opportunities to reference learning targets and criteria for success, and they should be expected to engage in peer review as part of the formative assessment process. An opportunity to talk to each other about their learning should be especially encouraged in middle school.

- The remarkable potential for using new technologies, such as SMART Boards and related instructional software, should be developed by instructional leaders and embraced by teachers. However, no tool in and of itself will embody the fundamental value and practice of formative assessment. It is up to teachers and students to act intentionally as proponents of learning, regardless of which technologies may or may not be available at a given time. Formative assessment is as old as the art of teaching and does not depend on new devices; it should be understood on its own terms before specific tools are relied on to enact it.

References

Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.

Bulgakov-Cooke, D. (2011). *Status of Formative Assessment with Technology Project, 2010-11*, WCPSS Data and Accountability Department. Cary, NC: Wake County Public School System.

Council of Chief State School Officers, (2008). *Attributes of effective formative assessment.* Retrieved from http://www.ccsso.org/Documents/2008/Attributes_of_Effective_2008.pdf

Appendix

Attributes of Effective Formative Assessment

From Council of Chief State School Officers (2008)

Learning Progressions

- Learning progressions should clearly articulate the subgoals of the ultimate learning goal.
    - Learning progressions enable teachers and students to use formative assessment to locate the students' current learning on the continuum and to identify the next steps.

Learning Goals

- Learning goals and criteria for success should be clearly identified and communicated to students.
    - This involves discussing the goals and criteria for success with students using terms they can understand and providing examples of how the criteria for success can be met.

Embedded in Instruction

- Teachers should evoke evidence about learning during instruction using a variety of methods that result in adjustments to teaching and learning.

Specific Feedback

- Students should be provided with descriptive feedback that is linked to the intended instructional outcomes and criteria for success.
- This feedback identifies the gap between current learning status and desired goals at a level of detail to stimulate action by teachers and students for improvements in learning.
- Feedback should focuses on enhancing student learning without assigning grades or scores.

Collaboration

- A classroom culture in which teachers and students are partners in learning should be established.
    - This involves ongoing interactions between teachers and students regarding learning goals, outcomes, achievements, and adjustments in learning activities. Students engage in self-assessment about how their learning is progressing toward desired goals. They are active agents in learning, working with teachers to close the gap between current learning status and desired goals.

Self- and Peer-Assessment

- Both self- and peer-assessment are important.
    - In self-assessment, students monitor their own learning using established criteria that indicate what a successful performance looks like, and they adapt their learning in order to achieve success.

    - In peer-assessment, students analyze each other's performance using established criteria and provide descriptive feedback to each other for continued improvement.

Formative Assessment Classroom Observation Checklist

Teacher-related Indicators	Present (√)	Use of Smart Board	Comment
Posts Learning Target and Criteria for Success			Board__ Handouts____
Targets and Criteria for Success are aligned			
Targets and Criteria for Success are written in student-friendly language			
Communicates about/shares Targets and Criteria for Success (refers to these during lesson)			
Provides instruction that is aligned with Criteria for Success			
Provides formative assessment strategies that are aligned with Criteria for Success			
Gathers evidence of learning			Whole class__ Small group___ Indiv. __
Collects evidence of learning using a variety of strategies (e.g., thumbs up, observation, questioning, math talk, exit tickets, etc)			Strategies used
Documents evidence of student learning (other than mental notes)			Strategies used
Analyzes evidence of student learning (e.g., teacher reflects on student info provided)			
Provides descriptive feedback on student work based on Criteria for Success			Whole class__ Small group___ Indiv. __ 1-2 3-5 more than 6 most all
Adjusts/modifies instruction based on evidence collected			
Checks for impact of feedback on student learning by providing other learning experiences			
Climate conducive for formative assessment; student and teachers partner in learning			

Student Behavior	Present (√)	Use of Smart B R	Comment
Reference Target in explanation			
Use technology to demonstrate learning (Type:)			# of students
Explain process/method used in work			
Use feedback to advance learning			
Provide feedback to others to advance learning			
Use criteria to explain “where he/she is” in learning			