ABSTRACT

HAYES, ANN MILLIGAN. From Assessment to Instruction: The Impact of Online Formative Assessment in Reading on Teachers’ Planning and Instruction in the Middle School English Language Arts Classroom (Under the direction of Carol A. Pope, Ed. D.).

Over the last two decades, there has been renewed interest in formative assessment, in large part due to the increasing pressures and prevalence of “high stakes” summative assessments. As states try to meet the requirements of the No Child Left Behind law, teachers and administrators are realizing that formative assessment offers an effective means of increasing student achievement and identifying specific learning gaps. In response to this recognition, many companies and institutions have designed online programs to assess students formatively. However, there has been little empirical research in how teachers use these programs in their planning and instruction, especially in the field of English Language Arts. This study examined how four middle school English Language Arts teachers used information generated by one or more online formative assessment programs to plan and implement reading instruction in their classrooms. The overarching research question for the study was: What is the impact of online formative assessment in reading on teachers’ planning and instruction in the middle grades English Language Arts Classroom?

Data sources included teacher interviews, classroom observations, and documents pertaining to the lesson and online formative assessment programs. Four themes emerged from the study, including obstacles teachers face in using online formative assessment systems (OFAPs), the manner in which they encourage metacognition in students, their abilities to leverage their experience for effectiveness, and their commitment to putting students first. Findings indicate that there is still a wide gap between the need for technology
in schools and its availability and that information concerning OFAPs, their availability, their purposes, and their features is not always complete; teachers consistently gather and analyze data from multiple sources and use it to inform all areas of their planning and practice; that teachers consciously use the components of formative assessment, such as feedback, student self-reflection, and goal-setting, to help students become better learners, and that teachers used formative assessment as an integral part of the classroom. Implications for English Language Arts instruction and suggestions for future research are provided.
From Assessment to Instruction: The Impact of Online Formative Assessment in Reading on Teachers’ Planning and Instruction in the Middle School English Language Arts Classroom

by
Ann Milligan Hayes

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

Curriculum and Instruction

Raleigh, North Carolina

2015

APPROVED BY:

_______________________________
Dr. Carol Pope
Committee Chair

_______________________________
Dr. Candy Beal

_______________________________
Dr. Peter Hessling

_______________________________
Dr. Carl Young
DEDICATION

Every great dream begins with a dreamer. Always remember, you have within you the strength, the patience, and the passion to reach for the stars to change the world.

--Harriet Tubman

To the children of my family who are also dreamers – may you always dream, and may you know the joy that comes not just from the fulfillment of your dreams but also the hard work you did to achieve them.
BIOGRAPHY

When she left Ohio for college in the Blue Ridge Mountains, Ann Milligan Hayes also adopted North Carolina as her home. Although her first loves were literature and writing, she was a generalist in college, studying sociology, religion, and English, vowing to never become an English teacher as her mother had – too many hours spent grading essays! At the last minute, she threw graduate school to the winds to become a lateral entry fourth grade teacher in eastern North Carolina. The next year, she earned a Master of Arts in Education from East Carolina University with a specialty in the burgeoning field of special education. She spent the next six years teaching learning disabled, mentally handicapped, hearing impaired, and academically gifted middle school students. Ann took the next nine years off to raise her two children, but during this period she taught preschool and was a teacher substitute for all grade levels. Upon her return to teaching, she taught middle school at-risk students in the Roanoke Rapids Graded School District, Chapter I program and began the inevitable transformation to English Language Arts teacher. Assessment for learning has always held a special interest for Ann, as she tried to find effective means for discovering what students had learned and what they still needed to learn. To this end, she was a member of the district Paideia implementation committee and the authentic learning task force while in Roanoke Rapids. Her work with at-risk eighth graders appeared in Peter Sacks’ (1999) book *Standardized Minds: The High Price of America’s Testing Culture and What We Can Do to Change It*. In 2002, Ann moved to Raleigh to teach at Centennial Campus Middle School on the campus of North Carolina State University and to pursue her dream of earning her doctorate in curriculum and instruction. After leaving teaching, she began a career in
formative assessment, first at North Carolina State University and then at Measurement Incorporated, where she oversees the content of five online formative writing sites. In 2013, she was awarded the first Margaret Dulaney Memorial Scholarship for Middle Grades English Language Arts/Social Studies for her student work and her contributions to education and NC State University. Her passions remain: students, learning, and English Language Arts. Her mission is, as always, to help students learn.
ACKNOWLEDGMENTS

The dream was always running ahead of me.

To catch up, to live for a moment in unison with it, that was the miracle.

--Anais Nin

This journey began with a dream that grew bigger and more insistent. At times, I doubted that the dream could come true. It was formed in large part with the students whom I had the privilege of teaching. My life is forever richer for having taught them, worked with them, and learned from them. The teachers with whom I worked refined the dream, as they shared the peaks and valleys that are an integral part of teachers’ lives. They, too, taught me more about teaching than they will ever know. My family, friends, and coworkers helped me visualize the dream. They offered love, support, and encouragement, and most important, believed in the dream even when I almost stopped believing. And Meg, who held my hand as we first pursued the dream together, and then continued to hold it when I had to go forward on my own. I am also grateful to members of my committee, Dr. Peter Hessling and Dr. Carl Young, who helped to shape the dream, generously sharing their time and expertise. Dr. Candy Beal, dream defender, was indomitable in fighting off doubt and despair. Dr. Carol Pope kept the dream in sight, gently prodding and pulling, coaxing and coaching, ever leading me forward, even when I resisted with all my might, so that one day I might catch it. Thank you.
TABLE OF CONTENTS

LIST OF TABLES ............................................................................................................ ix
LIST OF FIGURES ....................................................................................................... x

CHAPTER 1: Introduction ............................................................................................ 1
  Background .................................................................................................................. 1
  Emergence of Formative Assessment ................................................................. 2
  Data-Driven Instruction ....................................................................................... 9
  Online Assessment .............................................................................................. 10
  Teachers and Formative Assessment .............................................................. 11
  Statement of the Problem .................................................................................. 12
  Purpose of the Study .......................................................................................... 13
  Significance of the Study .................................................................................... 13
  Research Question .............................................................................................. 17
  Overview of Research Approach .................................................................. 18
  Definitions of Terms .......................................................................................... 19
  Organization of the Study .................................................................................. 21
  Summary ............................................................................................................... 21

CHAPTER 2: Literature Review .................................................................................. 22
  Background .......................................................................................................... 22
  Formative Assessment ......................................................................................... 24
  The Formative Assessment Model ................................................................. 31
  Formative Assessment and Adolescent Learning ......................................... 44
  Teaching Reading ............................................................................................... 56
  Teacher Planning, Decision-Making and Judgment ....................................... 62
  Concerns for Teachers ....................................................................................... 68
  Summary ............................................................................................................... 82

CHAPTER 3: Methodology ......................................................................................... 83
  General Introduction ............................................................................................ 83
  Background .......................................................................................................... 83
  Choice of Methodology ...................................................................................... 84
  Theoretical Framework ....................................................................................... 87
CHAPTER 5: Discussion

Introduction

Discussion of Findings
  The technology and information gap
  Making sense of data
  Helping students become better learners

Problems for OFAP Implementation in the English Language Arts Classroom

Policy and Budget Implications

Suggestions for Future Research

Conclusion

References

Appendix A

Appendix B: Interview Protocol: Pre-Observation

Appendix C: Interview Protocol: Pre-Observation
LIST OF TABLES

Table 1. Characteristics of Assessment vs. Grading .......................................................... 29
Table 2. Differences between Formative and Summative Assessment ............................. 30
Table 3: Teacher Participant Demographic Information (2012-2013) .................................. 92
Table 4. Demographic Information for Research Sites - 2012/2013 ................................ 96
Table 5. Assessment Programs Used by Teacher Participants (2012-2013) ......................... 155
LIST OF FIGURES

*Figure 1:* Diagram of Mastery Learning Tests and Adaptive Instruction .................. 26

*Figure 2:* The Formative Assessment Model (Hayes, 2013) ........................................ 33
CHAPTER 1: Introduction

Background

Assessment has become a pervasive element in the educational culture. It is perhaps the most discussed element of education today, and reported results to schools and districts have the power to retain students, affect teacher evaluations, and either provide or withhold millions of dollars. These results shape the thinking of legislators, administrators, and policy-makers as well as the general public. Just the mention of standardized testing can evoke intense feelings in students, teachers, administrators, parents, researchers, other educational stakeholders. As the number and impact of standardized, high-stakes assessments have increased, so have concerns about their legitimacy, relevance, and even negative impact on both students and teachers (Kohn, 2001).

In fact, many educators are pushing for an end to other kinds of summative assessments, such as student report cards, end-of-unit tests, and even graded classwork (Canady & Hotchkiss, 1989; Kohn, 1994). They argue that summative assessments, intended to “sum-up” a student’s achievement, do no more than compare one student to another, and require a situation in which some students pass and some must fail (Kohn, 1994). Stiggins (2002) further claims that “our current assessment system is harming huge numbers of students” (p. 758) and also posits that the increasing pressures of high-stakes testing means that students will begin to realize that success is unattainable and thus, instead of improving learning, these tests will result in students who no longer want to try.

Assessment, or testing, is not always summative, however, and most educational leaders do not wish to end altogether the accountability it can provide. Stiggins (2002)
suggests a reframing of assessment and its purposes. He points out that assessment itself is not bad, but we must change the way its results are used. Assessment reformers, such as preeminent experts Black & Wiliam (1998b) and Stiggins (2002) want the high-stakes summative tests replaced with formative models, which are given while students are still learning and guide teachers in what still needs to be taught or in what instructional interventions are needed. According to the State Collaborative on Assessment and Student Standards, a committee formed by the Council of Chief State School Officers (CCSSO), 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” (McManus, 2008, p.3). Thus assessment’s best purpose would be as a formative tool to guide learning.

**Emergence of Formative Assessment**

A type of formative assessment is used often in classrooms as teachers check for understanding through questioning, exit and entrance tickets or quick quizzes. However, the use of the more formal formative assessment model, including instruction, assessment, feedback, intervention, and reassessment, has become increasingly popular in the last two decades as a way to incorporate assessment into the learning process (Black & Wiliam, 1998a). This model is quite different from summative assessment, administered at the end of an instructional unit, a grading period, or a course to measure what the student has learned during that period. While formative assessment takes place during the unit of study so that an individual student’s growth and change in learning are clear, summative assessments come at the end of a unit of study, a year, or a course for the purpose of evaluating what the student
knows and does not know. Further, summative assessment scores or grades are awarded based on a range. The design of the range necessitates a distribution of scores, so that some are average, some are above average, and some are below average. Canady and Hotchkiss (1989) call this distribution “sorting and selecting” (p. 71). However, this arbitrary assignment does little to inform students, parents, or teachers about the information or skill the student is missing (Kohn, 2001). In contrast, formative assessment emphasizes the ongoing progress of the individual student and the ability of the student to continue to learn (Stiggins, 2004).

According to a policy brief on formative assessment published by the National Council of Teachers of English (2010), excellent formative assessment emphasizes the quality of student work instead of the quantity; coaching and encouragement for improvement rather than grades; student-teacher dialogue as opposed to lectures; and as many chances for success as necessary, concentrating on a few increments of understanding at a time.

To assure clarity, many educational researchers (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Stiggins, 2004) prefer the term assessment for learning to formative assessment because of the former’s clear differentiation from summative assessment, or the assessment of learning. The term assessment for learning clearly indicates the crucial role assessment should have in learning. Students are assessed as they learn so that teachers and students know what comes next – more practice, instructional intervention, or moving on to the next concept (Black & Wiliam, 1998b). In fact, an article by the Assessment Reform Group (2002) explains that assessment for learning is “the process of seeking and
interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there” (p. 2). Advocates of this approach claim that the use of formative assessment is a powerful tool for increasing student learning, and, as a result of better learning, student performance on standardized tests increases as well (Chappius & Chappius, 2008).

Studies have also revealed that formative assessment has a significant impact on teacher expertise, their planning and questioning practices, and their job satisfaction. In his study of an elementary school that implemented a school-wide formative assessment model over a period of several years, Baccellieri (2009) found that there was an increase in teacher-ownership of student learning as well as an understanding of how to plan best for student needs. For example, Bakula (2010), reporting on action research on using formative assessment in her science classroom, found that she became more responsive to student misunderstandings and more flexible in her planning in order to meet the specific needs of her students.

In a six-year study at a New Zealand elementary school incrementally implementing a school-wide formative assessment model, teachers reported that they had gained more expertise in their content areas because, instead of trying to meet general goals, they had to parse out the specific learning outcomes they expected from their students. This recognition also meant that they had to change the way they planned for instruction in order to meet students where they were in their understandings and move them forward from there instead of the arbitrary starting points in teachers’ guides. Finally, as a result of their shifts in practice and formative assessment processes, teachers began to feel that their focus became
more on student understandings and progress rather than the amount of material they covered (Jones & Moreland, 2005).

Formative assessment is actually not new in education. The philosophy, that ongoing evaluation of learning be partnered with immediate remediation, appeared in the educational theories of 17th and 18th century educators such as John Comenius and Johann Pestalozzi (Guskey, 1980) and Jean-Jacques Rousseau (1762/2013). During the eighteenth century, Rousseau claimed in both his novel Emile and later writings that educators must acknowledge that students go through different stages of cognitive development. Within each stage of development individuals differ, and thus the instructional approaches to their education need to be differentiated. According to Rousseau, each mind is unique, and thus teachers must teach what children can understand at their own level, without overwhelming them with more than they can grasp (Rousseau, 1762/2013). Building on these ideas, Russian psychologist Lev Vygotsky (1978) developed the theory of the ZPD in the early twentieth century, upon which much of formative assessment and its many predecessors are based. According to Vygotsky, the ZPD represents what children can do on their own as well as what they can do with help and support from an adult or older peer. Vygotsky considered this ZPD a more accurate measure of a child's intelligence than tests requiring children to perform tasks or answer questions on their own. The more scaffolding and support children receive from mentors, such as adults or more knowledgeable peers, the more they can achieve within the ZPD, and the better individual children will be able to perform in school (Chew, Jones & Turner, 2008). For educators, this theory means that new instruction should challenge the student so that the new task is above the student’s independent level but
still within the ZPD. In this model, the instructor is more of a facilitator who continually challenges individual students while guiding and helping them practice and learn (Chew, et al., 2008). The ZPD model is similar to the formative assessment model, in which teachers challenge students with new material while continuously giving feedback on their progress and checking to see where more support is needed in the form of re-teaching or remediation.

How children learn was also the focus of Jean Piaget’s (1963) seminal work in cognition which illustrated that learning is intrinsic and that children go through stages in their intellectual development so that learning occurs in a specific order, following a process called equilibration. It is through equilibration that children constantly correct and expand their understanding of the world. Piaget’s theory supports the formative assessment model, in that children learn to analyze their mistakes and correct them, leading to self-regulation, which, according to Piaget, is a natural state.

During this period, theorists were also developing theories based on Vygotsky’s (1978) Theory of Proximal Development. Michael Scriven (1979), an educational evaluation theorist, coined the terms formative and summative to differentiate between the two halves of evaluation – one with the intent to improve a product and the other that accepts or rejects it. John Carroll (1963) went against the prevalent thinking of the time that a child’s aptitude determined the amount he/she could learn by claiming that children differed in the time it took to master a concept, rather than in whether or not they could master it. His “Model of School Learning” set out five variables for learning as summarized below:

- aptitude, meaning the amount of time the student needs to learn a concept,
• opportunity to learn, meaning the amount of time given to the student to learn the concept
• perseverance, or the amount of time the student willingly spends on learning it
• the student’s ability to understand the instruction
• the quality of the instruction

This model foreshadowed the current formative assessment model in its emphasis on the variable of time.

Greatly influenced by Vygotsky, Scriven, and Carroll, Benjamin Bloom (1971) postulated that assessments used at that time merely supported what teachers already knew – that the material they had taught had been learned successfully by some students. He also acknowledged that most teachers taught all students in one group and in the same way. Given the diversity of learning styles and abilities, Bloom theorized that differentiating instruction to match the needs of each student, adding time for those who needed it, and re-teaching would allow all students to be successful. He maintained that assessments should be used to give students feedback and to serve as guides for correction (Bloom, 1968). Bloom could well be describing formative assessment, as it, too, uses assessment to provide student feedback and possible guides for correction (Black & William, 1998b).

Bloom’s (1971) research led, in part, to Outcomes Based Education (OBE), an approach in which specific students’ expected outcomes are identified before any instruction begins (Spady, 1994). OBE teachers focused on whether or not the student met the goal, or achieved the “outcome.” Re-teaching occurred as often as needed until the outcome was satisfactory. The origins of today’s formative assessment model can be clearly seen in OBE.
In both, teachers determine whether the student has met the goal, give feedback and then remediate as needed. While not a prerequisite of OBE, Bloom’s (1968) model for Mastery Learning, with its emphasis on short units of learning and assessment used as a learning tool, is usually an integral part of OBE models (Evans & King, 1994). Although OBE was controversial with some conservatives and suffered from the non-academic outcomes educational stakeholders incorporated into the design of the program, its inclusion as a predecessor of formative assessment is clear (Spady, 1994; Manno, 1994).

The current formative assessment model Stiggins (2002), Black and Wiliam (1998b) as well as other formative assessment proponents advocate components from all of these predecessors. Though some researchers use the term formative assessment interchangeably with assessment for learning, Stiggins (2002) differentiates the two because his model of learning assessment includes the student as a partner in the process, as do the current formative assessment models, whereas the formative assessment models in OBE and Mastery Learning do not. According to most formative assessment experts, the student’s role is this process is essential (Hebert, 1998; Stiggins & DuFour, 2009).

Sadler (1989) explains that for formative assessment and the feedback therein to be effective,

the learner has to (a) possess a concept of the standard (or goal, or reference level) being aimed for, (b) compare the actual (or current) level of performance with the standard and (c) engage in appropriate action which leads to some closure of the gap (p.121).
Thus students can, with support from teachers, set goals, track their progress, and become skilled in metacognitive strategies. Black and Wiliam (1998b) emphasize that “feedback to any pupil should be about the particular qualities of his or her work, with advice on what he or she can do to improve, and should avoid comparison with other pupils” (143). For the purposes of this study, the more familiar term, formative assessment, will be used to describe a model including feedback to students on their performance.

**Data-Driven Instruction**

Integral to formative assessment are data representing student performance. Since *No Child Left Behind* legislation of 2001 schools in increasing numbers have moved toward “data-driven instruction” (Schmoker, 2008). Education leaders hope that by gathering various kinds of data, analyzing and disaggregating it and then using it to inform instruction, student achievement scores on high-stakes tests will improve (Schmoker, 2008).

The kinds of data used to inform instruction in schools today is as varied as it is plentiful, resulting in educators’ trying to discover which strategy or tool will most likely give them the preferred results, and researchers’ questioning the validity of the data (Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006). Realizing that end-of-year state tests cannot provide teachers with information that allows for either planning instruction or effective intervention and to produce more immediate data, many districts have implemented interim, or benchmark, assessments, usually in the form of multiple-choice tests, to be given to all students several times a year (Olson, 2005; Heritage, 2007). Although the original intention for these assessments might be formative, they are of little use to the classroom teacher, giving too little information too late to affect learning. Since they are given only a few times
during the year, the benchmarks cover broad amounts of material, which makes intervention and remediation on specific objectives almost impossible. Teachers complain that students find the benchmarks intrusive, that there is little time left in which to teach and that the assessments have little or no impact on learning because they are too far removed from instruction (Olson, 2005; Martin, 2012). As a result, many educators describe these as being no more useful in guiding instruction than the end-of-the-year summative tests (Olson, 2005). In fact, according to Heritage (2007), they function best in predicting student performance on the end-of-year state tests. “What is missing in assessment practice in this country is the recognition that, to be valuable for instructional planning, assessment needs to be a moving picture” (Heritage, 2007, p. 141); that is, assessment must be continuous and ongoing rather than a one-time snapshot.

**Online Assessment**

As the thirst for data and awareness of formative assessment’s effectiveness have increased, educational materials and testing companies have entered to provide a plethora of assessments to be used formatively. Because of advances in and availability of technology, online assessments, both summative and formative, are quickly taking the place of paper-pencil ones. Online programs offer the advantages of testing many students at once while providing individual student modifications such as large text and read-aloud, as well as immediate scoring and feedback. Further, most of these programs offer various ways to quickly disaggregate the data produced to give teachers a more complete picture of each student’s progress (Bulkley, et al., 2010).
However, the ability of an online program to assess students during instruction and give immediate feedback does not make it formative. As Bulkley, et al. (2010) explain, **how** the assessment is used determines whether or not it is truly formative. An effective formative assessment is part of a cycle, or process. It is given during the instructional unit, provides detailed student feedback (not just scores), is analyzed to ascertain next steps in improving student understanding of concepts, and forms the basis of follow-up instruction and/or intervention (Nicol & Macfarlane-Dick, 2006). Students are then reassessed to determine their improvement. This kind of assessment actually matches what Bloom (1971) suggested and what effective teachers try to do in classrooms in a less formal way – determine gaps in each student’s understanding and intervene as needed to fill those gaps.

**Teachers and Formative Assessment**

Educators, researchers, and leaders in the field of formative assessment, point out that although districts can set requirements, provide training, and even choose the strategies, tools and/or programs, in the end, “learning is driven by what teachers and pupils do in classrooms” (Black & Wiliam, 1998b, p.139). Only by teaching more effectively can teachers help their students learn now and enable them to continue to learn (Black & Wiliam, 1998b).

The increasing demands for high standards, more instructional content, more accountability, and fewer resources, combined with the ever-growing numbers of students and students with special needs in classrooms can be overwhelming. Many states have approved including student achievement scores as a part of a teacher’s evaluation. Teachers are often the ones who must conduct the data analysis on data gleaned from the district
benchmarks and then show that they have modified their instruction based on the data (Olson, 2005). Adding still more assessment, whether formative or otherwise, may just seem to be too much (Heritage, 2007).

**Statement of the Problem**

As summative high-stakes test pressure has increased, the use of formative assessment as an effective online tool in the classroom has garnered increased attention (Bulkley, Oláh, & Blanc, 2010). Also of ongoing concern is reading proficiency at the middle school level. Based on Black and Wiliam’s (1998a) literature review of classroom formative assessment, the use of the formative assessment model has proved to be beneficial in increasing student achievement and responsibility for learning. Brookhart, Moss, and Long (2008) and Bakula (2010) describe specific schools and classrooms in which student achievement was increased through the use of formative assessment strategies. Online programs for formative assessment are a natural addition to the many commercial formative assessment programs available.

Web-based formative assessment programs allow instant evaluation and feedback for the student and teacher, achieving the critical timely feedback required for an effective formative assessment model. Despite widespread adoption by school systems and claims of increased student achievement, there is a gap in the literature on how teachers are actually using these online assessments in the classroom to impact instruction. This study sought to fill that gap with information about how teachers in middle school English Language Arts classrooms use the information generated by the assessments to plan and implement instruction in reading.
Purpose of the Study

To address the use of online formative assessment programs (OFAPs) in teaching reading, this study explored how four middle school English Language Arts teachers used information generated by one or more OFAPs to plan and implement instruction in reading. Effective online formative assessment programs are designed to follow the research-based formative assessment model, in which teachers use constant, ongoing formal (tests, essays, quizzes) and informal (questioning, entrance and exit tickets, etc.) assessments to track student progress and then give specific feedback to students. Teachers then plan modified instruction or other kinds of intervention to fill identified gaps in student knowledge and understanding. The model is completed when instructors reassess students through these same methods to check on whether or not the gaps have been eliminated. This study will use this recommended formative assessment model, using an online program to assess students, provide information on student achievement of the instructional objectives to both teachers and students, and for reassessment.

The study, therefore, will contribute to the educational knowledge base of best practices, especially in reading and formative assessment, by revealing teachers’ experiences in using online formative assessment programs.

Significance of the Study

With the advent of the landmark No Child Left Behind legislation of 2001 (Public law 107-110,2002) and its subsequent revisions, an increasing amount of attention has been focused on student literacy and mathematics achievement (Duncan, 2013). Conley and Hinchman (2004) identify a gap between the research findings about what impacts adolescent
literacy and the focus of plans to address the problems. Further, while intervention programs for grade three and below have been funded, money for adolescent students is concentrated on creating accountability systems and funding the transformations of schools that do not make the adequate yearly progress (AYP) toward certain educational goals as measured by mandated state tests. According to Haynes (2011), our nation’s fourth grade students score extremely well when compared to fourth graders of other nations. However, based on their performance on the National Assessment of Educational Progress (NAEP), students between the ages of thirteen and seventeen are lagging. In fact, the literacy of these adolescents “has remained stunningly low” (p. 10).

Middle school teachers face problems unique to the their level. They deal with larger classes than those in elementary school with smaller percentages of the student’s day devoted to literacy, making individualized or even small-group instruction extremely difficult. In a 2007 study, McMurre found that elementary students receive 503 minutes of ELA instruction per week compared to 331 minutes per week for middle school students. Of the time, middle school students are expected to master a total of eighty standards in reading literature, reading for information, writing, language (grammar, vocabulary, and conventions), speaking and listening, and literacy in history/social studies, science, and technical subjects as determined by the Common Core State Standards (2014) which have been adopted in forty-four states (Academic Benchmark, 2015). Technology integrations and research are included in these standards.

Another problem for ELA teachers at this level is the emphasis on literature over reading. This focus often results in organizing instruction by literary genre rather than by
specific learning goals and objectives. One unit, for example, might be short stories and include instruction in plot, characterization, and theme as well as point of view, figurative language, and literary devices. In the next unit, perhaps on poetry, much of the same subject matter is repeated. This practice has required that multiple goals and objectives be covered during the course of the unit of study, resulting in an overwhelming number of skills to be mastered by the student and monitored by the teacher. Worse, when they assess what students know and do not know, teachers are unable to target specific skills, because they are assessing multiple skills at once. The data they receive from assessments of an entire genre is often too general, and certainly too broad, to be of real help in identifying a specific gap in a student’s understanding. Commercial publishers, with their collective ears to the ground for shifts in theory and practice, have published anthologies organized by theme and learning objective for over a decade, and numerous school administrators require that teachers post the day’s learning objective in their classrooms. Still, many English Language Arts teachers are resistant to a more objective-based teaching style, believing that it may lessen the impact of the literature they love. There appears to be little research on this subject, but until this change happens, students will likely not make the kinds of gains on reading scores policy-makers expect.

Perhaps the largest hurdle for middle grades ELA teachers is the lack of experience in teaching the skills of reading. Not only do middle school ELA teachers get little or no training as pre-service teachers, but often colleges and universities do not require to provide any reading methods course for middle school teachers. Further, most reading instruction now ends at fourth or fifth grade (Allington, 2011). Yet many students entering middle
school have not mastered basic reading skills. What little intervention is given struggling readers at the middle school level is most often instruction in decoding, or translating the printed word into sound, a misguided approach, Allington points out. While students at this level are not necessarily proficient in phonics and phonemes, they have much more difficulty with vocabulary and comprehension (2011).

Heritage (2007) claims that high stakes, summative assessments have caused assessment to be seen as a way to “rank” schools, rather than as a resource teachers can use, and that the “reciprocal relationship between teaching and assessment has been lost from sight” (p. 140). Based on her experience in professional development at the University of California’s National Center for Research on Evaluation, Standards, and Student Testing, Heritage states that even teachers have begun to consider assessment as something that is completely separate from what happens in the classroom. In order for any formative assessment model to be effective, both teaching and learning must be modified, and thus, the changes a teacher makes in instruction will greatly affect its success (Black & Wiliam, 1998b).

Study after study has clearly shown that formative assessment is effective in increasing learning outcomes for students in kindergarten through college (Stiggins & Dufour, 2009; Black & Broadfoot, 2004). Assessing specific learning goals during instruction, giving detailed, individual feedback to students about their performance, and re-teaching or remediating for those who need it often not only affects classroom learning but raises achievement scores on summative assessments (Black & Wiliam, 1998b).
To date, however, there have been few studies on specific elements of teacher planning and instructional change. Existing studies have primarily focused on science and mathematics classrooms or whole-school initiatives. For example, Ruiz-Primo and Furtak (2001) studied classrooms in which science teachers incorporated formative assessment strategies, and Wiliam, Lee, Harrison, and Black (2004) conducted a study with science and mathematics teachers from six different schools in two different school systems in England who implemented formative assessment practices. Others have focused on foreign-language learning, technology, and elementary school classrooms (Jones & Moreland, 2005; Baccellieri, 2009). To fill an existing gap in the research on OFAPs and reading, this study will examine middle school English Language Arts teachers’ planning and practice as they use OFAPs.

**Research Question**

Creswell (2007) describes constructivists as those who, in trying to understand their world “develop subjective meanings of their experiences . . . formed through interaction with others (hence social constructivism) and through historical and cultural norms that operate in individuals’ lives” (pp. 20-21). Teachers are constructivists in that by constantly reflecting on their practice, their students’ abilities, learning styles, strengths and weaknesses, successes and failures, and incorporating their own experience, knowledge and a plethora of strategies, methodologies, and theories, they construct meaning in order to make instructional decisions.

In this study, I look at the results of this construction by participants using online formative assessment data, their own knowledge of their students, and of their subject to make instructional decisions about the best kind of intervention to increase student
understanding. Thus, this study answers the question: What is the impact of online formative assessment in reading on teachers’ planning and instruction in the middle school English Language Arts classroom?

**Overview of Research Approach**

This research study will take a qualitative, social constructivist, multiple-case study approach following four middle school English Language Arts teachers as they use student data from an online assessment program to help guide curriculum decisions. Qualitative studies are identified by four essential characteristics:

- The study is based on a search for understanding and meaning, rather than to prove or disprove a theory.
- The primary instrument is the researcher.
- Inductive, rather than deductive, investigation strategies are employed.
- The final product is characterized by rich description (Merriam, 2009, p. 39).

In qualitative research, the terms *interpretist* and *constructivist* are often used interchangeably (Merriam, 2009, p. 9). The interpretist/constructivist realizes that there is no “single, observable reality . . . Researchers do not ‘find’ knowledge, they interpret it” (Merriam, 2009, pp. 8-9), and meaning is constructed by those experiencing the phenomenon. The qualitative researcher, serving as the instrument, interprets data to find themes that lead to possible substantive theory (Merriam, 2009). In this study, I will look for common elements in the process the teacher participants use to determine appropriate practice to modify instruction in response to formative assessment data.
The study lends itself to the case study approach, as only four teachers and their use of an online formative assessment system will be studied. According to Hatch (2002), “this quality of being bounded is that which sets the case study apart from other qualitative studies” (p. 30). The study’s participants will be middle school language arts teachers who are fully licensed, who regularly use an online assessment system formatively in their classrooms, give feedback based on the data it provides, and use class time to re-teach or remediate based on its data.

For the purposes of this study, the formative assessment cycle (FAC) will be defined as having six parts: pre-assessment, analysis of data, feedback to students, planning for modified instruction, intervention and assessment. Through data collection and analysis, common themes describing how teachers use data to modify instruction will emerge.

**Definitions of Terms**

For the purposes of this study, these terms will be used as defined below:

*Assessment* – an evaluation of what students know and can do on identified educational goals. *Feedback* – information given to students about their performance and/or achievement on a particular assignment or assessment is called feedback. Feedback may be given orally or in written form; formally, as in a grade or rubric; or informally, as a comment or discussion with individual students or a group of students. Feedback may be used for both formative and summative assessment, but is particularly helpful for formative assessment.

*Formative assessment* – according to the State Collaborative on Assessment and Student Standards, formed by the Council of Chief State School Officers (CCSSO) to define and study the term, 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” (McManus, 2008, p.3).

*Metacognition* – knowledge individuals have about their own cognitive processes as well as the control they have over these processes, including organizing, monitoring, and adjusting them as a part of leaning. Metacognition includes the ability to self-regulate and reflect on the process of learning, and is necessary for students to become independent learners.

*Online formative assessment program* (OFAP) – Internet-based programs designed to give immediate feedback to teachers and students on student progress toward specific, incremental educational goals in reading, mathematics, and/or other content areas. Therefore, most offer multiple choice assessments that can be scored electronically, but formative assessments for writing are also available. Assessments may be prepackaged or constructed by the teacher and usually contain only a few questions or items covering only one or two learning objectives. Reports for entire classes and individual students should be available to show progress by standard and over time.

*Reading comprehension* – a literary experience during which meaning is made from text.

*Summative assessment* – assessments that come at the end of a unit of study, a year, or a course for the purpose of evaluating what the student knows and does not know. These assessments often result in a grade or a final score. Standardized tests fall into this category.
**Organization of the Study**

Chapter Two provides a review of pertinent literature, including the development and purpose of formative assessment, impact of formative assessment, the pedagogy and practice changes in teachers necessitated by formative assessment, and uses of computer-based programs in formative assessments.

Chapter Three outlines the methodology which will be employed in the study, including the rationale for the methodology, the types of data that will be collected in order to discover how middle school language arts teachers change instruction in response to data. The chapter includes an overview of how the data will be collected and analyzed. In addition, subjectivity, validity, reliability, and ethical issues, as well as the limitations of the study are discussed.

**Summary**

This chapter has explained the background and rationale for this research. It has included the statement of the problem of needing to know more about the process of teacher analysis, planning and modification of instruction in response to formative assessment data, the purpose and significance of the study and the questions guiding the research. Terms necessary to the understanding of the research were defined, and the overall approach was explained.
CHAPTER 2: Literature Review

Background

While high-stakes, summative assessments seem to be a permanent fixture in the current educational system, concerns continue to grow about their effectiveness in accurately assessing student growth and proficiency, their influence on the educational climate, and their impact on student learning. The worries of educational scholars about these kinds of tests are not new, yet the number of these tests and their impact on students, teachers, schools, districts, and states has only increased (Black, Harrison, Lee, Marshall, & William, 2004; Wiliam, 2010).

In his overview of the history of high stakes testing for accountability, Wiliam (2010) traces concerns with the practice as far back as 1886 and Emerson E. White’s Elements of Pedagogy, who charged that this kind of testing created an environment where learning was mechanical, students learned to cram and memorize instead of think, made students and teachers less than honest, and increased pressure on both.

The recent increase of high-stakes testing is mainly due to the No Child Left Behind legislation of 2001. Based on this Congressional Act, students can be held back a grade based on their performance on state-administered summative tests, and pressure to perform well permeates most of their school years. However, students are not the only ones held accountable by the tests. Mandates to prove that students have mastered the standards in their state’s curriculum have become a burden to all stakeholders. The tests carry heavy weight on teachers’ performance evaluations, and, thus, their jobs. Districts and states are granted federal money based on the same tests (“Public law pl 107-110,” 2002). Further,
tests are now being required in the primary grades. In their report on reading assessments given in kindergarten through third grades, Paris and Hoffman (2004) conclude that “the gulf between what teachers value as informal assessments and what is imposed on them in the form of standardized testing appears to be broadening. Although performance assessments and portfolios were popular in the 1980s and 1990s, the trends today are to increase high-stakes testing for young children, to remove teacher judgment from assessment, and to streamline assessments so they can be conducted quickly and repeatedly” (p. 205). Sacks (1999) notes three other serious problems with high-stakes testing: Their ability to predict academic success is questionable, their correlation with socioeconomic class is high, and they reward “passive, superficial learning, drive instruction in undesirable directions, and thwart meaningful educational reform (pp 7-8).

Many researchers and educational experts agree and contend that these summative, end-of-year, one-time snapshots are of little or no use in assessing what students really know and can do, and worse, actually have a negative impact on students, teachers, and learning (Kohn, 2001). While some experts suggest doing away with current methods of testing and grading, including standardized tests and even classroom grades (Canady & Hotchkiss, 1989), others push educators to use assessment for learning (Broadfoot, et al., 2002). Assessment for learning, or formative assessment, has become popular as a way of checking on student learning during instruction, so that interventions can be made immediately to correct any misunderstandings (Black & Wiliam, 1998b).

The question this study seeks to answer is “How does online formative assessment in reading impact teachers’ planning and instruction in the middle school English Language
Arts classroom?” To provide background for the study, this chapter, the literature review, focuses on four main topics. The first explains formative assessment, its development, purpose and model. Within this broad section is a section on adolescent learning theory and how it relates to formative assessment, given that the study will focus on middle school classrooms. Because this study is concerned with formative assessment in reading, the second section addresses the teaching of reading. The study will examine teacher planning in response to formative assessment data, so teacher planning and concerns for teachers using formative assessment make up the third section. Because the study deals with online formative assessment programs (OFAPs), a section on online assessment has also been included.

**Formative Assessment**

Formative assessment has been on the educational horizon for many years. In fact, according to McManus (2008), Scriven (1967) first called it *formative evaluation*, and identified it as the opposite of summative evaluation that occurs at the end of a curricular program (p. 19). As early as 1963, however, Carroll (1963) had theorized that, with the right amount of time, all students could master the necessary concepts of a given course as well as overcome five other personal and experiential differences: perseverance, aptitude, ability, opportunities to learn, and quality of instruction. Thus differing achievement levels in students could be overcome through support in the form of more time, even when other factors important to learning were missing. The current practice of testing all students on the same concepts at the same, predetermined time puts too much emphasis on time and too little
on mastery. Allowing students time needed to master concepts is one of the essential tenets of formative assessment.

In 1968, Bloom’s “Learning for Mastery Model” further developed Carroll’s ideas to show how each of the five differences Carroll identified could be addressed in the classroom (Zimmerman & Dibenedetto, 2008). Bloom had been researching how and why student achievement differed among individual students, which led to his finding “that there was much variation in students’ academic learning in traditional classrooms, with their scores typically forming a normal distribution or bell-shaped curve” (Zimmerman & Dibenedetto, 2008, p. 208). He also found that teachers generally used little variety in their instruction, allowing the same amount of time for all students to master concepts and move on to the next topic whether or not they had, assuming that student differences could not be overcome. Bloom maintained that there should, instead, be variety in instruction and little similarity in student mastery of concepts (Guskey, 2007).

This research eventually resulted in Bloom’s Learning for Mastery Model in which students would work on the same topic until their mastery level is at least 80%. His model included four elements: defining, planning, teaching and grading for mastery (Bloom, 1971). Bloom’s model also incorporated formative evaluation as one its three steps (Bloom, 1968). Figure 1 shows how these steps were used within the classroom context (Zimmerman & Dibenedetto, 2008, p. 209).

Bloom thought that this model would allow almost all students to achieve at the level that only the highest level students had before, and that weaker students would need less time to master learning objectives once they had mastered basic skills as well as the mastery
learning method (Bloom, 1968). Although Bloom used Scriven’s (1967) term *formative evaluation*, the term eventually evolved to *formative assessment* (McManus, 2007). Bloom’s model, shown in Figure 1, shows the same steps and the important circular characteristic as formative assessment. For clarity, the terms *formative assessment* and *assessment for learning* will be used interchangeably for the purposes of this study.

*Figure 1:* Diagram of Mastery Learning Tests and Adaptive Instruction (Zimmerman & Dibenedetto, 2008, p. 209).

To support the formative assessment model, researchers often refer to Vygotsky’s ZPD (ZPD) Theory (Poehner, 2005; McManus, 2007). This theory holds that the difference between what a child knows and can do independently, and what that child can learn must be
supported or scaffolded, with help from an adult or peer (Shabani, Khatib, & Ebadi, 2010). Thus, teachers should start at the point of current knowledge, coach the student through the new material, giving whatever assistance is the student’s ability to transfer this knowledge or skill to other tasks (Poehner, 2005). This theory is integral to formative assessment in that these assessments show the ZPD, so that the teacher can then help the student cross to mastery.

Outcomes-based education was another instructional model similar to previous models focusing on student mastery of very specific learning goals. In Outcomes Based Education (OBE), learning is measured by what students can do at the end of any given curriculum. William Spady (1994) describes OBE as producing “actions and performances that embody and reflect learner competence in using content, information, ideas and tools successfully” (p. 2). Spady cites familiar models of OBE, including military training programs, scouting merit badges, parenting, flight schools, and professional licensure of doctors and lawyers, to name a few. These examples are more concerned with “clearly defined performance” and “what and whether students learn successfully [rather] than when and how they learn it” (Spady, 1994, p. 14).

The influence of OBE still appears in the way learning objectives are written. “When defining and developing outcomes, educators must use observable action verbs - describe, explain, design, or produce – rather than vague or hidden non-demonstration processes – like know, understand, believe, and think” (Spady, 1994, p. 13). Curricula over the last two decades have been rewritten to incorporate these terms.
As formative evaluation or assessment evolved, researchers used other terms to express the practice of assessing students within the unit of study to help form ongoing instruction, such as Black and Wiliam’s preferred term *assessment for learning* (Black, et al., 2004; Wiliam et al, 2004). Whatever the term, researchers make clear distinctions between formative and summative assessments. Summative assessments are for accountability purposes for students, but also for teachers, schools and curricula, as well. They come at the end of an instructional unit, grading period, or school year, include teacher-designed as well as standardized tests, and result in grades, scores, reports, and at times, decisions on promotion and graduation. Formative assessment is for the purpose of increasing student learning and understanding (Black & Wiliam, 1998b). According to Bloom (1971), “The most important value of formative evaluation, in our view, is the aid it can give the student in his learning of the subject matter and behaviors for each unit of learning” (p. 129).

This view is supported by Pope and Beal (2002), who explain that grading students is often like pronouncing a final verdict on their learning, when learning is ongoing. They point out differences between their formative methods of assessments and the more common summative ones in Table 1 below.
Table 1. *Characteristics of Assessment vs. Grading*

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Grading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proficiency focus</td>
<td>Deficiency Focus</td>
</tr>
<tr>
<td>Instruction</td>
<td>Verdict</td>
</tr>
<tr>
<td>Avoids Separation</td>
<td>Categorizes</td>
</tr>
<tr>
<td>Motivates</td>
<td>Limits</td>
</tr>
<tr>
<td>Aids Learning</td>
<td>Aids Adults</td>
</tr>
<tr>
<td>Builds Scaffolds</td>
<td>Creates GPAs</td>
</tr>
<tr>
<td>Stimulates</td>
<td>Controls</td>
</tr>
<tr>
<td>Encourages</td>
<td>Stymies</td>
</tr>
<tr>
<td>Open System</td>
<td>Closed System</td>
</tr>
</tbody>
</table>

Pope & Beal, 2001, pg. 68

Researchers have approached the characteristics of formative assessment in several ways. In her study of a balanced assessment program including formative assessment and its impact on mathematics instruction, McManus (2008) differentiated between formative and summative with characteristics of time, feedback and definition, as shown in Table 2.
Ruiz-Primo and Furtak (2007) divide formative assessment into formal and informal and planned and unplanned. Formative assessment for Heritage (2007) consists of three categories: on-the-fly assessment, planned-for interaction, and curriculum–embedded assessment (p. 141). On-the-fly assessments are unplanned and are the result of a teacher spotting a misconception during a lesson. Planned-for interactions are assessments that teachers have planned to use to gather information, such as questions they will ask during a lesson to elicit information.

Curriculum-embedded assessments can be either those designed to determine progress at specific points in the school year or those that a regular part of classroom activities. These assessments may appear as quick quizzes or exit tickets, informal questioning, class discussions, group work, homework, written classwork, writing assignments, projects, games, formal quizzes, portfolios, work folders, and tests. The assessments may also appear as more than one type. For example, written classwork will most likely be identified as
planned-for interaction, but in some cases, it might be an on-the-fly assessment. Together, these descriptions begin to form a formative assessment model.

**The Formative Assessment Model**

While almost any activity may be used as a formative assessment, teacher feedback to students, remediation, and reassessment are essential. Other experts agree that assessment as used in formative assessment is ongoing, occurring constantly in the classroom. Heritage (2007) further claims that for assessment to contribute to instructional success, it “needs to be a moving picture – a video stream rather than a periodic snapshot” (p. 141). Formative assessment, according to Greenstein (2010) helps teachers

- Consider each student’s learning needs and styles and adapt instruction accordingly
- Track individual student achievement
- Provide appropriately challenging and motivational instructional activities
- Design intentional and objective student self-assessments
- Offer all students opportunities for improvement (p. 26).

McManus (2007) delineated four steps in the process of formative assessment. In her study of four high school mathematics teachers and their implementation of formative assessment, McManus (2007) concluded that these steps were essential (pp. 80-81).

Step 1. Identification of learning targets and criteria for success.

Step 2. Elicitation of Evidence of Learning.


Step 4. Implementation of Action(s) to Close the Gap
Another model is presented by Pope and Beal (2001, p.70). They use the acronym D.A.N.C.E. to represent their formative assessment model.

- Describe what is in the student work
- Account for what is in the student work by relating it to the expectations of the assignment, the goal of the work.
- Nudge students by questioning, guiding, suggesting, probing, and prodding
- Compromise with students by negotiating gaps, goals
- Envision with students, set plans for action collaboratively.

Although different models of formative assessment have been devised, the common components include clearly communicated learning goals, instruction, assessment, descriptive feedback to students, instructional adjustment or modification (opportunity for students to improve), and reassessment. Further, these components are cyclical. These components and their relationship to one another are shown on the following page in Figure 2.
In their well-known two-year study, Bell and Cowie (2001) were interested in developing a model for the process of formative assessment. Their research, focusing on the practices of ten middle and high school science teachers in New Zealand, included researcher observation of assessment activities and examination of teacher and student views on assessment. The third part, carried on throughout the study, consisted of researchers and
teachers participating together in professional development, discussion reflection, and work toward a model. Qualitative data collected included surveys, interviews with students and teachers, observations, and taping the professional development discussions.

Bell and Cowie’s (2001) formative assessment process model has five parts: 1) teachers use both planned and interactive formative assessment; 2) formative assessment is a complicated, multi-faceted task requiring teacher knowledge on many levels; 3) purpose of the assessment dictates the type of assessment (planned or interactive) used; 4) formative assessment is a valuable, necessary part of instruction; and 5) teacher awareness of their formative assessment practices and the results of the study changed teacher reflection and their understanding of formative assessment.

The Impact of Formative Assessment in Classrooms

Not only did formative assessment practices result in changes for teachers, as Bell and Cowie’s model illustrates, but researchers also point to increased student achievement through formative assessment (Black & Wiliam, 1998a; McManus, 2008; Stiggins & Dufour, 2009; Bakula, 2010). Block and Anderson’s (1975) research review of mastery learning found that mastery learning techniques improved student learning in spite of differences in classroom settings, subject content, and class size.

Kulik, Kulik, and Bangert-Drowns (1990) also found an increase in student achievement. Their meta-analysis of mastery learning programs found that, in 55 of the 108 of the studies, there was not only an overall increase in student summative examination scores, but the gains were even more profound for lower-achieving students in mastery learning classes than in traditional classes. The greater gains for the lower-achieving
students resulted in a significantly lower variation in the scores for mastery learning students, effectively narrowing the perennial achievement gap. Kulik, et al. (1990) note that, “few educational treatments of any sort were consistently associated with achievement effects as large as those produced by mastery teaching” (p. 292).

Stiggins and DuFour (2009) describe a school in Virginia that used formative assessment to change their school in two years from one whose achievement scores were well below the state average to one nationally recognized with proficiencies that went as high as 100%. The school incorporated a revised schedule to allow targeted remediation on each skill by the teacher most successful in teaching that skill, while others provided enrichment for the rest of the students.

Targeted instruction was also part of two studies described by Bloom (1984). In the parallel studies replicated four times in two different subject areas, student proficiency was measured on summative tests after eleven sessions of instruction over three weeks to determine if the type of the instruction made a difference in student learning. In each replication, students were divided into three groups: whole-class instruction with 30 students using summative assessment, whole-class instruction with 30 students using formative assessment, feedback and remediation, and one-on-one tutoring with formative assessment, feedback and remediation. Unsurprisingly, the tutored students scored highest, but the formative-assessment whole class scored significantly above the students who were only given summative assessments.

Student performance was also the focus when Ruiz-Primo and Furtak (2007) studied the practices in formative assessment of three science teachers to find out if consistent use of
formative assessment strategies would increase student success on the end-of-unit assessment. The mixed-methods study incorporated coded video tapings of classes and comparison of student performance on pre-and post-assessments. These authors defined a cycle of assessment as having four steps, which they termed ESRU: Elicit (teacher asks a question); Student (student response); Recognizes (teacher recognizes the student responses); and Uses (the information the teacher gathers from the student response is used to support student learning). This model collapses the formative assessment model into four steps: Elicit: assessment; Student: assessment; Recognizes: data analysis; Uses: feedback and targeted intervention. Ruiz-Primo and Furtak also described the two types of formative assessment as formal and informal. Ten lessons per teacher were studied, transcribed and coded. Students of all three teachers were given a pre and post-test, and results were analyzed. While the study focused on only three teachers, the number of videotaped class sessions indicates a definite link between the model and student success on the post-test. The results further indicated that the more consistently a teacher used the formative assessment model (ESRU), the better students did on the post-assessment.

Further support for formative assessment was reported by Wiliam, Lee, Harrison and Black (2004), in a study of British math and science classes. Twenty-four teachers at the middle and high school levels volunteered to use certain types of formative assessment, or teaching for learning, with one of their classes. They changed their practice from that of teachers bearing all responsibility for learning to one in which responsibility for learning was shared by teachers and students alike. Results were measured by written teacher reflections as well as comparison of scores on standardized tests between experimental classes and
control classes. Improved questioning, feedback without grades, peer and self-assessment, and formative use of summative testing were all methods used and reflected on by teachers as well as the researchers. The gain on end-of-year summative tests averaged 0.03 standard deviations, a significant gain. However, the teachers realized that the changes they made were effective long before the end of the year and began using the strategies in their other classes as well. Although the authors began the study with science and mathematics teachers, they expanded the study to other content areas (Wiliam, Lee, Harrison & Black, 2004).

Other positive results from the use of formative assessment have been documented as well. A study conducted by Treagust, Jacobowitz, Gallagher and Parker (2001) investigated the use of embedded assessment (formative assessments used within the regular curriculum) to increase student learning in a middle school science class. The class of 23 eighth graders was about evenly made up of Caucasian, African-American and Hispanic students, with a few Asian Americans and Native Americans. Students were from working-class families and their abilities in science were varied. The teacher was recognized as accomplished and experienced. She was involved in a professional development project in conjunction with a local university and several schools in her district, studying ways to use assessment more effectively. Until this project, she had never used pre-assessment and had not incorporated certain other kinds of formative assessment, feedback and discourse with her students. Research consisted of observation of the class daily over a three-week period and frequent interviews with both teacher and students.

The researchers then analyzed the lessons and established five “assertions,” which outlined the structure and types of assessments used. These assertions were that formative
assessments were defined by 1) use of activities across a broad spectrum; 2) use of pretests to determine instructional direction and content; 3) use of a wide variety of integrated writing assignments; 4) use of questions to refine student understanding; 5) student responses that varied with the tasks among individual student (pp. 143-153).

Tregust, et al. (2001) concluded that the change to embedded assessment changed the entire culture of the classroom from test–based to assessment-based. The results of the ongoing assessment dictated how the lesson was designed and delivered. The teacher reported that she was pleased with the unit and the increased knowledge she had of her students and their understanding as a result of her assessments. Further, students reported that they felt their opinions were valued, their understanding increased, and that they experienced a more caring atmosphere.

Parents as well as students sense the impact of formative assessment. In their study of first and fifth graders, Campos and O’Hearn (2007) found both students and parents believed that making the students more knowledgeable about their progress toward goals resulted in higher achievement. This phenomenological study looked at making students more responsible for their own learning by keeping them informed as to the mathematics goals for the year and their progress towards them. Pre- and post- surveys were given to both parents and students to determine attitudes toward math and knowledge of important math skills. The subjects were a class of 21 first graders and one of 21 fifth graders. The researchers were teachers at the school who were interested in changing their instruction to help students self-monitor learning with the hope that it would increase student learning and achievement.
The teacher-researchers first clearly communicated learning goals in student-friendly terms to both students and parents, then used formative assessment to track student progress, taught students to keep portfolios to track their own progress with graphs and then reflect on their progress, and provided time daily to remediate specific problems. Even the first graders were able to reflect successfully on their work and progress.

Formative assessment was also found to affect the attitudes of high school students in McManus’s (2007) study of the impact of implementing formative assessment in two low-performing high schools. McManus (2007) noted that both students and teachers reported positive changes in the classroom environment, and students became more involved and demonstrated more positive attitudes in their classes due to their increased share in the responsibility of developing rubrics, self and peer assessment, and learning. In addition, teachers reported that students were more respectful and turned in more homework. In these studies, student motivation increase as teachers became partners with their students, giving more control and responsibility of learning to the students.

**The School -Wide Impact of Formative Assessment**

The impact of formative assessment has been felt school-wide, as well. Studies such as one by Brookhart, Moss, & Long (2008) looked at whole-school change, finding that formative assessment was successful in impacting whole-school culture. With the dual goals of increasing communication between students and teachers and self-regulation in students, six teachers in one Pennsylvania district began working together in what eventually became a district-wide, three-year project to implement formative assessment and is now an integral part of education in the district. The district has seen an increase in achievement,
motivation, time on task, and engagement as a result of teachers’ use of formative assessment (Brookhart, et al., 2008).

School-wide impact was also documented in Baccellieri (2009)’s three-year study of a Chicago elementary school which used formative assessment as a way to increase student achievement. Hoping to increase student achievement on the required end-of-year standardized testing, the school began using formative assessment and collaboration to change their approach to instruction. Over several years, the assessments and the processes of collaboration were revised to fit the needs of the school and its students. Student achievement improved dramatically. Other changes as a result of the cycle of assessment, analysis and collaboration, remediation, and assessment included an increase in teacher satisfaction, development and learning, in addition to an increase in teamwork and teacher-to-teacher discussion. Perhaps even more important was an increase in teacher sense of ownership, or that they were making a difference, in the success of students. Baccellieri’s (2009) study is particularly interesting in that its concentration is primarily on the change in the professional climate of the school after implementing formative assessment and collaboration. While the results were based on teacher reports, it is clear that the implementation had significant impact on almost every aspect of teaching and learning.

Jones and Moreland (2005) found similar results in their multi-year New Zealand project, which began in 1998 as a study of eighteen schools that were incorporating technology as a curriculum for the first time. The researchers wanted to “investigate, develop and enhance teachers’ technology education teaching, learning and assessment practices” (p. 194) as the curriculum was implemented. Teachers were trained in formative
assessment practices, but also discussed their formative practices with students. Evidence was gathered through classroom observations, interviews with students and teachers, student work and “teacher documents” (Jones & Moreland, 2005, p. 194). According to the data collected, instructional changes in technology increased students’ achievement and understanding in all areas of technology, but also in their metacognition, their own questioning, interest and motivation. Both teachers and students reported that students were able to see what they still needed to learn and what they already knew.

At the request of the staff of one participating elementary school, the project was extended an additional three years and expanded to include electronics, science education and environmental education “because the principal and teachers saw the changes that were occurring in both teacher practices and student learning in technology and wanted to upscale these practices across the whole school” (Jones & Moreland, 2005, p. 194). By the end of the third year of the second study,

it became apparent that school-wide changes in assessment practices had begun to develop, so a further research project was undertaken by the authors in that year to explore the factors that contributed to practices at the individual teacher and school-wide level, as well as in subject areas other than technology (p. 194). Through interviews, the researchers found that teachers developed more expertise in their subject areas because they needed to focus on specific learning outcomes, rather than more general goals. Teachers themselves described changes in their student reports, explaining that they were “thinking about the actual learning rather than saying they [the students] are good at something” (Lucy, in Jones & Moreland, 2005, p. 203). Of special
interest in this study was the fact that teacher-participants reported that not only had they changed their assessment practices, but also their planning, their questioning, and their knowledge levels as well.

Formative Assessment with Portfolios

One of the most predominant and effective formative assessment tools is the portfolio, a collection of student work representing individual student progress toward learning goals. The use of portfolios in the English Language Arts classroom has evolved over the last thirty years from a way of collecting student work to a method of assessing students, including their performance on course objectives, growth over a period of time, and goals for the future. Additionally, in its best use, a portfolio provides opportunities for students to reflect on their own work and growth over time and to set goals for themselves.

Callahan (1995) neatly describes their value for education:

Within pedagogy, portfolios are seen as a way of encouraging process writing, collaboration, and student responsibility for learning. Within writing assessment theory, portfolios are widely viewed as the most potentially valid test of direct writing, and within curriculum design, portfolios are often proposed as an effective way to encourage faculty development and to build coherent, effective writing programs (p. 295).

The portfolio has seven specific features, according to Yancey (1996a).

1. It is a collection of work.

2. It is a selection of work, culled from the archival collection, usually supplemented by additional texts created specifically for the portfolio. Such texts include
reflective letters, annotations on individual texts, and other contextualizing texts such as a table of contents.

3. It includes reflection, which typically allows the portfolio composer to guide the reader through the portfolio and assist in its evaluation.

4. It presumes development, although texts demonstrating development aren’t always included in the portfolio.

5. It documents diversity-both in its contents, which are various, and in its ability to show how different our students are, one to the next-individually, cognitively, culturally, institutionally.

6. It is communicative in the sense that a portfolio always shares what is important to the portfolio’s composer, what is valued in the context in which that student works, and so on.

7. It is evaluative, as suggested before: The portfolio itself tells its observers what is valued by the participants who shaped it (p. 130).

Given these seven characteristics, the portfolio is a comprehensive formative assessment tool. The addition of student selection and reflection give the portfolio even more value, enabling learning beyond that brought about by completing the individual portfolio tasks (Abrami and Barrett, 2005). Requiring that students critically reflect on their work make the portfolio a learning experience for students as they evaluate the differences between what they create and what the learning objectives are, how their work has changed over time, and even the differences in their reflections on the same pieces of work. This student reflection is key in any formative assessment process. Gorlewski (2010) contends that “student self-
assessment has been referred to as the most overlooked, yet possibly most valuable, aspect of assessment for students at all levels” (p. 100). As the learner reflects on what his or her achievements are, what gaps there are in understanding, and goals for the future, “the assessment process is in itself a learning experience” (Wilcox, 1997, p. 36).

The portfolio, however, might best be used as an overview of student learning. It is limited in its ability to give teachers quick, immediate information about student understanding of and proficiency in one particular standard covered over the previous few days. Using both the portfolio and other types of formative assessments that offer quick information gives the teacher the most complete information for students. An advantage of online formative assessment programs (OFAPs) is their immediacy for assessment, scoring, and feedback. Since they provide reports for both proficiency and growth, they also allow students to reflect upon their own learning. The immediacy of online formative assessment is the feature that sets it apart from other forms, and this study sought to determine if teachers were taking advantage of this unique feature.

It should be noted that the term portfolio is also used to describe the collection of all the data for one student in a particular online formative assessment program (OFAP). However, since the student work and data in these OFAPs is limited to that completed in that program, they are not interchangeable with the portfolios described above.

**Formative Assessment and Adolescent Learning**

Because my study focused on the practices of teachers in middle school, it is important to note how learning theories support formative assessment for adolescents. Therefore, this section reviews the development theorists pertinent to FA in middle grades.
Of significance to formative assessment is Len Vygotsky’s (1978) theory of the ZPD, the area between what a child can do independently and what he/she can do with help from someone more knowledgeable, whether a peer, a tutor or a teacher. Vygotsky believed that the ZPD represented a truer measure of a child’s intellectual potential. The more scaffolding and support a child receives in the ZPD, the better that child will be able to perform in school. New instruction should challenge the student so that the new task is above the student’s independent level but within the ZPD. In this process, the instructor is more of a facilitator who continually challenges the student while guiding and helping him practice and learn (Chew, Jones & Turner, 2008). In addition, Vygotsky claimed that environment for students was everything – and that learning occurred when children and adults interacted. He called this process cognitive socialization (Beliavsky, 2006). In fact, Vygotsky claimed that learning takes place on dual planes, the intermental and intramental. "Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people [intermental] and then inside the child [intramental],” (Vygotsky, 1978, p. 57).

According to Yildirim (2008), “Vygotsky’s theory basically offers that if we want to understand learning and development, we have to focus on process instead of product” (p. 302). Vygotsky argued that “the only appropriate way of understanding and explaining ... forms of human mental functioning is by studying the process, and not the outcome of development” (Lantolf & Thorne, 2006, p.28.)

Vygotsky’s (1978) theoretical framework is an underlying one for formative assessment. The formative assessment model includes a pre-assessment to determine a
student’s understanding, followed by targeted instruction based on the pre-assessment. A post-assessment measures the distance between the first and second assessment in the student’s understanding. If the student still has not achieved the necessary understanding, specific, targeted instruction is again applied. This provides the scaffolding in the ZPD that Vygotsky claimed that students need in order to perform well. The formative assessment model further follows closely Vygotsky’s belief that “instruction and assessment should be inseparable” (Yildirim, 2008, p. 302).

The dynamic assessment model described by Yildirim (2008) as a derivative of Vygotsky’s ZPD theory is a formative assessment model. Lantolf and Thorne (2006) clearly explain the difference between summative and formative assessment, using the terms of non-dynamic and dynamic assessment:

What makes a procedure dynamic or not is whether or not mediation is incorporated into the assessment process. In other words, fill-in-the-blank, multiple-choice, open-ended essay, or even oral proficiency tests in themselves may or may not be dynamic. Their status is determined by the goal of the procedure and the format in which it is subsequently administered. In other words, there are no dynamic assessment instruments per se; there are only dynamic assessment procedures” (p. 331).

Otero (2006) contends that Vygotsky’s theory of concept development is the theory of choice for formative assessment. Vygotsky believed that the formal academic concepts a student learns in school can be connected to the understandings he has gained through experience. Further, as the learner becomes conscious of these experience-based concepts, he is able to transfer understanding across situations. Since the prior knowledge a student
brings to class, a combination of both academic concepts and experience-based concepts, is not always fully developed or completely correct, the teacher must explore what that prior knowledge is before making instructional decisions. Earlier, Dewey (1938/1977) also identified the critical role experience plays in learning in is concept of the continuity of experience. According to Otero (2006), “recognizing, describing, and using students’ prior knowledge in instruction is the formative assessment process” (p. 250).

In the sixties and seventies, John Carroll (1963), Michael Scriven (1979) and Benjamin Bloom (1971) all contributed to the theories and practice of formative assessment. Carroll is credited by some as creating the original true formative assessment model with his paper “Model for School Learning,” published in 1963. Carroll went against the prevalent thinking of the time that a child’s aptitude determined the amount he could learn by claiming that children differed in the time it took to master a concept, rather than in whether or not they could master it. His definition of aptitude mirrored that of Vygotsky, claiming that aptitude was established by the amount of time a child needed to learn the material (Guskey, 1980). Carroll’s idea was that the kind of instruction a student received was as important as that student’s innate ability to learn. Carroll argued that five factors determined how much a student learned. These included a student’s perseverance, the student’s “opportunity for learning,” or the time given in the classroom for learning, the student’s aptitude, the student’s ability to understand the instruction, and the quality of the instruction (Carroll, 1989). Too often, the focus in education has been limited to student aptitude and ability to understand the instruction and its quality, without consideration of the student’s perseverance and time allowed for learning.
In writing about evaluation, education evaluation theorist Michael Scriven (1979) coined the terms *summative* and *formative* to differentiate between evaluation that results in acceptance or rejection of the product and evaluation for the purpose of improving it.

Benjamin Bloom, influenced by both Carroll and Scriven, developed the Learning for Mastery model in 1968, later shortening it to Mastery Learning (Guskey, 2007). Bloom’s model was developed through two types of research: how successful pairs of tutors and their students worked together; and characteristics of teaching that made students academically successful. He saw that assessment as it was being used in schools did little more than to tell the teacher who could answer the questions correctly and who could not, which teachers usually already knew. In contrast, Bloom found in his study that assessment by tutors was frequent and included feedback to the students about how they were doing and providing activities to remediate any lack of understanding (Bloom, 1968).

Bloom applied these findings to his studies, which were based on Carroll’s five learning variables, and which he described in his 1968 article “Mastery Learning.” First, instruction was divided into small parts, after which students took an assessment called a diagnostic-progress test or formative evaluation test. No grades were given, and students who showed mastery on the formative test were given follow-up enrichment activities. Students who did not show mastery of the instructional content were offered several methods to reach full understanding, such as tutoring, rereading original materials, studying alternative texts, workbooks, and audio-visual materials. However, the most successful method was reviewing in groups of two or three. A second assessment was given to see if students had mastered the concepts. A comparison of this method with one in which students went through the
curriculum in the more conventional way showed significantly greater gains for students in the Mastery Learning program (Bloom, 1968): “Under these conditions, Bloom believed that virtually all students could learn well and truly master the subject material” (Guskey, 1980, p. 107).

Bloom and his students also worked over the summer of 1972 with the faculty of a Chicago junior college, Olive-Harvey, which previously had had dismal results with their low-achieving students, though many different programs had been tried. As a consequence of coursework Director of Research and Evaluation Emmett Jones did with Bloom, fifteen instructors began using Mastery Learning in their classrooms (Jones, Gordon, & Schechtman, 1975). Their informal research was focused on the changes instructors could make in the classroom to ensure that all students could achieve at the desired level, ignoring the conventional wisdom that intelligence, economic status, parental support, or any other aspects were limitations. The result was “significant improvement in student achievement, reduced rates of class attrition, and more positive attitudes toward learning on the part of both students and faculty” (Guskey, 1980, p. 107).

The Chicago City College System built upon these successes by developing a system-wide program for all eight colleges using Mastery Learning. Results were similar to those attained by Olive-Harvey (Guskey, 1980). Public school programs followed, in which teachers teamed up to develop the materials needed. Although this resulted in slow implementation, it also provided for program quality (Smith & Katims, 1977).

Several educational programs share aspects of the Mastery Learning model of pre-assessment, instruction, correctives or remediation, and post-assessment. These include
Understanding by Design and Response to Intervention (Guskey, 2010), as well as the Assessment for Learning models supported by Black and Wiliam (2010) and Stiggins (2005). The core philosophy behind Mastery Learning, formative assessment, Assessment for Learning, and other such models is that every student can not only learn but master the material that is taught under the “appropriate instructional conditions” (Guskey, 1980). While this theory is widely proclaimed by educators today, practice does not always follow.

Vygotsky’s influence is also seen in Howard Gardner’s (2011) theories. Gardner agreed with Vygotsky that intelligence was both a result of genetics and the number and characteristics of the instruction, or chances for learning, a student receives, (Gardner, 2011). Gardner also followed Vygotsky’s beliefs that the student learned with the help of other people, or the more knowledgeable other, or MKO (Vygotsky, 1978). The theory of Multiple Intelligences grew out of Gardner’s work with patients who had suffered some sort of brain damage. He found that while some abilities were seriously compromised or lost completely, others were unaffected. His further research convinced him that there is a collection of intelligences that are independent of one another and that the strengths and weaknesses of these intelligences differ for each human being, including musical-rhythmic, verbal-linguistic, visual-spatial, logical-mathematical, bodily-kinesthetic, interpersonal, intrapersonal, and naturalistic (Gardner, 2011). Thus educational pedagogy and assessment must acknowledge the “many different ways of acquiring and representing knowledge” (Beliavsky, 2006, p. 6). Although his use of the word “intelligence” to describe each of these competencies was controversial and his theory was in contrast to traditional theories of intelligence that focused on verbal and mathematical abilities, Gardner argued that there are
at least eight different kinds of intelligence, which he defined as “the capacity to solve problems or to fashion products that are valued in one or more cultural settings” (Gardner & Hatch, 1989, p. 5).

In Gardner’s view, teachers must give opportunities for students to learn through the different intelligences and assess them in ways that allow them to express what they know in a variety of ways. Since end-of-year summative assessment tests focus solely on verbal-linguistic and logical-mathematical, Gardner’s theory suggests that many students are left out of any true assessment of their learning. While teachers cannot design the summative assessments, in their classroom assessments they can make use of movement, music, groups, hands-on activities, and visuals such as paintings and videos to appeal to the strengths of all students. These assessments of course, would not be the more common paper/pencil tests, but would be what Gardner calls “intelligence-fair assessment” (Gardner and Hatch, 1989, p. 6). This type of assessment includes many opportunities to “stretch [students’] knowledge” and “many projects in which students are engaged for significant periods of time and which lead to genuine products,” which are much more helpful in increasing students’ understanding and preparing them for the world (Gardner, Siegel, & Shaughnessy, 1994, p. 564). Further, students who receive ongoing feedback for their work and chances to improve will soon “insist that all instruction pursue [the] goal” of real understanding (p. 564)

Gardner also was influenced by the ideas of John Dewey (1916), often called the Father of Progressive Education (Berube, 2000; Moula, Mohseni, Starrin, Scherp, & Puddephatt, 2010). Dewey came of age toward the end of the nineteenth century and the beginning of the twentieth, a time of great change in the United States, and his theories of
education were implemented to solve the problems of the “new age.” Education, to Dewey, was simply for the purpose of learning, enabling individuals to live “every stage of present development” and, that in “giving the child command of himself starting in the present, the future will take care of itself” (Sullivan, 1966, pp. 391-392). Dewey agreed with Vygotsky that social interaction is how students learn. In his book *Democracy and Education*, Dewey explained, "Since growth is the characteristic of life, education is all one with growing; it has no end beyond itself “(1916/1966, p. 53). Further, Dewey believed that instead of educators imparting knowledge to their students, learning was to be developed from within, as a result of teachers nurturing a love for learning that would last throughout students’ lives. People with normal abilities have “a vast capacity for learning” (Gibboney, 2006, p. 170). Dewey’s ideas fit well with the formative assessment model, in which one goal is for students to become responsible for their own learning and one tenet is that all students can learn with the appropriate time, scaffolding, and feedback.

According to Dewey, “it is the teacher's obligation to know each child intimately” (Sullivan, 1966, p. 393). Understanding the whole child - his “capacities, interests, habits, and instincts” - is essential in being able to foster the growth of knowledge, and by offering various activities based on the child’s abilities, interests, and needs, as well as his experiences, the student learns on his own in his own way (Sullivan, 1966, p. 392). Dewey’s ideas about the importance of experience also dovetail with those of Vygotsky. His theory of experience explains that a student’s learning is dictated by two principles: continuity and interaction. A student’s present and future is determined by the continuity of both his past and present experiences. It is the interaction of the student’s past experience with the current
situation that determines the present experience. Thus teachers must understand how subjective each student’s experience in the classroom is in order to teach effectively, as everything a student experiences is filtered by his past experiences (Dewey, 2007/1938). The only criteria for success is whether or not the student develops a lifelong love of learning; thus successful schools are simply those that meet this goal, a goal not measurable on any existing test (Gibboney, 2006, p. 170). Because the formative assessment model affords the teacher a vast amount of knowledge of a student’s learning, it allows the teacher to provide the kinds of experiences that will promote growth in knowledge and encourage a lifelong love of learning. In addition, the active role students play in formative assessment, in which they are partners with the teacher in setting goals, assessing their own work, and analyzing their own data, aligns well with Dewey’s theories.

Jean Piaget’s (1963) studies of children convinced him that their errors could give him valuable information about the way they learned. Early in his career while teaching at a school run by Alfred Binet, creator of the Binet Intelligence Test, Piaget helped mark papers for Binet. He began to see that the wrong answers students gave often provided more information about their cognition and cognitive development than correct answers did, and that children of the same age often gave similar types of wrong answers (“Jean Piaget: Champion of Children’s Ideas,” 2001, p. 43). This finding is fundamental to formative assessment. When teachers study both correct and incorrect answers students give to discern how students think, what they understand, and what understanding is missing, teachers gain information on which they can effectively act.
Piaget also claimed, contrary to prevailing early to mid-twentieth century theory, that the desire to learn in children is intrinsic, that “children think and learn because they are built that way” (Flavell, 1996, p. 200). When children encounter new information, according to Piaget (1963), they try to fit it in with their existing understanding, in a process called assimilation. If it does not fit, they experience disequilibrium. Since disequilibrium is not comfortable, the new material forces children to develop new schemas. It is this disequilibrium which makes learning happen, as children constantly adjust their thinking to assimilate and then accommodate new material, a process Piaget called equilibration. Piaget further ascertained that children who are given the opportunity to make mistakes often find and correct them, or even discover new ways to solve them, again allowing them to construct learning for themselves (Piaget, 1963). Allowing students the freedom to make and correct mistakes regularly themselves is also an important part of formative assessment, leading to student self-regulation. Thus Piaget’s theory of equilibration, resulting in the construction of knowledge fully supports the practice of formative assessment.

Abraham Maslow (1968), like Piaget, recognized stages in human development, a concept at the heart of formative assessment. With his needs hierarchy, that he called a “health and growth psychology,” Maslow postulated that humans must have certain needs met before they can move to a new stage in development, and that this is true for all people from all cultures (Goebel & Brown, 1981, p. 809).

Romig and Cleland (1972) explain that, according to Maslow, “the first task of the school, then, is to facilitate the child's attempts to have his basic needs met while teaching the child how to meet his own needs, growing away from an immature dependence on parents
and other adults” (p. 290). One of the basic tenets of formative assessment is that of student self-regulation. The specific, timely feedback a student receives, along with practice in goal-setting, moves students toward independence (Campos & O’Hearn, 2007; Black, et.al, 2004).

When the needs for food, water, shelter, oxygen, and safety/security are met, the needs of belongingness and self-esteem must be met to allow students to finally reach self-actualization. They can then learn, create, and fully realize their potential to meet this need, teachers must understand students’ abilities (Romig & Cleland, 1972). As with the models of Vygotsky, Bloom, and Piaget, the teacher first designs the curriculum to present to the student tasks that challenge, but are not too difficult and scaffolds tasks so that each is challenging but accessible with support. The goal is that as the student learns to take more responsibility for his learning, he is able to structure his own learning (Maslow, 1968). Formative assessment allows teachers to do just this, while enabling students to move toward self-actualization through regulation of their own learning, reflecting on their success and failures, and setting goals for the future. This concept is at the heart of formative assessment.

It is clear that the theories of these educational theorists provide strong support for the concept of formative assessment in one form or another, and in some cases, provide the very foundation for it. According to these men, some of them the best-known in educational theory, setting goals, performance assessment, teacher feedback, and student responsibility and reflection are important in students’ education, and each is intrinsic to one or more of these theories. This illustrates that the use of formative assessment in the classroom is one of the bedrocks of successful instruction. Thus, how teachers use online formative assessment in English Language Arts classrooms is an important addition to the literature.
Teaching Reading

Today there is still a widespread lack of understanding as to how to ensure students can read, though theory, methods and strategies have been identified and studied for fifty years or more. As a result of this confusion, non-readers and poor readers proliferate at the middle and high school levels. States and districts vacillate between programs, trying to find the one-size-fits-all program that will solve the problem from kindergarten through college. With the *No Child Left Behind* legislation, schools have increased their efforts and their expenditure of resources in this area. For example, for the last decade, teachers in North Carolina must take a certain number of professional development hours in teaching reading in order to keep their license current (Public Schools of North Carolina, 2014).

The most dominant theory for teaching reading through literature during the twentieth century was the New Criticism approach which prevailed in schools and colleges for most of the twentieth century. Under the New Criticism approach, literature was viewed almost scientifically. Characteristics of genres and the text structure were analyzed rather than the meaning of the text to the individual reader. Thus began the practice of teaching literature by genre, still in wide use today (Beach & Swiss, 2010).

Louise Rosenblatt brought Reader Response theory to the forefront with her ground-breaking book *Literature as Exploration* (1995). Reader Response, or transactional, theorists believe that the uniqueness of each reader makes the reading of the text unique for each, and that the reader *interacts* with the reading. The meaning of a text will differ for individual readers, depending on their beliefs, emotions, socio-cultural heritage, their prior knowledge, and many other factors. Further, the same reader reading a text for the second or third time
will not necessarily respond in the same manner as the first time, since the contributing factors in that reader will differ.

Research in reading takes two paths. One path, more traditional, leads researchers to question the skill set of reading – phonics, decoding, spelling, fluency, etc. A more contemporary path leads to an investigation of the reader, depending on a view of what the reader is doing while reading. For the first, the correlation of reading fluency, or the ability to read orally on grade level, and reading comprehension, the ability to understand what is being read, at one point are considered to be one and the same. Fluent readers are readers who comprehend what they read. However, studies such as that by Denton, Barth, Fletcher, Wexler, Vaughn, Cirino, Romain, and Francis (2011) have shown that the correlation between fluency and comprehension is not as clear as was supposed. This study found that while paragraph reading fluency is more highly correlated to reading comprehension than is reading word lists, reading fluency does not always indicate reading comprehension.

Applebee and Langer (1983) stressed that the best way to teach the cognitive skills and strategies necessary for reading comprehension is scaffolding, when the teacher offers heavy support at the beginning of instruction and withdraws it as students advance, echoing Vygotsky’s theory of the ZPD. It is also one of the tenets of formative assessment. Applebee and Langer (1983) advocated a process approach. Scaffolding occurs when teachers first set learning goals for students and give strong support and feedback, then guide them to become goal-setters for their own learning, needing less teacher support and feedback.

Scaffolding is also important for readers who do not form images in their minds while reading, as successful readers do (Wilhelm, 2007). Helping students construct these images,
using such activities as improvisation, prediction, drawing, and puppetry, enables students to learn that reading can evoke images and to begin forming them on their own.

Understanding that reading comprehension is a transaction between reader and text has brought about change in the types of texts being used in the classroom. The English Coalition Conference and the Pacemaker Project grew out of concern on the part of the National Council of Teachers of English and other organizations charged with the teaching of English. These entities helped bring about the move away from the use of previous “canon” pieces of literature and toward the use of literature that speaks to specific subcultures and experiences and allows students to explore significant issues. Moreover, both groups stressed that assessment was to be an ongoing, integral part of both teaching and learning (Applebee, 1996). Both were based on a pedagogy of constructivism. Skills and knowledge would develop as an outgrowth of their engagement in the literature, rather than by teaching these skills directly and independently through literature chosen primarily for that purpose.

Garbarino (1982) emphasizes the importance of identity formation in adolescence, and explains that adolescents make judgments about themselves based on comparisons they make with others and their perceptions of how others see them. Thus texts that show others overcoming familiar problems help adolescents form positive views of themselves. Partially as a result of The English Coalition Conference and the Pacemaker Project and the mounting research about adolescents and identity, self-selected texts and young adult novels with themes that confront issues that adolescents deal with in their lived lives have become the backbone of many English Language Arts classrooms.
This approach has also been influenced by the work of Frank Smith (1994), who postulated that reading comes naturally to students, and that the most important thing teachers can do is to provide accessible reading material of interest to students. Teaching reading strategies is interruptive to the reading process, according to Smith, and does more harm than good (Atwell & Mayher, 2006). Using texts that deal with adolescent issues as well as self-selected texts was further encouraged by Nancie Atwell’s book In the Middle (1987), in which she described her English Language Arts classroom as one in which students spent their time reading self-selected novels. She credited Frank Smith with developing the reading and writing workshop model (Atwell & Mayher, 2006). Atwell (1998) used mini-lessons during each class period and individual, interactive journal letters to increase skill, motivation, and comprehension. She reported that students not only gained dramatically in reading skill, but that the depth of their understanding, their internalized strategies, and their interest in reading also increased.

Harvey Daniels’ (2002) description of the literature circle and its use in his classrooms also increased interest in young adult novels. Daniels created certain roles that students adopted and rotated within their groups to ensure good discussion and group understanding. In this way, students could select a novel of their choice from a small list provided by the teacher, grouped by theme or setting or some other commonality. Vygotsky’s (1978) theory of the Zone of Proximal Development states that learning based on language and communication is a social activity and that students learn when there are others who are more advanced to model and support their learning at first. Learning, according to Vygotsky, took place on two planes, the intermental and the intramental. That is, learning is
an experience that occurs not only within an individual, but also as a result of interaction with others (Wertsch & Tulviste, 1992). In Daniels’ (2002) model, students who are more advanced are able to support the understanding of their less-advanced peers.

Literature circles also reflect the theories of Howard Gardner (2011), whose theory that there are at least eight different intelligences, contradicts traditionally held notions that verbal and mathematics ability indicated one’s intelligence quotient. Gardner advocates using learning groups made up of students with varying abilities in the identified intelligences. In this way, stronger students in each intelligence both model and support the others in the group, resulting in growth for all group members in several areas (Beliavsky, 2006). Daniels (2002) took this further to develop literature circles, where students take on specific roles, including artist (visual-spatial), director (interpersonal), word wizard (linguistic), literary luminary (linguistic), and puzzle master (mathematics), to help one another view a text from various viewpoints. Roles can also be invented to meet specific needs of students or of the instructional unit.

Teachers of adolescents need to tap into what adolescents already do well and what adolescents find interesting. According to the NCTE Guideline (2004):

Adolescents are already reading in multiple ways, using literacy as a social and political endeavor in which they engage to make meaning and act upon their worlds. Their texts range from clothing logos to music to specialty magazines to Web sites to popular and classical literature (p. 2).

Further, teachers need reading material that will “tap students’ diverse interests and represent a range of difficulty” and they need the kinds of training and support that will help
them meet adolescents’ needs (NCTE Guideline, 2004, p. 3). Then, they themselves will be able to:

- [Create a] bridge between adolescents’ rich literate backgrounds and school literacy
- Teach literacy in their disciplines as an essential way of learning in their disciplines
- Recognize when students are not making meaning with text and provide appropriate, strategic assistance to read course content effectively
- Facilitate student-initiated conversations regarding texts that are authentic and relevant to real life experiences.
- Create environments that allow students to engage in critical examinations of texts as they dissect, deconstruct, and re-construct in an effort to engage in meaning making and comprehension processes (NCTE Guideline, 2004, p.3).

Though much research has been done on the process and teaching of reading, ensuring adolescents become proficient readers is still a challenge. Through the work of practitioners and researchers such as Rosenblatt (1995), Applebee (1996), Smith (1994), and Daniels (2002), reading education has evolved over the last fifty years from the presentation and analysis of the classic canon to meeting specific needs of adolescents through the recognition of individual reader-text interaction, student scaffolding, self-selected texts, and peer support. Though much research has been done on the process and teaching of reading, ensuring adolescents become proficient readers is still a challenge. Through the work of practitioners and researchers such as Rosenblatt (1995), Applebee (1996), Smith (1994), and Daniels (2002), reading education has evolved over the last fifty years from the presentation
and analysis of the classic canon to meeting specific needs of adolescents through the recognition of individual reader-text interaction, student scaffolding, self-selected texts, and peer support.

According to proponents of online formative assessment, not only can online formative assessment replicate these new models of instruction, but it can do so efficiently and effectively, relieving teachers of much of the time-consuming preparation these models require and giving them more time to work directly with students.

**Teacher Planning, Decision-Making and Judgment**

In this study of teachers’ use of an online formative assessment system, teachers will be observed as they plan instruction to remediate specific gaps in student understanding. Teacher planning and decision making are integral pieces of the instructional protocol, and according to Clark and Lampert (1986), “Teacher planning is a major determinant of what is taught in schools. The curriculum as published is transformed and adapted in the planning process by additions, deletions, interpretations, and by teacher decisions about pace, sequence, and emphasis” (p. 28).

While experts agree that teacher planning is an essential part of the teaching process, how teachers plan and what guides the decisions they make during planning are areas that are still being researched. Until 1970, research focused on observable behaviors of students and teachers (Clark and Yinger, 1977). But “research on teacher thinking, through its descriptions of the functions, models, and sheer amount of teacher planning, provides us with a set of insights and the beginnings of a language in which to think about and explain how teaching may at once look simple yet feel and be very complex and demanding” (Clark &
Lampert, 1986, p. 28). Part of the difficulty of understanding planning is that it is a “solitary activity” and is often conducted informally (Clark & Lampert, 1986, p. 28). Glatthorn (1993) found that teachers do a great deal of mental planning. This mental planning goes on over the summer and while school is in session: they mull over student needs, objectives, materials, and likely problems, developing mental scenarios and alternatives for action. They consider . . . plans as a guide for making crucial decisions. (p. 2)

An analysis of teacher planning using a “think-aloud” method by Peterson, Marx, and Clark (1978) found that teachers made more statements about the instructional process and “Lower-Order Subject Matter,” or facts than about goals and objectives or “Higher-Order Subject Matter,” including such things as concepts (p. 424). Teachers also plan continually.

In Jackson’s groundbreaking book *Life in Classrooms* (1968), he concluded that teachers used preactive, interactive, and post-active decision making, “that is, occurring before, during and after teaching (Westerman, 1991, p.292). Teachers must plan on many different levels too, including yearly, term, unit, weekly and daily planning (Glatthorn, 1993).

Clark and Yinger (1977) describe a five-month case study Yinger conducted in order to develop planning process models. The case study followed a first and second grade teacher as she planned and illustrated that:

The teacher’s goal conceptions, her knowledge and experience, her notion of the planning dilemma, and the materials available for planning interact to produce an initial problem conception worthy of further exploration. The second stage in the planning process was problem formulation and solution. The mechanism proposed for
carrying out this process was the "design cycle." In this cycle, problem solving was characterized as a design process involving progressive elaboration of plans over time. Elaboration, investigation, and adaptation were proposed as phases through which plans were formulated. The third stage of the planning model involved implementation of the plan, its evaluation, and its eventual routinization. This stage emphasized the contribution of evaluation and routinization to the teacher's repertoire of knowledge and experience. (Clark and Yinger, 1977, pp. 283-285)

Planning served three purposes for teachers in a study conducted by Clark and Yinger (1979). It helped them mentally prepare for teaching, psychologically prepare for teaching, and helped them organize the process of teaching and learning. Clark and Lampert (1986) identified three areas of knowledge necessary for a teacher to plan and teach effectively. The first of these is contextual knowledge. When planning, teachers must take into account student behaviors, abilities and disabilities, class size, and availability of materials and tools” in addition to the activities, school mandates, standards, that are necessary to instruction (Clark and Yinger, 1977).

Contextual knowledge allows a teacher to consider the classroom climate, the individual students and their behaviors, abilities, and motivation, and other circumstances that may be germane. The extreme complexity of teaching requires that teachers have and use a great deal of contextual knowledge. Doyle (1977) explains that teachers must deal with the multidimensional, in that there were a variety of purposes, events and processes in each classroom; simultaneity, as the teacher had to be aware of many things at a time, such as pacing, fairness to all students, relevance and logic of answers to questions, as well as
behavior, work involvement, and outside distractions; and unpredictability, including student reactions and interactions, interruptions, and equipment malfunctions. “What one decides to do today has a great deal to do with what happened yesterday and what effects such a decision will have tomorrow, next week, and next month” (Clark and Lampert, 1986, p. 29).

The second area of knowledge identified by Clark & Lampert (1986) is interactive knowledge, which allows teachers to know how to present instruction to achieve their goals in a way that is not only positive for the student “but also convince[s] the student to see it that way as well” (p. 29). A teacher is faced with “interrelated and competing decision situations” during both planning and teaching, and “there are not perfect optimal situations to these decisions” (Clark & Lampert, 1986, p.28). Teachers must weigh one gain or loss against another, while juggling all the other responsibilities of instructions, thus the teacher is constantly experimenting. Teachers therefore learn practical ways to conduct their classrooms in such a way that learning can take place (Clark & Lampert, 1986).

Third, Clark and Lampert (1986) identify speculative knowledge as necessary because of the myriad of decisions teachers must make both before and during instruction. These include long-range decisions, such as how to present the curriculum, which materials to use, the assignments to give, and how to assess student learning. Even more difficult are the decisions that must be made during class, including how to talk to an unruly student, how to redirect student attention after an interruption, what should be done about those who didn't do the homework and those who clearly do not understand it, and whether more questioning will elicit better understanding. There are so many things out of teachers’ control, whether it
be what happened to a student on the bus or an unexpected fire drill, teachers are like juggling circus performers balancing on a rolling barrel.

Downey (2001) contends that instructional planning should consist of three parts: learning objectives, assessment and analysis, and instruction strategy selection, based on Bloom’s (1971) model of mastery learning. However, Downey points out that while this model for planning and instruction is still the most widely researched, it is not commonly used.

Further, although most preservice teachers learn that objectives should be identified first in planning, and, in spite of the increased emphasis on state and national standards, teachers do not tend to make standards or objectives their first priority in planning (Clark & Yinger, 1977). Glatthorn (1993) noted that the teachers he studied were first concerned with their students’ needs, abilities and interests. Next in importance was the task, including content, activities, materials and classroom context, and finally, objectives. Similar results were found in the study by Peterson, et al. (1978), who reported that that teachers spent most of their planning time concerned with the content or subject matter. A much smaller proportion of time was devoted to strategies and activities, and the smallest dealt with objectives.

Clark & Yinger (1977) report that multiple studies show that student motivation and involvement were the first considerations in teacher planning, followed by cognitive outcomes, the difficulty of subject matter, and affective outcomes. However, it is interesting to consider whether the small amount of time devoted to objectives is because objectives are predetermined and mandated by the state or district, and/or because these have become
intrinsic with their content and pedagogy for experienced teachers. Berliner (1986) posits that successful teachers become expert in blending objectives with other aspects of planning, almost unconsciously. Instructional standards and/or objectives data are routinely reported by OFAPs, thus bringing focus back to these essentials during analysis and planning. Analyzing formative assessment data, teachers are immediately brought back to the learning objectives.

Clark and Yinger’s (1977) article includes two other important aspects of teacher thinking that affect their teaching as well as their planning: judgment and decision-making. Shavelson, Cadwell, and Izu (1977) found that teachers paid close attention to the reliability of information they received about students and were able to revise their judgments as needed when additional information was available. Interestingly, a study of the general population found that the opposite was true – people were not influenced by the reliability of the information, and they also did not usually revise their judgments, even when presented with more information (Tversky & Kahneman, 1974). However, Clark and Yinger point to other studies, such as that by Peterson, et al. (1978), that conclude that teachers were not very successful in predicting the specific achievement levels of their students, and that often these judgments were based on what teachers described as “student participation,” though they were clear about what specific behaviors this included. Given the absence of a clear conclusion, it seems even more essential that teacher receive accurate, ongoing formative information about their students so that they can make the kinds of decisions necessary for student success.
Concerns for Teachers

Based on the study by Jones and Moreland (2005), in which teachers reported changes in planning, questioning, and knowledge levels when they implemented formative assessment practices, there is certainly a need for various kinds of support for teachers embarking on the formative assessment journey. If formative assessment is to succeed, teachers must become experts in it. According to the findings of Black and Wiliam (1998) in their review of 681 studies, teachers are the most significant factor in implementing successful formative assessment. This finding is borne out by the previously mentioned study by Furtak, et al. (2008). In this study, the researchers concluded that the most significant variable for student achievement was the teacher. How teachers used formative assessment was critical to its success, just as how they approached the curriculum and taught it made similar differences.

In her very clear and thorough description of what formative assessment should look like, Heritage (2007) maintains that to use formative assessment successfully in the classroom, teachers need specific knowledge and skills. Four basic elements of teacher knowledge are critical: 1) domain knowledge, 2) pedagogical content knowledge, 3) knowledge of students' previous learning, and 4) knowledge of assessment. (p. 142)

Similar to these four basic elements are the four steps McManus (2008) identifies that are critical to the process of formative assessment and dependent upon teacher understanding and skill:

Step 1. Identification of learning targets and criteria for success.
Step 2. Elicitation of Evidence of Learning.


Step 4. Implementation of Action(s) to Close the Gap (p. 81).

McManus (2008) states that a major obstacle to widespread use of formative assessment is lack of teacher training. Her sixteen-week study of mathematics teachers at two high schools included four “modules” of professional training and observations of subsequent use of formative assessment in the classroom. Of the original sixteen teachers, four completed the study. Through her grounded theory approach, she went on to identify three elements necessary for implementing formative assessment. McManus calls these elements “inter- and intra-related” to the four formative assessment process steps previously identified in this review:

Element 1. Teachers must have a high level of content and pedagogical content knowledge.

Element 2. The classroom environment must be an open, trusting environment where students are partners in learning process.

Element 3. Discourse about [curriculum] content must be dialogic. (pp. 82-83)

Though her study focused on high school mathematics classrooms, these elements seem to be applicable to any effective use of formative assessment in the classroom, and their necessity to effective formative assessment are borne out by many of the studies reviewed here.

Furtak, et al. (2008) noted that in the past, teachers were often surprised and disappointed with student performance on assessments. Upon reflection, the teachers realized that they were not clear at the beginning of the unit as to what they really wanted and
expected students to know. That many teachers are not clear about the learning outcomes they expect from students has been documented in other studies as well, along with the finding that clarity of purpose is an essential component of formative assessment, and works best when it is shared with the students (Black & Wiliam, 1998; Heritage, 2007; McManus, 2008). McManus (2008) further cites Sadler’s (1989) claims that communicating specific, well-defined learning goals to students is necessary so that they understand their current status, their short-term goals, and the method for achieving them (pp. 23-24).

To use formative assessment correctly, teachers also need to know how to look at the data generated by the assessments and to be able to see what it means in terms of student understanding. In a study of twenty-five third and fifth grade mathematics teachers who were to use district-wide interim assessments formatively, Oláh, Lawrence and Riggan (2010) found that the district gave teachers a data analysis protocol with the following guidelines:

- designed to help teachers identify weak points in their students’ performance; and articulate strategies for regrouping, reteaching, and reassessment
- In addition, it asks teachers to reflect on how they can better differentiate their instruction to meet the needs of all students. . . . [However,] there seemed to be little guidance for teachers about how to act upon their analyses of interim assessment data (p. 239).

In this Ohio district, teachers were expected to use five days following the interim assessment for necessary reteaching. Yet individual teachers made the decisions as to how to use this time. Most concentrated on those items that most students missed and the students who performed at the lowest level. Students performing at high levels often were used as
peer tutors or were given extra independent work (Oláh et al., 2010). The researchers found that teachers used the data to find students and content in the following way:

[To] focus during reteaching week, and that they did an excellent job of translating the standards into content. However, for the most part, teachers were unable to learn very much from the responses students made to the items as they “did not use the interim assessments to make sense of students’ conceptual understanding of the content, nor were [the assessments] helpful for diagnosing errors in anything beyond a procedural way” (Oláh et al., 2010, p. 244).

The researchers concluded that “limitations in [teachers’] analyses of [the] data ultimately led to a relatively superficial approach to instructional planning and response” (Oláh et al., 2010, p. 244).

Equally important for teachers is understanding how to incorporate feedback to students. A study by Margaret Brown of King’s College, London, reported by the National Research Council (1999), found that “more than half of the teachers studied emphasized summative assessment over formative assessment; indeed, many considered formative assessment an unnecessary addition to their workload. These teachers taught according to the prescribed curriculum and did not adjust their teaching based on assessment-related feedback” (p. 17). Further, “the top teachers in the study, as gauged by student gains, were focused on assessment all the time” (National Research Council, 1999, p. 17).

Another change that may be difficult for teachers to make is that differences between summative practice for both individual teachers and the school as a whole when understood
as part of a process, formative assessments do not require grading, as the learning is still ongoing. Using formative assessment correctly means that students and teachers become partners in reflecting on and monitoring learning, and students learn to self-regulate their own learning (Heritage, 2007). Grading can then work against, rather than for, the process.

Participants in the study conducted by Black, et al. (2004) critically examined the purpose and process of grading work. While grading is appropriate as a summative measure, many assignments should be ungraded, giving students opportunities to check their own understandings and correct mistakes, while giving teachers information on what still needs to be done to fill instructional gaps. Other assignments might be both summative and formative, with grades assigned, but with generous feedback given to the student on how to increase his performance in the future. This is already common practice in the teaching of writing.

In another example, Vaden-Goad (2009) reports on his use of summative quizzes and tests in formative ways. Vaden-Goad used the summative tests and quizzes his university students took as formative assessments. Students who improved on their performance in later tests were able to use the better scores on the same material for their final score.

Finally, although the purpose of formative assessment is improvement in student learning, effective use of it often requires changes in belief as well as skill. Black and Wiliam (1998) contend that there are two major concerns affecting teachers who may want to use formative assessment. Both have to do with teacher beliefs. The first is what the teacher believes about learning. If the teacher believes that learning is transferred to the student by the teacher, with understanding following when the student is mature enough, formative
assessment will not be successful in her class. The second concern is what the teacher believes about students and their abilities. If the teacher believes that learning capacity and/or ability is set, then formative assessment in her class will be meaningless, and can lead to issues of fairness, bias, and stereotype. These concerns go to the very pedagogy and belief systems of teachers and (Black & Wiliam, 1998).

**Online Assessment**

Formative assessment, used correctly, can be a daunting task for already-overworked teachers. The time necessary for effective planning, construction, organization, administration, scoring, and analysis of useful formative assessments, added to that needed for feedback and modified, individualized instruction is enough to make many teachers hesitant.

According to Brown, Hinze and Pelligrino (2008),

In diagnostic assessments of individual learning . . . significant amounts of information must be collected, interpreted, and reported. No individual, whether a classroom teacher or other user of assessment data, could realistically be expected to handle the information flow, analysis demands, and decision-making burdens involved without technological support. (p. 253)

The unique characteristics of the Internet offer the opportunity to quickly assess and then immediately score and disaggregate data. These characteristics allow teachers to concentrate on teaching, analysis of student misunderstanding, feedback, and remediation. In their article in *Journal of Research on Computing in Education*, Bonham, Beichner, Titus, and Martin (2000) point out that the Internet’s flexibility allows for assessing the same
concept in multiple ways, providing simulations to which students can respond, accepting submissions of essays and graphics, and for the use of “nonexclusive” multiple choice items with a bank of possible choices, as well as the more traditional “exclusive” multiple choice items with a narrow set of answer choices (p. 36). Bonham, et al. (2000) give eighteen examples of web-based assessment programs, a sample of the programs in existence at that time.

According to the National Research Council (1999), Technology can help with the design, administration, and scoring of performance assessment and can help reduce the burden on teachers when more formative assessment is added to the instructional process. O’Neil, for example, illustrated that computer software can facilitate administration, scoring, and reporting of concept map assessments, can reduce scoring costs, and can provide helpful feedback on student proficiency in specific cognitive areas (p. 25).

Online testing also has the advantage of becoming an integral part of the curriculum. Since online assessments can be accessed by individual students at various times, be scored automatically, and the data collected and disaggregated for the teacher immediately and easily, they offer more flexibility in the classroom. “The goal is to integrate assessment into daily instruction to support, rather than just measure, student learning” (Olson, 2006). Hunt and Pelligrino (2002) explain that “disruptive testing, sometimes derisively referred to as ‘drop in from the sky’ testing, has traditionally been used for certifying student
accomplishment and predicting future performance. This is the sort of testing that is at the focus of debates about accountability” (p. 73).

One of the most important characteristics of formative assessment is the necessity of timely feedback to students. According to Hunt and Pellegrino (2002), “There is no point in formative assessment by a teacher if the teacher cannot identify, analyze, and respond to the problems of individual students” (p.75). As the demands of ever-expanding curriculum, multiplying diversity of student abilities, cultural backgrounds and language, and increasing expectations of administration, identification, analysis and response to the information gathered through formative assessment is often the stumbling block (Hunt & Pelligrino, 2002).

Online testing can help teachers overcome this challenge, as it often allows for quick reporting of assessment data delivered digitally, thus allowing the timely feedback that is so crucial for formative assessment, and providing various kinds of reports so teachers and administrators can view data in different ways (Shirley & Irving, 2015).

In addition, online testing can limit the time used for testing. Feng, Heffernan, & Koedinger (2009) point to the fact that some critics of the No Child Left Behind legislation already call it the “No Child Left Untested” legislation, and “that every hour spent assessing students is an hour lost from instruction” (p. 243). Since teachers do not need to administer them, the assessments can be given to small groups of students while the teacher meets with another group, or individual students can take an assessment while the rest of the class is
otherwise engaged. Also, because these assessments are designed to be given frequently over small “bites” of material, they are made up of just a few questions.

Some controversy exists over online assessment because of the necessity of using multiple choice questions for standardization and immediate evaluation. Some researchers insist that multiple choice questions encourage surface learning. However, other research has found little or no difference in the depth of learning between constructed response (written responses) and multiple choice questions. In fact, according to a study by Scouller and Prosser (1994), the type of learning a student generally employs, surface or deep, is the primary factor in the determination of the effectiveness of a test. Students who “surface learn” do not improve on multiple choice tests; nor do students who use “deep learning” do more poorly on multiple choice tests than any other.

Further, Kniveton (1996) states that “the performance of one student on two essays in the same examination is no more consistent than a student’s performance on an essay and a multiple-choice test taken on different occasions” (p. 81). Bull and Stephens (1999) report that Question Mark software used at the University of Luton has been successful with both students and professors, providing students more frequent assessments, and giving professors quick information about a student’s current understanding of instructional goals without requiring the investment of time more conventional assessments would require. The same software was used successfully for formative assessments, giving students opportunities for reflection on their learning through immediate feedback (Bull & Stephens, 1999).

While no one would suggest that multiple choice online tests be the sole tool for assessment, whether summative or formative, use of quick, prepackaged quality assessments
can be a boon to both teachers and students. Black, et al. (2004) posit that it is unrealistic to expect teachers and students to completely separate summative and formative assessments. It is how summative tests are used that makes them formative (Black, et al., 2004; Chappuis & Chappuis, 2008). Almost any assessment can be made formative if the teacher uses the data from the assessment to check on student understanding of the concept taught and then remediates as needed, instead of simply marking the assessment and moving on to the next unit of study. Several researchers and practitioners have examined and/or used computer-based summative assessments in formative ways. (Dobson, 2011; Wang, 2007; Hunt & Pelligrino, 2002).

However, teachers and administrators must be clear about the function of formative and summative tests. Many commercial companies are simply relabeling their summative tests as formative, and school districts are administering interim or benchmark assessments with the idea that they be used for formative purposes. Chappuis and Chappuis (2008) explain:

In reality, this level of testing is often little more than a series of mini-summative tests, not always tightly aligned to what is taught in the classroom. There is nothing inherently formative in such tests—they may or may not be used to make changes in teaching that will lead to greater student learning (pp. 15-16).

In order for an assessment to be formative, teachers must be able to access the data during instruction of the material assessed and then must use that data to inform instruction. Often the commercial and interim assessments do not provide this immediacy, and thus, by the time the data are available, students have moved on to another group of instructional
goals. Moreover, formative assessments must assess *only* the material that is being currently taught. Many so-called formative and interim assessments cover material taught over a grading period or semester, thus negating any possibility of timely feedback and modified instruction (Chappuis & Chappuis, 2008).

Increased interest in assessment over the ensuing ten years, due in large part to the *No Child Left Behind* legislation, has resulted in the availability of many more programs for online assessment, and more are in the process of development. In addition to companies which focus only on assessment, such as Measurement, Inc. and the non-profit Education Testing Service, some of the largest educational publishing companies have developed online assessment programs to supplement their other products (Olson, 2004). Pearson Education, Harcourt-Brace, Renaissance Learning and Vantage Learning are just a few of the well-known companies jumping on the formative assessment bandwagon (Olson, 2004; Doe, 2005). According to Tom Wiley, a senior analyst for a Boston market-research firm, "Assessment providers are looking to get more value out of the vast amounts of assessment data that they have at their fingertips now" (Olson, 2004, p. 7). In fact, Eduventures, Wiley’s firm, predicted that the sales from formative assessments would bring in $323 million for these vendors during 2004 alone, and that this revenue would continue annually (Olson, 2004, p.7). These companies do provide their own research to prove the effectiveness and reliability and “an aspect of validity” of their programs; however, it is more difficult to find unbiased research (Brown, et al., 2008, p. 251). Evidence of studies on frequency of assessment do exist and are included in the article by Brown, et al. (2008), but the studies are
two decades old, and the authors make it clear that further research needs to be conducted on the efficacy and use of online formative assessments.

This lack of unbiased research makes assessing how effective online formative assessments difficult. Until recently, what research has been done has generally occurred in professional and university settings. A search of the literature on the subject brings up many studies of online assessment in psychology (Buchanan, 2002) and higher education, especially in medicine and teacher education (Gikandi, Morrow & Davis, 2011; Kibble, 2007). Of these studies, most are concerned with summative assessments, according to Gikandi et al. (2011), who “applied qualitative thematic criteria in selecting and reviewing the available literature from which they focused on identifying and analyzing the core themes that are central to the concept of formative assessment” (p.2333). In their review of eighteen studies, these researchers included all types of formative assessment, including quizzes, portfolios and projects, most of which were collaborative. These tended to be generated by individual professors or departments for both online and on-site classes. Nine of the studies included only teacher education courses, and seven more include teacher education courses as one of the two or three disciplines studied. The authors of the review concluded that the “online formative assessment has a potential to engage both teachers and learners in meaningful educational experiences,” but that the “paucity of studies and the fact that most were found in teacher education also indicates that such inquiry is likely to require improved professional development for faculty in other disciplines” (Gikandi, et al., 2011, p. 2347).

The teachers in this study used a variety of online programs in formative ways. A description of these programs, along with their intended purposes, follows:
Case 21 – A product of TE21, providing benchmarks only. Benchmarks, also called interim assessments, are usually given by the school or district at specified times during the school year. Results can be aggregated, unlike formative assessments, so these assessments are often used to track progress specific goals or predict performance on summative assessments (Perie, et al., 2007). According to their website, “Teachers want and need immediate feedback on student skills to ensure that they are ready for testing” (TE21, 2015, “Case Assessments”). The assessments must be sent to TE21 to be scored, but are returned to the schools within 48 hours of receiving them. Schools or districts must pay for the program.

ClassScape – Provided free by the North Carolina Department of Public Instruction (NCDPI) as a part of the teacher and student information system and developed through a partnership between the NCDPI and North Carolina State University for classroom formative assessment... It later added benchmark capability. When North Carolina hired Pearson to provide its state assessments, ClassScape was cloned and inserted into the Pearson shell program called Schoolnet, as a part of the products it provided to North Carolina. Both ClassScape and Schoolnet are available to teachers. Teachers can use prepackaged quizzes of ten questions assessing a single standard or create custom quizzes assessing one or more. Assessment results are immediately available to teachers and administrators. Some aspects of it, such as the way the reports appear, differ from Schoolnet.

Mastery Connect – Allows teachers to share their own assessments with other teachers and administrator. Based on mastery learning and formative assessment concepts, There is a free version and two paid-subscription ones.
Quia – Offers teachers the ability to create assessments, games and other activities, inputting their own selections and questions as well as accessing content other teachers have created. Reports can be seen by whole class or individual students. Teachers buy a subscription.

Study Island - Offers assessments based on individual state standards. Its website describes it as an “instructional and diagnostic tool” to prepare students for their state’s summative tests. It offers program customized to individual state standards. It rewards students with ten-second games to keep them engaged, and offers reports. Study Island can be used on many platforms, including cell phones.

Schoolnet – the formative and benchmarking tool provided free to schools through NCDPI and Pearson’s PowerSchool, a student information system which includes a parental portal, grade book, attendance reporting system, and other tools for teachers and administrators. Schoolnet contains the same assessment items as ClassScape. As a part of the 2013 state contract with Pearson Education, Schoolnet was to replace ClassScape, but many schools preferred the ClassScape format and reporting, so both are available.

While research shows that using online formative assessments has significant advantages for both students and teachers, the combined lack of studies of this tool in all subject areas as well as in the K-12 arena underscores the importance of my research. States, districts, schools, and teachers are using online formative assessment, yet how they are using it and whether its advantages truly translate to the English Language Arts classroom is unknown.
Summary

From this literature review, it is clear that both teachers and students benefit from formative assessment, that it is an important tool for use in reading, and is supported by learning theory. It is also clear that formative assessment is underused for a variety of reasons, including lack of teacher time for all of the components of formative assessment as well as the lack of professional development in formative assessment. Online systems have proven useful in helping teachers in overcoming these obstacles to frequent, ongoing formative assessment so necessary to student educational growth. According to proponents, these systems can save time, give individualized specific feedback by objective, track student progress by objective and over time, and organize data for easy reference. Theoretically, using an online formative assessment system can free teachers to spend more time on instruction. In addition, teacher planning can benefit from the specification formative assessment provides in terms of objectives match to instruction (Bulkley et al., 2010).

However, missing from the literature is the process English Language Arts teachers in reading use to bring all of their training, knowledge and expertise together to construct meaning from the data gathered through formative assessment using an online evaluation tool. This study seeks to explore that process and to fill a gap in the literature about how middle school language arts teachers modify their practice as a result of analyzing formative assessment data.

Chapter 3 outlines the methodology which will be used to conduct the study, data collection and analysis, as well as the study’s theoretical framework and issues of validity, reliability, subjectivity, and ethics.
CHAPTER 3: Methodology

General Introduction

This chapter describes the methodology that was used for the research study to determine the impact of online formative assessment in reading on teachers’ planning and instruction middle school English Language Arts classroom. It examines the appropriateness of the methodology, the research question, the theoretical framework, site and sample selection, data collection, and analysis procedures. The chapter also includes the justification of a qualitative approach, as well as the validity and reliability of the study. A discussion of the author’s subjectivity as it correlates to this study, ethical issues, and the study’s limitations round out this chapter.

Background

Although teachers use formative assessment to check student understanding every day in classrooms, most of this assessment is informal. Teachers often, for example, ask questions orally, use bell ringers, entrance or exit tickets, homework and all kinds of class work to check understanding. However, the lack of formality and standardization in these strategies may make it difficult for teachers to align what they learn through these assessments with learning objectives, and to identify and track common strengths and weaknesses. In fact, Bentz and Fuchs (1993) found that teachers in their study were not very accurate in their judgment of actual student proficiency.

Another problem for teachers is that they lack the time to develop and score what Stiggins (2005) calls “short-cycle” assessments, those that give immediate feedback to both teachers and students and are vital to formative assessment. Results for formative
assessments must be “available quickly enough to enable teachers to adjust how they're teaching and students to alter how they're trying to learn” (Popham, 2006, p. 8). Without the time to create and score these assessments, teachers may begin to administer fewer assessments over larger units of instruction, making remediation more difficult. For these reasons, a tool which provides pre-made questions or assessments and then automatically scores them can support teachers’ integration of formative assessment in the classroom, thus making within-instruction assessment and timely feedback possible.

Because of the onslaught of new technologies, online formative assessment has become an option for teachers to access not only pre-packaged assessments but data that are immediately available. Computers have ushered in new ways for students to immediately access a plethora of information from thousands of sources. In this era of quick information access, data-driven instruction, and increased demand for testing modifications for special needs students, assessment via online tools has become the new norm, even for state-administered end-of-year summative tests (US Department of Education, 2009).

**Choice of Methodology**

For this study, to investigate the relationship between the use of online formative assessment in reading and English Language Arts (ELA) teachers’ planning and instruction, I used a qualitative multiple-case approach. In designing this study, I chose to look for answers to a question rather than to first formulate a hypothesis that would be proved or disproved. Qualitative studies use this open-ended approach to research. For example, Miles and Huberman (1984) describe qualitative research as “essentially an investigative process” (p. 37). Through examining the planning and instruction of teachers, I wanted to
understand how four teachers interpreted and used data to inform their practice, and as Merriam (2009) explains in her description of qualitative research, “to achieve an understanding of how [teachers] make sense out of their lives, delineate the process of meaning-making, and describe how people interpret what they experience” (p. 14, emphasis in original). I was particularly interested in the thinking of teachers as they analyzed the data and the pedagogical choices they made based on that analysis. Therefore, I planned to talk to teachers about their thinking and their choices, allowing me, as Hatch (2002) noted, in order “to understand the world from the perspectives of those living in it.” (p. 7). Based on the criteria set forth for qualitative studies set by experts Miles and Huberman (1984), Merriam (2009), and Hatch (2002), my study best fit the qualitative approach.

The goal of this study was to discover the impact of online formative assessment in reading on teachers’ planning and instruction in the middle school English Language Arts classroom. Although OFAPs have been available for several years, there is a gap in the literature concerning the impact these programs are making in the classroom. Talking with teachers and observing them in their classrooms was the most effective way to understand how teachers use OFAPs to analyze student data, process this information, and then plan and implement instruction. A multiple-case study design, therefore, best fits the goals for this study. As Creswell (2007) explains, the case study requires that the researcher “[explore] a bounded system (a case) . . . over time, through detailed, in-depth data collection involving multiple sources of information . . . and reports a case description and case-based themes (p.73, emphasis in original).
Using a case study approach also allowed me to peel back layers of unexplored teacher practice to gain unique insights into their thinking as they analyzed data and planned and implemented instruction. Merriam (2009) notes that a case study “offers insights and illuminates meanings that expand its readers’ experiences (51) and that it “can bring about the discovery of new meaning, extend the reader’s experience, or confirm what is known” (p. 44). Further, since the practice of teachers cannot be separated from the context in which they teach, I knew that I would need to look at the cases holistically. Observing teachers in their classrooms as they worked with students allowed me to study them in what Yin (1981) calls their “real-life context (p. 59) and was important in understanding the application of teachers’ data analysis and planning. This study, therefore follows Yin’s (1981) suggestion that a case study examine: (a) a contemporary phenomenon in its real-life context, especially when (b) the boundaries between phenomenon and context are not clearly evident.” (p. 59).

Four cases were used in this study to address recommendations of qualitative experts. For example, Miles and Huberman (1984) point out that “multiple-site [case] studies are especially appealing because they can purposively sample, and thereby make claims about, a larger universe of people, settings, events, or processes than can single-site studies” (p.37). Merriam (2009) agrees that more cases mean a more “compelling” study (p. 49). Creswell (2007) suggests that researchers include no more than four cases to avoid weakening the analysis of the study. Therefore, using four cases allowed me to use both within and cross-case analyses, which afforded this study a broader generalization and resulted in forceful and convincing findings.
Theoretical Framework

This study of four teachers and how their planning and instructional practice was influenced by the use of online formative assessment systems will be framed by social constructivism. Creswell (2007) describes constructivism as a world view in which “individuals seek understanding of the world in which they live and work” (p. 20). The constructivist researcher uses experience and reflection to interpret rather than to simply report, with the goal of making clear that which has been unclear. In qualitative research, constructivists seek “a complexity of views” and rely “as much as possible on the participants’ views of the situation” (Creswell, 2007, p. 20). Using the constructivist paradigm, I entered the worlds of the study’s four teachers as they analyzed data, and planned and implemented instruction. As Hatch (2002) explains, “It is through mutual engagement that researchers and respondents construct the subjective reality that is under investigation” (p. 15).

Site Selection and Participants

The study’s design required that participants (1) teach English Language Arts in a North Carolina middle school; (2) use an online formative assessment program regularly; (3) use data from that program to plan and inform their instruction. I sought to include subjects from at least two different middle schools that differed in student demographics and represented a range of teaching experience so that the experiences and perspectives of the subjects and their settings would provide a richer, broader response to the research question. Miles and Huberman (1984) maintain that sampling includes not only participants but settings, events, and social processes as well. I planned to use purposeful, or criterion,
sampling to find these subjects, settings, events, and social processes. According to Patton (2002), “the logic and power of purposeful sampling lies in selecting information rich cases for study in depth. Information rich cases are those from which one can learn a great deal about the issues of central importance to the purpose of the inquiry, thus the term purposeful sampling” (p.230, emphasis in original). Therefore, the researcher first determines what criteria will “reflect the purpose of the study” and facilitate finding the cases that are “information-rich” (Merriam, 2009, p. 78). Therefore, I not only sought subjects who would fit the three primary criteria, North Carolina English Language Arts teachers who regularly used an online formative assessment program as well as the program’s data to plan and inform their instruction, but who would also fit the secondary ones of diversity of teaching experience and of student demographics.

To select the research sites and participants, I first contacted principals and other administrators in North Carolina districts representing a variety of student demographic to find middle schools using OFAPs. Then I asked these same contacts for their help and permission in finding volunteers in middle school English Language Arts. From my research and knowledge of the availability of a state-wide OFAP, I had assumed that most schools in the state were using one or more of these programs, especially since the state-sponsored one is now free to students, teachers, and administrators. It was a surprise to learn that while many schools had at one point used OFAPs, these same schools were either not using them consistently at the time of the study nor were they using them school- or district-wide. Further, while principals seemed happy to help with my research, they were understandably reluctant to allow any direct access to teachers. Because of this disconnect between the
researcher and teachers, identifying potential subjects was problematic. The administrators preferred to send general emails to their teachers asking if any were interested in participating in a graduate study on using online formative assessment.

Therefore, I depended on a combination of criterion sampling and convenience sampling to secure subjects. According to Merriam (2009), in convenience sampling, sites and subjects are chosen based on “time, money, location, availability of sites or respondents and so on (p. 79) without consideration of teacher preparation or background I secured two of the participants through a former colleague who was a district administrator in one of the selected districts. She introduced me to a middle school assistant principal in the district, who then provided two subjects. The other two participants were former colleagues who I knew had used at least one OFAP in the past. They were still using OFAPs and volunteered to participate in the study. I secured formal permission from their district and school administration and then emailed the teachers to give an overall view of the study’s design, the time required of the participants, procedures, and methods. Each subject was also provided with the North Carolina State University Informed Consent Form for Research enclosed (Appendix A).

Participants

At the time of the study, all four participants in this study taught reading using the Common Core English Language Arts reading curriculum at the middle school level. Two participants are licensed in both middle grades and high school English Language Arts and are veteran teachers. Two participants are in the first few years of their career and are corps members in the Teach-for-America program. Two of the teachers have advanced degrees,
and one entered teaching after a successful career as a lawyer. Two of the subjects teach fifth grade, one teaches seventh grade, and one teaches eighth grade. All four teachers teach advanced classes as well as regular ones.

**Mrs. Williams.** Mrs. Williams is a Caucasian female in her fifties who had 16 years of experience as a lawyer in the Northeast when she relocated the western part of North Carolina. Although she had a BA in political science as well as a law degree, she decided to change careers when she saw the critical need for teachers in her new home state. She completed a state-sponsored program designed to train and support those who enter the teaching field after working in other fields and is now state-licensed in middle grades English Language Arts, middle grades Social Studies, and high school English. At the time of this study, Mrs. Williams was in her tenth year of teaching. Although at the time of the study, she was teaching 8th graders at Dorothy Edwards Middle School, she had taught 7th grade as well.

**Ms. Quinn.** Ms. Quinn is a Caucasian female in her early twenties who was in her first year of teaching at the time of this study. She was placed in Nolan Reid Middle School by Teach for America, an organization whose mission is to help “provide an excellent education for kids in low-income communities” (Teach for America, 2015, “A Solvable Problem”). Individuals, often recent college graduates who may or may not have degrees in education, are recruited, given six weeks of training over the summer, and then placed in high-poverty schools across the nation to serve a minimum of two years (Teach for America, 2015). Although Ms. Quinn has a Bachelor’s degree in political science and anthropology and has received some teacher training from Teach for America, she is not state-licensed.
Her decision to join Teach for America was so she could help underserved students. Student population for grades five through eight at Nolan Reid totals 257. Due to the fact that teachers are funded by the state based on student population, there are only three fifth grade core teachers to cover the four core subjects of English Language Arts, Math, Science, and Social Studies. Therefore, Ms. Quinn teaches a combination of English Language Arts and social studies during her three class periods. At the time of this study, she was looking forward to her second year and using the knowledge she gained in her first year.

Mrs. Kilpatrick. Mrs. Kilpatrick, a Caucasian female in her early twenties, also came to Nolan Reid Middle School through Teach for America. Mrs. Kilpatrick is state-licensed in elementary education K-6 and English for Speakers of Other Languages. She chose not to teach in her home state due to the political climate surrounding education there. She thought she could make a significant contribution through Teach for America. At the time of this study, Mrs. Kilpatrick was in her second year at Nolan Reid, teaching fifth grade, and planned to return for at least another year. Since there are only three core fifth grade teachers at her school, Mrs. Kilpatrick shared responsibility with Ms. Quinn for teaching English Language Arts. Although her teaching assignment was science, Mrs. Kilpatrick was also responsible for covering the fifth grade Common Core State Standards in reading for information.

Mrs. Jackson. Mrs. Jackson is a Caucasian female in her early forties and at the time of the study was in her eleventh year at Turtle Creek Middle School. She has a bachelor's degree in Education/English Language Arts and a Master’s degree in Educational Technology. She holds North Carolina licenses in Middle Grades Language Arts 6-9.
English 9-12, and Academically and Intellectually Gifted K-12. At the time of the study, Mrs. Jackson had twenty years of experience teaching English Language Arts, most of which had been at the middle school level. During this study, Mrs. Jackson was teaching two regular classes and two advanced classes of seventh grade English Language Arts. In addition, she served as chairperson for the school’s English Language Arts department and for the 7th grade and had various other leadership roles. She also served on a district committee for staff development and curriculum revision.

The following table provides a visual representation of identifying characteristics of the participants including age, ethnicity, experience, and teaching assignment.

<table>
<thead>
<tr>
<th></th>
<th>Mrs. Williams</th>
<th>Ms. Quinn</th>
<th>Mrs. Kilpatrick</th>
<th>Mrs. Jackson</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>School</strong></td>
<td>Edwards</td>
<td>Nolan Reid</td>
<td>Nolan Reid</td>
<td>Turtle Creek</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td>50-55</td>
<td>20-25</td>
<td>25-30</td>
<td>40-45</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td><strong>Grade Assignment</strong></td>
<td>8th</td>
<td>5th</td>
<td>5th</td>
<td>7th</td>
</tr>
<tr>
<td><strong>Teaching Assignment</strong></td>
<td>ELA</td>
<td>ELA/SS</td>
<td>ELA/Science</td>
<td>ELA</td>
</tr>
<tr>
<td><strong>Years of Experience</strong></td>
<td>10-15</td>
<td>1-5</td>
<td>1-5</td>
<td>20-25</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td>J.D.</td>
<td>B.A.</td>
<td>B.A.</td>
<td>M.A.</td>
</tr>
<tr>
<td><strong>Teaching Certification</strong></td>
<td>ELA and Others</td>
<td>None</td>
<td>Elementary Education</td>
<td>ELA and Others</td>
</tr>
</tbody>
</table>
Settings

The three schools serving as sites for this study are all public schools representing three different school districts and three different geographical areas of the North Carolina, which allowed the study a broader representation of participants and students. One school is located in the rural eastern part of the state, one is in an urban area in the central part of the state, and the third was in located in a western suburban area.

Edwards Middle School. Edwards Middle School, where Mrs. Williams teaches, is a suburban school in the western part North Carolina. It is a traditional calendar public school in a district with several other middle schools. There are 40 teachers and 661 students in grades 7 and 8. Two thirds of the students at Edwards are Caucasian, 20% are Hispanic, and another 13% are African-American or Asian. 49% of the students qualify for free or reduced lunch. In 2012-2013, almost half of Edwards’s eighth grade students were proficient in reading and slightly fewer students were proficient in math, based on the state End of Grade reading test, somewhat higher than the state average. These scores are consistent with scores from previous years. All teachers are fully licensed, and almost all are highly qualified. Edwards enjoys a low 3% turnover rate for teachers, compared to 16% statewide. Daily student attendance matched the state average (NC Report Cards, 2013). All classrooms at Edwards are connected to the Internet, and students have small school-supplied laptops. There is also a computer lab in the media center.

Nolan Reid Middle School. Nolan Reid Middle School, where Ms. Quinn and Mrs. Kilpatrick teach, is a part of a small public four-school district in eastern North Carolina. Nolan Reid follows a traditional calendar and serves 257 students in grades 5-8, 98% of
whom are African-American, while the rest are other minorities. Nolan Reid is recognized by the state as a high-poverty school, with 91% of the students receiving free or reduced lunch. In 2013, less than five percent of fifth grade students at Nolan Reid were proficient in reading and math. Nolan Reid struggles to recruit and retain teachers who are fully licensed, and only two-thirds of their 21 teachers were fully licensed in the 2012-2013 school year, compared to the state average of 95%. Further, 47% of the teachers in 2012-2013 had fewer than four years of experience, more than twice the state average. Teacher turnover, at 27%, was also almost twice the state average. Average daily school attendance is 96%, compared to 95% statewide (NC Report Cards, 2013). Each grade has one team of teachers. Since teachers are allocated by the state based on enrollment, fifth grade, which at the time of this study had fifty students, is allocated three teachers, one of whom teaches Math, one Science, and one English Language Arts. In her English Language Arts class, Ms. Quinn must cover the Social Studies standards through curriculum integration. Mrs. Kilpatrick uses her Science class to teach the standards for reading informational text. Students have three 90 minutes core classes, two electives, and a 45 minute class for review, for which they rotate each day among the three teachers. All classrooms are connected to the Internet and the ratio of students to instructional digital learning device (e.g., desktops, laptops, or tablets) is a little less than 1:1. All classrooms have a single classroom desktop computer, and teachers have school-supplied laptops (NC Report Cards, 2013). There is also a computer lab.

Turtle Creek Middle. Turtle Creek Middle School, where Mrs. Jackson teaches, is part of a suburban public school district located in the central part of North Carolina. It is a public middle school serving grades six through eight and follows a traditional calendar.
There are 740 students and 43 teachers. Almost fifty percent of Turtle Creek’s students are Caucasian and the rest are evenly divided between African-American and Hispanic. 54% of the students are eligible for free or reduced lunch. All teachers at Turtle Creek are fully licensed by the state and are highly qualified, a federal standard indicating that they have met certain criteria, most important of which is that they are state-licensed in the area in which they teach. Teacher turnover for Turtle Creek is somewhat above average for the state, which in 2013 was 16%. Almost half of 7th graders at Turtle Creek were proficient in reading in 2013 and only slightly fewer were proficient in math, based on the End of Grade test given by the state, a percentage that was about the same statewide. Average daily school attendance matched the state average. The school is organized around the middle school concept of teams, with one teacher for each of the core subjects of Math, English Language Arts, Science, and Social Studies. Students attend four core classes a day of 70 minutes each in addition to two elective classes. Each student has a laptop provided by the school, and Internet access is available throughout Turtle Creek. The ratio of students to instructional digital learning device (e.g., desktops, laptops, or tablets) is 2:3 (NC Report Cards, 2013).

The table on the following page provides a summary of the demographics of the three schools.
Table 4. Demographic Information for Research Sites - 2012/2013

<table>
<thead>
<tr>
<th></th>
<th>Edwards</th>
<th>Nolan Reid</th>
<th>Turtle Creek</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grades Served</td>
<td>7-8</td>
<td>5-8</td>
<td>6-8</td>
</tr>
<tr>
<td>Setting</td>
<td>Rural</td>
<td>Rural</td>
<td>Suburban</td>
</tr>
<tr>
<td>Calendar</td>
<td>Traditional</td>
<td>Traditional</td>
<td>Traditional</td>
</tr>
<tr>
<td>Student Population</td>
<td>661</td>
<td>220</td>
<td>740</td>
</tr>
<tr>
<td>Teacher Population</td>
<td>40</td>
<td>19</td>
<td>43</td>
</tr>
<tr>
<td>Fully-Licensed Teachers</td>
<td>100%</td>
<td>63%</td>
<td>100%</td>
</tr>
<tr>
<td>Average Class Size</td>
<td>25</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>Student Ethnicity</td>
<td>64% Caucasian</td>
<td>98% African-American</td>
<td>47% Caucasian</td>
</tr>
<tr>
<td></td>
<td>20% Hispanic</td>
<td>2% Other Minorities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7% Asian</td>
<td>24% Hispanic</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6% African-American</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free/Reduced Lunch</td>
<td>49%</td>
<td>91%</td>
<td>54%</td>
</tr>
<tr>
<td>1:1 Laptops</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Teacher Participants</td>
<td>Mrs. Williams</td>
<td>Ms. Quinn</td>
<td>Mrs. Jackson</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mrs. Kilpatrick</td>
<td></td>
</tr>
</tbody>
</table>

**Data Collection**

To learn as much as possible about the effect online formative assessment has on the planning and instruction of these four teachers, three types of data were collected and analyzed: interviews, observations, and documents. Three types of data allowed me to use triangulation to increase the validity and reliability of the data. Using interviews, observations and various documents also allowed for the focus and breadth of information
needed, while limiting the amount of time the participants had to spend on participation in
the study. Use of these three types of data collection is an accepted method for case study
research (Merriam, 2009).

**Interviews.** During the study, I conducted two interviews with each participant, one
before the observation and one afterwards. The interviews were semi-structured in nature,
based on an interview guide (see Appendices B and C), which ensured that I asked questions
that would elicit complete information while providing freedom to probe for more
information or to further pursue a pertinent topic. The pre-observation interviews were
either conducted on the day before the observation or on the same day. The post-observation
interviews were conducted on the same day as the observations.

With the exception of Mrs. Jackson, who was unable to meet for a pre-observation
interview, all interviews were conducted in person and digitally recorded. I obtained
permission from the subjects to record the interviews and additionally, I took field notes as
we talked. The field notes I took during the interviews were invaluable in that they later
reminded me of what at the time seemed most important in what each participant said and
allowed me to record impressions, thoughts, and interpretations as the interview progressed.
Since a face-to-face interview was not possible for Mrs. Jackson, I interviewed her by phone
the night before the observation, during which I took extensive field notes. At this time, I
also emailed the pre-observation interview questions to her, to which she responded in
writing and emailed back to me the same day. The purpose of these interviews were two-
fold: to establish a rapport with each participant and to elicit background information about
their training, teaching experience, experience with online formative assessment programs,
and challenges and benefits of using them. For each interview, I followed an interview outline (Appendix A) but supplemented it with clarifying questions to ensure understanding. All interviews lasted between forty-five minutes to an hour.

The pre-observation interview began with a short exchange of information to establish rapport. This was especially important with Ms. Quinn and Mrs. Kilpatrick, whom I had never met. I reiterated the purpose of the study and thanked them for their time. In the pre-observation interview, I asked questions to determine which OFAPs the participant used and how they were used, the frequency use, as well as the participant's knowledge and experience with formative assessment, perception of its effectiveness, and plans to remediate missing skills. I also asked about the choices for instruction the participants had made based on the data they had analyzed, how the lesson plan for the observation fit into the teacher’s overall plans and instructional goals and how it fit into the formative assessment cycle, and. During this interview, I gathered information about each participant’s education, teaching experience and licensure, knowledge and experience with OFAPs, and her opinion of their credibility in measuring student proficiency. Of paramount importance to the study was how these teachers used the OFAPs and the data the programs provided. Therefore, most of the time during these interviews focused on this discussion.

Each of the post-observation interviews, which lasted between 45-60 minutes, was conducted in the teacher’s classroom on the same day as the observation. Again, I used an interview outline (Appendix C) and asked additional questions for clarity, understanding, and to further explore the responses of the participants. I also asked questions that emerged as a result of the class observation, such as the experience students had with certain strategies,
why the teacher chose a certain strategy during the class, or what surprised her as she worked with students. During these post-observation interviews, I also learned how the teacher planned to continue the formative assessment cycle, whether by providing more practice, more targeted intervention, more data analysis, or a new assessment.

Observations. A second method of data collection was observations of teacher participants as they conducted their classrooms and implemented the instruction they planned as a result of their OFAP data analysis. Observations allow researchers to view the phenomenon under study in context and to add to the rich, thick description that is characteristic of the case study (Merriam, 2009). At two of the sites, I observed two class periods each. These classes were fifty to seventy minutes in length, and each participant used the same basic lesson plans for both classes. At Nolan Reid Middle School, I observed one class conducted by Ms. Quinn and one by Mrs. Kilpatrick. These classes lasted 90 minutes each. During these observations, I took extensive field notes that included my observations on a laptop computer of the classroom setting, student demographics, instructional focus, and activities. Observing these teachers as they conducted the lessons they had planned as a result of their analysis of reading data from an online formative assessment program (OFAP) allowed me to “enter into” their world, viewing it as much as possible the way they experienced it. It also allowed me to see exactly how their plan was implemented and to some extent, its success. Further, as Creswell (2007) explains “the [constructivist] researcher listens carefully to what people say and do in their life setting” (p. 21). Observing these teachers in their environment helped me to understand the descriptions they had given in the pre-observation interview. Included in these field notes were
descriptions of the schools and classrooms as well as impressions and interpretations that helped to enhance my understanding of the data I collected. For example, I noted my impressions of the relationships between students and teachers, the pacing of the class, and how the attitudes toward teaching the participants displayed. I also included my interpretations of how the class and the teacher interacted and the attitudes toward learning exhibited by the students.

**Documents.** The third data source and strategy for attaining an understanding of teachers’ use of OFAPs were classroom instructional documents and OFAP materials. I collected and examined student handouts and answer keys of the printed assessment used during the observation. I also inspected sample reports of student performance provided by OFAPs. These reports can include results for individual students or entire classes on single assessments or on a group of assessments. Finally, I examined available information on the web about the OFAPs teachers were using. This information included the stated purposes of the programs (e.g., test preparation, practice, and formative assessment), the features they offered, and their applications for the classroom. The examination of the OFAP-specific information helped me create a context for the way in which students and teachers used each program.

**Data Analysis**

Several types of data analysis were used to increase the reliability, validity, and rigor of the study. Sources of data included two interviews and a classroom observation for each participant as well as documents, such as class handouts, lesson plans, and teacher notes. The nature of qualitative research, which “describes people acting in events” (Firestone,
1987, p. 19) depends on the interpretation by the researcher of the events, subjects’ responses and interpretations of those events (Merriam, 2009, p. 210). It is, therefore, essential that the researcher “understand the perspectives of those involved” in the study, “to uncover the complexity of human behavior in a contextual framework, and to present a holistic interpretation of what is happening” (Merriam, 2009, p. 215). How a researcher collects, analyzes and interprets data, therefore, is crucial.

In order to build a clear perspective on the complexity of using OFAP data in planning and instruction, this study included interviews, observations, and documents as sources of data. All interviews were digitally recorded and listened to numerous times. After each interview, I transcribed the data using a computer word processing program. Transcribing the interviews myself allowed me to become familiar with the contents of each interview and, as a result, came to feel that I intimately knew each participant. Field notes from the interviews and observations were also transcribed using a word processing program on my computer, and then they, along with the information gleaned from the documents, were coded and sorted using tools on both the word processing and spreadsheet programs on my computer.

To ensure confidentiality, pseudonyms were assigned to all four teacher participants as well as to their schools, and any identifying names were removed from the recordings. No student names were viewed or collected by the researcher at any time. All transcriptions and notes were kept on a password-protected computer to which only the researcher had access. In analyzing data, Merriam (2009) suggests that the researcher use the step-by-step method of analysis, moving from broad categories to more narrow ones. Therefore, I carefully
examined each piece of data, mining it for details that would emerge from close analysis and lead to identification of emergent themes. I used open coding for this part of the data analysis, assigning general labels to small pieces of information that correlated in some way to the research question (Merriam, 2009). I used in-document notes at first, but found that using color to code each kind of information using my computer’s word processing program was more efficient and effective. I then copied these bits of data onto a computer spreadsheet in columns labeled by participant and code. I followed this procedure with each set of documents and then reviewed the data again with fresh eyes.

Throughout this analysis, I was careful to use triangulation, comparing within and between the transcripts and documents. Triangulation is a method for “checking for consistency” in the data (Patton, 1990, p. 464) by cross-checking within and between methods to validate interpretation and themes. According to Merriam (2009), the use of triangulation can help ensure validity and reliability. Huberman and Miles (2002) state that “the triangulation made possible by multiple data collection methods provides stronger substantiation of constructs and hypotheses” (p. 14). Furthermore, using both between-methods triangulation and within-methods triangulation strengthens the reliability of the study (Jonson & Jehn, 2009).

Accordingly, I began to see broader categories emerging from within both the individual cases and across cases, so was able to group the data within cases to narrow the emerging categories. I followed this same procedure across cases to determine which categories/themes occurred in more than one case, using the constant comparison method to find similarities and differences in themes. Again, I grouped categories to find broader and
broader themes, eventually identifying four emerging themes. According to Merriam (2009), this method, both inductive and comparative, is particularly well-matched to qualitative studies. Similarly, Silverman (2006) maintains constant comparison of data provides an effective way to ensure validity in qualitative research. Beginning with broad categories, then breaking them into smaller categories, and finally building them once again into broad categories, I was able to view the data from varying perspectives.

**Research Reliability and Validity**

In addition to the triangulation and constant comparison, I used member checks to ensure reliability and validity. I shared initial findings by email with each participant and invited them to respond with corrections, additions, or questions to ensure my interpretations “[captured] their perspective” (Merriam, 2009, p. 217). Lincoln and Guba (1985) contend that using member checks is an especially important technique for researchers with a constructivist perspective, since the constructivist tries to construct meaning in partnership with the research participants. The constructivist researcher works to create meaning in collaboration with participants. Member checks allow participants another opportunity for this collaboration, and “inviting participants to give feedback on interpretations is a vital step in co-construction” (Hatch, 2002, p. 188). In fact, Maxwell (2005) calls member checks the single most important way of ruling out the possibility of misinterpreting the meaning of what participants say and do and the perspective they have on what is going on, as well as being an important way of identifying your own biases and misunderstanding of what you observed” (p. 111).
Lincoln and Guba (1985) agree, claiming that member checking is “the most crucial technique for establishing credibility” (p. 314). Three of the participants requested no changes or additions. The fourth participant requested only one minor change in my interpretation of what she meant. This change was made and I sent her the corrected findings, which she approved.

**Safeguards against Researcher Bias**

The essence of qualitative research is “understanding the meaning individuals construct in order to participate in their social lives” (Hatch, 2002). The researcher, then, interprets data that includes both observable external actions and “inner states” (p. 9). As pure objectivity in qualitative research is impossible, I had to be cognizant of my subjectivities in order to reflect as accurately as possible the findings of the study. Recognizing subjectivity is to delve into the experiences, characteristics and knowledge that have shaped the researcher. Important to note is that for the researcher, subjectivity “is a factor to neither to minimize nor to embrace,” (Hays & Singh, 2011, p. 145), as it both benefits the study and makes it more complicated.

According to Ladkin (2005), one way researchers can deal with subjectivity is to use *critical subjectivity*. “Critical subjectivity encourages inquirers to notice the particular frames of reference they bring to any inquiry arena, including, among others, their political, racial, cultural, or gendered orientation” (p. 118). Other researchers refer to critical subjectivity as *reflexivity*. Reflexivity is the researcher’s way of trying to accurately represent the lived experience of the subjects of the research while acknowledging that as human beings, this is virtually impossible (Pillow, 2003). Lincoln and Guba (2003) define
reflexivity as “the process of reflecting critically on the self as researcher” (p. 283). Using reflexivity to acknowledge my “biases, values, and experiences” (Creswell, 2007, p. 243) was of paramount importance.

A useful strategy to address, yet embrace, researcher subjectivity is bracketing, or setting aside one’s subjectivities (Bednall, 2006; Crotty, 2002). Ladkin (2005) further suggests that bracketing is one way to “‘reach out’ beyond [our] subjectivities . . . putting aside one’s preconceptions, expectations, or culturally determined interpretations in order to encounter the ‘essence’ of a phenomenon” (p. 119). Following these suggestions increase the reliability of my study, I used bracketing to ensure that my perspectives and experiences, as well as any biases, did not sway the interpretation of the study’s data.

I understand that I bring certain biases, values and experiences applicable to this research study. I am a twenty-seven-year veteran of the middle school classroom, with a master’s degree in special education and a variety of teaching experiences, including different configurations of special education classes, at-risk, academically and intellectually gifted, and regular classes, as well as all core subjects. Most of my career, however, was spent as an English Language Arts teacher. For many years, I believed that good teaching would be enough to enable my students to succeed in whatever they were called upon to do, but the increasing number of students in my classes who were English-language learners, students with identified special learning needs, and students who could not read on grade level made me realize that good teaching was not enough.

My interest in finding better methods of teaching and assessing students led to work in both authentic tasks and assessments, with further work as a member of the Governor’s
Committee on Portfolio Development from 1998-2000. I regularly used performance tasks and authentic assessment in my classrooms. From 2004-2008 I served as a member of my district’s data team, which studied ways to use data from various sources to enhance instruction and student success. During that period of time, I also came to better understand the formative assessment model and the impact it can make on student learning. I learned even more about formative assessment through my work in various capacities of assessment development for online formative assessment programs. No one is more surprised than I that my work after teaching would be in the area of assessment. Although my work currently involves writing assessment, my overall passion lies in seeing that assessment is always used to enhance instruction and student understanding.

My experience as a teacher and in assessment also brings strengths to the study. My familiarity with the classroom, the curriculum, special needs of many different students, and time constraints make me a knowledgeable and empathetic researcher. My familiarity and experience with various forms of assessment also allowed me insights into the participants’ use of OFAPs and the formative assessment in general, as well as the challenges they face in the classrooms of today.

It is also important to explain that, although I did not have a personal relationship with any of the participants, I knew two of them slightly as former colleagues when I worked for a state university. It was due to this relationship and a common interest in formative assessment that they agreed to participate in this study.
Ethical Issues

Before beginning this study, I obtained the approval of the North Carolina State University Institutional Review Board for the Protection of Human Subjects in Research (IRB), following its guidelines to protect the identities of the participants and to ensure that the participants, their students, and their schools do not suffer consequences as a result of this study. Pseudonyms were assigned to participants, as well as their schools, immediately upon securing them. The digital recordings of the interviews sessions, field notes, transcriptions and class documents were kept either in a locked file drawer or a password-protector computer at my home.

I was transparent in my dealings with the participants, making sure that they were aware of my perspectives, biases, and experiences germane to this study. It was of primary importance to me that I treat the participants and what they shared with the utmost respect and consideration. They received no compensation for their time other than my extreme gratitude.

Limitations of the Study

Being able to “peek in” on the way teachers analyze student data, process this information, and then plan and implement instruction for their students is a rare opportunity. Though two teachers might have the same training, teach at the same school, have the same number of years of experience, and even the same teaching assignment, each will plan and deliver instruction in unique, personal styles. Verbalizing exactly what each teacher was thinking when she reflects and plans is often difficult, and adding even slightly to the understanding of teachers’ metacognitive dialogue is exciting.
Using four teachers for this study allowed me to interview and observe each in their classrooms and delve into the thinking, analysis, planning, and instruction each teacher conducted as a result of formative assessment data. Thus this study was one of depth rather than breadth. Although focusing on only four teachers’ experiences also limits this study, the findings will add to the field of knowledge about how middle school English Language Arts teachers help students increase their proficiency in reading through the analysis of data, planning, and instruction based their use of online formative assessment programs.

**Summary**

This chapter has described the methodology used in investigate how the use of an online formative assessment program in reading impacts teachers’ planning and instruction in the middle school English Language Arts Classroom. The next chapter provides a detailed discussion of the findings resulting from this study.
CHAPTER 4: Findings

Introduction and Framework for Analysis

This chapter deals with the analysis undertaken to answer the research question: How does the use of an online formative assessment program impact teacher planning and reading instruction in the middle school English Language Arts classroom? To answer this question, a multiple case study was conducted focusing on four middle school English Language Arts teachers who regularly incorporated one or more online formative assessment programs (OFAPs) in their reading instruction. The chapter is divided into sections representing each of the four cases, including descriptions of the subjects’ school settings, backgrounds, experience with OFAPs, data analysis and planning, targeted instruction and intervention, and the nature of formative assessment. Major themes that emerged in the cross-case analysis are described in the section that follows each of the case portraits. The chapter concludes with a summary of the chapter.

Multiple Case Study Details

This study focused on four middle school English Language Arts teachers who regularly used one or more online formative assessment programs as a part of their reading instruction. In order to understand as completely as possible the ways in which these programs influenced the planning and instruction of these teachers, I used three types of data: interviews, observations, and documents. I interviewed each of the four teachers before and after an observation of their classroom instruction and examined classroom documents. Because context is so important in understanding, I include descriptions of the subjects’ schools and school settings, including size, population, demographics, and past performance.
on state-administered end-of-grade tests. I also include background information about each subject, including education, experience, licensure, and prior experience with OFAPs as well as the various OFAPs they have used. I describe the processes each uses in determining next steps in instruction, based on OFAP data analysis and their resulting instruction. Finally I detail participant reflections on the value these programs bring to their instruction and their plans for using them in the future.

Case Study Profiles

This multiple-case study involved four in-service middle school English Language Arts (ELA) teachers who used online formative assessment programs (OFAPs) on a regular basis. Three sources of data, interviews, observations, and documents, were used to create complete profiles of each of these teachers. The documents included lesson handouts, OFAP data samples, and descriptions of various online programs from their websites. Included in each profile is the subject’s background, school, and classroom setting. To give context to each subject’s profile, I also include prior experience with OFAPs. I then describe each subject’s data analysis and planning process, the intervention lesson and instruction. Each profile concludes with a section describing how each subject adapts formative assessment to fit the circumstances and needs of her classroom.

Mrs. Williams. When this study took place, Mrs. Williams was teaching eighth grade at Edwards Middle School housed in a relatively new, one-story school building nestled in rolling hills. Edwards Middle School served 661 students in grades seven and eight, of which 64% were Caucasian, 20% were Hispanic and another 13% were African-American or Asian. Class sizes average twenty-five students.
The walls of Mrs. Williams’s large, sun-filled classroom were decorated with historical and humorous educational posters, maps, and inspirational quotations. Two double cabinets stood along one wall, with a large triple window stretching across another, with work tables and a teacher’s desk in front. At the front of the room were three upholstered chairs under a large white board and a door to the hallway. Two large bookcases and two smaller ones held novels and other books. A two-foot Eifel Tower stood along one wall. Students had small school-provided laptops which they carried from class to class. Flat-topped desks were arranged in threes in semicircular rows. “I often team students up and this seating enables me to do that quickly,” she explained. Of the thirty-two students in her advanced class, twenty-two were Caucasian, nine were Hispanic, two were Asian, and one was African-American. Males and females were evenly divided.

Mrs. Williams had a successful career as a lawyer before deciding to become a teacher and at the time of the study was in her tenth year of teaching. She is state certified in Middle Grades English Language Arts (ELA), Middle Grades Social Studies, and High School English. Mrs. Williams was teaching four classes of about fifty minutes each, of which two were advanced and two were regular education classes.

**Experience with OFAPs.** Mrs. Williams used several types of informal formative assessment on a daily basis. “I take all sorts of learning checks and journaling and [things like that], “she explained in the pre-observation interview. She had also used Study Island, she continued. “We used Study Island. I really liked it. I liked it very much. I liked it because it took each student where they were. I could [take students] down to a lower grade.
That was really great.” At the time of the study, Mrs. Williams was in her second year of using ClassScape and explained her experience with this OFAP.

I really like ClassScape. Overall, I think it’s excellent. I like it because of the rigor. I think the items are excellent. They’re well-developed [and] they hit at the common core standards they are supposed to. It’s challenging to students and there’s a range of items to choose from. I love the selections. I love the content. I think [it] is very, very useful, and the kids respond well to it.

She explained in the pre-observation interview that the immediate feedback offered by the program was a very important component.

As far as being able to see the results immediately, which is awesome for them, and being able to see what was the right answer and figure it out - that’s sort of very important to the learning process. They like that and it’s that’s very helpful to me. I think the more they use it the better it is.

Mrs. Williams, prior to this study, served as an item writer and reviewer for ClassScape. She wrote questions for reading passages for use in ClassScape, as well as reviewing questions other teachers wrote. In her pre-observation interview, she openly expressed, her view about how this work had helped her. “ClassScape has made me a better teacher by 1000 percent.”

In the same interview, she also described another advantage of her work with this OFAP. As with most online programs, ClassScape provides teachers access to the content for their grade level. But
... because I write for ClassScape, I have access to English II that I can use for 8th grade and I have access to 7th grade that I can use in 8th grade and so it’s the same stuff as Study Island – I can differentiate.

Although Mrs. Williams used ClassScape, the lack of computers at her school limited her access to the program online, so she often printed copies of the assessments to use in her classrooms. Despite this limitation, Mrs. Williams was clear in her interviews that ClassScape had been extremely beneficial to her students, requiring them to think about why answers are correct rather than just choosing an answer. Even though she was only in her second year of using ClassScape, she said that she could tell that it had made a difference on her students’ achievement on the EOGs.

Using online programs, she explained in her pre-observation interview, was not only helpful, but her district and school administrators openly encouraged teachers to use online programs to assess students. State summative assessments would eventually be online; however, moving toward a completely-online assessment system was slow. Although students at Edwards have laptops, the screens are small, making accessing the text online while answering questions difficult. Edwards Middle School had a computer lab with desktops in the media center, and Mrs. Williams had used these for assessment, since their screens were larger. However, the year of the study, Edwards Middle School had acquired two new online math programs which required students to use the media center’s computer lab on a regular basis. This situation meant that Mrs. Williams was unable to use the computer lab as much as she would like.
The district encourages us to use [ClassScape]. They want us to use it. I had a principal who wanted us to use it - I think he said twice a nine weeks - but I’m not able to get it in [to the computer lab] and now, with the [increased state testing] on computers everyone needs computers for everything.

Mrs. Williams said that if this were not the case, she “would do an assessment every two weeks, if not more. I probably did one six or seven times this year.” One advantage of printing assessments, she said in her post-observation interview, is the administration of the state summative assessment on paper. Printing some assessments on paper allowed Mrs. Williams to be sure students were comfortable using both online and paper/pencil methods, she explained.

[ClassScape] is very informative to kids about [how to use online systems] but as long as [the End of Grade assessment] is a paper and pencil test, the skills are not completely analogous. And so the kids like using it, and it’s a helpful diagnostic to me, but when we are at this point in the year, I have to go back to pencil and paper.

She explained that as long as the state assessments were printed on paper, she wanted her students to write on printed assessments, making notes in the margins as they read the passages, “... and so to get them to do this I need them to practice, and they do.” This was one component in her integration of standards, content instruction, and preparation for reading assessment.

Data analysis and planning. In the pre-observation interview, when I asked Mrs. Williams about how she analyzed data from the assessments and how the data informed her
planning, she said that she began both her targeted intervention and her analysis during the assessment.

I know you’re supposed to study the data and then go back . . . But I feel like my most effective time is when they’re taking it and I can work with them one-on-one. If they have trouble while they’re doing it, I usually will go and talk with them, because I think it’s helpful. And then they show me their score, and I say, ‘Okay, which ones did you miss?’ and ‘I need you to read that question again.’ Then they do their analysis and write [the analysis] down so I can look at it again. It takes longer but it’s so much better, so I try to individualize right there.

Observing and questions students as they took the assessment and then looking over the analysis students did of their answers to assessment questions helped Mrs. Williams see not only which students needed extra help in answering the questions, but also with what part of the assignment they were struggling. She determined whether struggling students were having trouble understanding what they were reading, understanding what the question was asking, or understanding how to determine the best answer to the question. She then took the opportunity to address student needs immediately. Mrs. Williams went on to explain that monitoring students and the way they answer questions during the assessment also allows her to get an overall idea of how the class is doing.

Now, before I even look at the data, I will know there are one or two questions that they are all missing because they’ve all missed #6. Some kids will get 3 out of 10; most kids will get 5 or 6, and so that I know that #8 is a problem. . . . I always have a dummy student [account in the program] or I can pull up [individual student] results
and show [students the problem]. Sometimes I can handle it on an individual basis in
the library and sometimes I just bring it back here and [project the question on the
whiteboard] in here. I can pull it up and say, ‘Oh, okay, let’s all talk about this as a
class.’

This vigilant analysis was also critical for the assessment itself. Although usually a question
that several students miss is an indication of a gap in understanding, Mrs. Williams was
careful to look at the questions themselves as well. She explained, “. . . and sometimes, #8 is
a badly-worded question and [that’s also important].”

Mrs. Williams had taught her students a five-step reading strategy she had devised to help
them understand the reading passage and analyze questions about them. This strategy was
posted on the wall, and included five steps:

1. Mark up the passage

2. Talk to yourself

3. Create a one-sentence summary of the passage

4. Determine what the question is really asking you

5. Find text-based proof of your answer.

This strategy was used any time students were assigned reading passages in her class. For multiple-
choice assessments, Mrs. Williams also required her students to explain why each of their answers
was the best choice.

The day before my observation, Mrs. Williams told me in her pre-observation
interview, she gave students a twelve-question ClassScape assessment that she had printed on
paper. Students used the five strategies that she taught them and explained why they chose
each of their answers. Mrs. Williams collected these assessments and looked over them, keeping a tally of right and wrong for a number of randomly-chosen questions. She explained that she could already see some questions that were problems for students.

You can see the patterns forming – and I looked back and I looked at those and [for each one] tried to [see] why it was such a tricky question. We’re going to go through all of them today because I really want to reinforce what they are doing right, since we’re so close to the [End of Grade] test. You know, I really want to say, “You really focused on this word. You got this right! Good for you! Keep focusing on it.” It’s just basically, you know, a pep talk.

Although some students had not finished the assessment due to illness or absence, Mrs. Williams said that she thought that it was more important for them to hear the discussion of the assessment than to finish it on their own.

**Targeted intervention and instruction.** After our initial interview, when Mrs. Williams had explained her plans for reviewing the assessment with students, I noticed in my observation that Mrs. Williams used a Maya Angelou quotation to model analyzing and reflecting on text and then guide students in practicing those skills. Students were also asked to make connections from the quotation to a current event, and to make historical, cultural, literary, and/or personal applications. Then the students shared these ideas aloud with the class. Mrs. Williams returned the twelve-question assessment students had completed the previous day. She explained to students that she had examined the assessments they completed the day before but had not scored the items for a grade.
I went through these pretty quickly, but I wanted to give you a sense of which [items on the assessment] you got right and which ones you got wrong. I was very pleased with your analysis, how you used the five steps and how you tailored it [to each question]. Take a look at the ones you missed. Were you surprised at any of the ones you missed? Take a look to see if these were the ones you had trouble with when you were taking it.

Mrs. Williams explained to students that when she reviewed their assessments the day before, she noticed which questions had given students the most trouble, and told them that they would go over the all twelve of the assessment questions and would take extra time with the questions identified as problems. While Mrs. Jackson’s main focus of the class period was to go over the assessments and eliminate gaps in understanding, she also took time to examine with the class elements of each selection that were not included in the questions. I observed that Mrs. Jackson kept bringing her students’ attention back to the passages as they went over the assessment. She also asked students to share the thinking process they had used to determine the correct answer. In the post-observation interview, she explained that she followed this process because she thought it was helpful to students to hear how others thought through questions and analyzed reading passages.

Throughout the class period, Mrs. Williams made connections to other literary works the students had read over the year. For example, in discussing one poem, she said, “Imagery of the staircase - where have you seen that before? Right, the crystal stair [from “Mother to Son” by Langdon Hughes].” She then drew comparisons to the two uses of imagery in each
of the poems. She continually asked questions about both the texts and the questions, such as:

1. What does this passage remind you of?
2. What connections can you make?
3. Why is A correct?
4. Why is it C?
5. What helped you in finding the right answer?
6. What thought process did you learn?
7. What word is so important in D?
8. What’s hard about this question? What’s harder about the answers?
9. When people got it wrong, they picked A – why is A wrong?

It seemed evident to me that Mrs. Williams’s emphasis throughout the class period was to help students learn to better analyze text and answer questions about text on their own by using the process she modeled in class: carefully examining and reflecting on the text, asking questions, making connections to other texts and their own experience, and eliminating answers to questions about the text that for one reason or another were not the best answers.

Continuing to model this process, Mrs. Williams invited students at the end of class to respond to the activity they had just completed: “Reflections on this – steps you used, processes - what worked?” One student admitted that each time he had found proof of the correct answer in the text, he got the question right. Mrs. Williams reiterated that: “Yes! Every question he found proof for he got right.” She ended the class with reminders of the “thinking through” strategies they had used throughout the class period.
The nature of formative assessment. In describing her process for analyzing formative assessment data and giving feedback to students, Mrs. Williams explained in her post-observation interview that she did not grade the answers students gave on formative assessments but did take grades on the analysis students did:

Now I probably do it a little differently. but the way I grade ClassScape - you know, it’s formative - I tell my kids they’re getting graded on participation, and what I have them do [is] to go in and see the wrong answer and the right answer . . . I have them take their time and work on it but then . . . where the real learning takes place is when they go back and [determine] why the right answer is the right answer. ‘What did you miss about the question?’ [I] have them do an analysis. ‘What did I miss? Why did I miss this?’ And I have them try to figure out why. ‘Oh, I didn’t read the question carefully enough.’

In addition to her own analysis and reflection on student understanding, analysis and reflection on the part of students also were important components of her grading system. In her post-observation interview, Mrs. Williams said,

I also have them – they keep track of their scores throughout the year – and major grades – and they do a reflection on how they are doing with ClassScape – and what they did well – why one test may have been harder than others and what types of questions they missed and all that. [They do an overall reflection] right after each assessment. So most of the year – we’re at a different time now [anticipating the state End of Grade test] - most of the year, I’ll give them [Common Core State Standard Reading Literature] 2, right? Which is a really important one – finding the central
idea and summarizing and all of that. So I’ll give them ten questions because that’s about how many – my better readers can finish [more] but [that’s enough for] my slower readers [for a fifty-minute class period]. . . and ten is a good number because usually they get most of them right.

Getting most questions right, she explained, made reluctant and struggling students feel successful and more willing to participate in learning and assessment.

**Ms. Quinn and Ms. Kilpatrick.** At the time of this study, both Ms. Quinn and Mrs. Kilpatrick taught fifth grade at Nolan Reid Middle School, a twenty-year old one-story school in a small rural district. Nolan Reid serves grades six through eight and has an enrollment of about 270 students, 99% of whom are African American. The school’s design supports the middle school concept of teams, with one hallway for each grade leading out of a central hub. Because Nolan Reid has so few students, three fifth grade teachers, including Ms. Quinn and Mrs. Kilpatrick, share responsibility for all four core subjects. Although Ms. Quinn’s nominal assignment is English Language Arts, she is also responsible for covering all of the North Carolina Essential Standards for social studies. Mrs. Kilpatrick’s assignment is science but is also responsible for covering the Common Core State English Language Arts Standards for reading for information.

Ms. Quinn and Mrs. Kilpatrick have mirror-image classrooms that are next door to one another. Both teachers, as well as their math colleague, designed their classrooms for group work. In each of their classrooms, seven or eight desks form a semi-circle in front near the white board and bulletin board. Groups of five or six student desks in the rest of the spacious rooms are pushed together to form flat tables. Along an adjacent wall are large
closets and a door leading to the hall. Two other doors, along the opposite walls, lead to an in-class restroom and outside to a tiny patio and the school grounds. Almost all available space on the wall is taken up with inspirational posters and large charts displaying student progress and improvement over the year as measured by their district benchmarks.

Nolan Reid students have historically struggled in all content areas on the North Carolina End of Grade (summative) tests. Reading and math proficiency for fifth grade students the year before this study took place averaged at less than 5%. The highest average for proficiency in any content area for any grade that year was 26%. For this reason, the district has placed a high priority on increasing proficiency through data-driven instruction, or instruction that is based on analysis of data provided through assessments or other means. An amalgamation of challenges has an influence on student proficiency levels. Attracting qualified teachers is a significant challenge for the district, in part because it is small and rural, with few of the opportunities for employment and socialization enjoyed by more urban areas. The district also struggles with resources, in large part due to its small size. Property taxes do not yield the kind of money larger districts enjoy and that can be used for more technology, supplies, and teacher supplements. Low scores, lack of opportunity, and money all make attracting teachers even more difficult. Both Ms. Quinn and Mrs. Parker were placed at Nolan Reid Middle School by Teach for America, which recruits recent college graduates and other volunteers who may or may not have degrees in education to teach for two or more years high-poverty schools across the nation. Teach for America gives their teachers six weeks of training during the summer and then provides support throughout the years. Its mission is to help “provide an excellent education for kids in low-income
communities” (Teach for America, 2015, “A Solvable Problem”). Individuals, often recent college graduates who may or may not have degrees in education, are recruited, given six weeks of training over the summer, and then placed in high-poverty schools across the nation to serve a minimum of two years (Teach for America, 2015).

Each classroom has one desktop computer, and each teacher has a district-furnished laptop. During the time of this study, Internet access was not available in the computer lab. For the year in which this study took place, the district administered Schoolnet benchmarks every four and a half weeks of the first semester. The benchmarks focused on the Common Core State Standards teachers had covered during those periods. Second semester, for reasons undisclosed to teachers, the district switched to Case 21 and administered only a nine-week benchmark. All benchmarks were administered on pencil and paper due to the lack of computers. Teachers had scored the Schoolnet benchmarks, but the Case 21 assessments were sent off to the company to be scored.

During the first semester at Nolan Reid, the first thirty minutes of each day was set aside school-wide to work on reading skills. In her post-observation interview, Ms. Quinn reported,

At the beginning of the year we had a reading block at the beginning of the day . . . it was a school-wide initiative where we would teach kids different strategies using like short passages and five multiple choice questions that go with it, and that was great practice but it wasn't totally aligned with the Common Core. That was kind of our form of remediation because [students] already need the extra support in reading comprehension anyway. Half-way through the year we got rid of the 30 minutes at
the beginning of the day where we would do that because [the administration was] just trying to see if [a special daily reading period would make a difference in students’ reading skills.] [But] they got rid of that and they added [the remediation period] to the end of the day.

This remediation period was also increased to forty minutes, and students were divided into groups by reading proficiency. Instead of having the same teacher each day, students rotated by class from core teacher to core teacher. One day students would have Ms. Quinn, for example, and the next day they would go to Mrs. Kilpatrick’s classroom. Teachers chose which skills in reading or math to have students work on, and whether the work would be a group or independent activity and had the freedom to reward student efforts by letting them play outside for the last ten minutes of the period.

Because Ms. Quinn and Mrs. Kilpatrick teach the same students, they work together and with the third team member, who teaches math, to make things as consistent as possible. In addition to setting up their classrooms identically, they follow agreed-upon procedures for class instruction, differentiation, and remediation. They carefully analyze and base their targeted instruction and remediation mainly on the data from the benchmarks. In her post-observation interview, Ms. Quinn explained,

[After each benchmark], each of the 5th grade teachers will individually call students back and tell them what they made overall . . . “so your overall projected level is a 3. So you’re doing really well in technology science, [and] social studies passages but we need to boost you in fiction and poetry.” [The benchmark report] actually it breaks it down into their projected level for each genre.”
The benchmark also provides information on Depth of Knowledge levels, which indicate the level of thinking required by the question. Ms. Quinn further explained in her post-observation interview,

I’ll look at [the Depth of Knowledge summary] and say, okay, this student is getting level -the first level pretty good but when you get up higher, they're not getting those highest order thinking questions.

Along with a traditional student score report, Case 21 reports a projection of how the student will perform on the state summative test. Schoolnet does not provide this prediction. Based on the student’s performance, Ms. Quinn reported, teachers determine groups and activities for the whole class as well as for individual students.

At Nolan Reid, the sixty fifth graders are divided into three teams. Each team represents a class that moves together from Ms. Quinn for English Language Arts (ELA) and social studies to Mrs. Kilpatrick for science and ELA, and the math teacher. Each of the three teams/classes has a name – Sharks, Lions, and Dragons, for example – and the benchmark progress of each class, shown as levels one through five, is charted on the walls of their classrooms. Additional notes by the teachers are added to the wall, such as, “Sharks are moving up!” Both classrooms communicate to students that teachers are interested in improvement, not just scores. Ms. Quinn and Mrs. Kilpatrick each explained in their pre-observation interviews that celebrating and supporting student progress was very important. As Mrs. Kilpatrick stated, “They know I don’t care if they got a 1, as long as they are moving up. If they are moving, that’s all I can ask.”
**Ms. Quinn: Experience with OFAPs.** At the time of the study, Ms. Quinn was completing her first year of teaching. She admitted that the first year was really hard, and that trying to teach both English Language Arts and social studies was difficult. In the pre-observation interview, Ms. Quinn reflected:

> I think every first teaching year is just kind - you just don't know what you're doing and you're trying different things, and then hopefully by your second year you've got some things under your belt and you forget about the things that didn’t work and move forward.

The district used Schoolnet during the first half of the year for benchmarks, administered every four-and-a-half weeks. Because there are so few computers available with Internet access, the benchmarks were created online in Schoolnet and then printed out for the students. Although Schoolnet offers immediate scoring, reports, and student feedback, these were not possible with the printed versions. Instead, the district used a scanning program to score the assessments and then import the scores so that they would be accessible to teachers and administrators. Despite the challenges of using an online program without computers, Ms. Quinn reported that Schoolnet was helpful in many ways for English Language Arts.

> It’s great because there are a variety of passages and for the most part [they] are vigorous and aligned to the Common Core. I had no problem with the types of passages . . . . In fact a lot of the passages I'd see have tied in very nicely with certain kinds of thing I need to teach because I teach language arts and social studies combined... so that’s one way Schoolnet has been very helpful.
Since the original purpose of Schoolnet was as a formative tool, each selection assesses only one standard so that teachers can administer small, single-standard assessments often, allowing for ease of targeted instruction. Assessing students on multiple standards, then, means additional selections. Unaware of this characteristic, when Ms. Quinn and her colleagues began creating a benchmark to cover four or five standards, the assessment had too many passages for the amount of time the students were allotted.

We would say ‘I would like 30 questions, five for each standard,’ [and] it would spit out this test that was so long - and it was too much - and the kids would have like 20 passages to read and only 30 questions. So my students - knowing my students and knowing the way my students - act, they're not going to be able to sit there and go through that.

Unfortunately, “that was for 75% of the year.” However, she said, “I love the passages. I loved the questions. For the most part they're aligned [to the Common Core State Standards for English Language Arts].” She further explained, in the pre-observation interview, the challenges of using an assessment with so many passages.

That did make it difficult for me to, you know, progress-monitor my students, and when it came to looking at the data, I knew that when it came to a regular assessment and there were 10-15 questions per passage they might do a little bit better. And they would go more in depth there not just scratching the surface of the passage.

Another frustration was the overlapping of selections and questions. During the first two quarters, Ms. Quinn would print out a passage and questions from Schoolnet for class practice. In the pre-observation interview, she reported,
Then when it came time or the summative assessment, they’d see the same passage, and the kids did really well. Well, that happened at least twice, so then I would think then we won’t use it [in class] because it's not going to give me accurate data.

Still another complaint Ms. Quinn had was the lack of language items in Schoolnet. Schoolnet has so much for [the Common Core] reading for information [standards] and so much for [the Common Core standards for reading] literature but there's nothing for [the Common Core] language [standards] - 400 or more items to choose from and nothing for language - and obviously [none for] speaking and listening [or] writing.

These frustrations led to her relief when the district changed to Case 21 for the third nine weeks benchmark.

I was so happy that we used it because you know it took maybe 5-6 passages - so it cut the number of the passages in half and it gave them aligned questions and there were, like five to ten per passage, so it was more realistic. The kids could read the passage and go more in-depth.

Seeking more help in formative assessment, Ms. Quinn discovered Mastery Connect, an online program that provides teacher-made assessments.

I’ve used Mastery Connect as way to look for formative assessments that other teachers have made. You can search by standard, by grade, and by subject. The reason I like Mastery Connect a lot is because you get a different variety.

Although students cannot work directly in Mastery Connect and assessments must be down-loaded and printed out, as Ms. Quinn says, “Without computers, it works well.”
also reported that she likes using printed copies of the assessment. “I want to give them the opportunity to write on the passage . . . so Mastery Connect is one way, early on, that I was finding formative assessments for my kids.”

However, the Schoolnet and Case 21 assessments available to her were still benchmarks instead of in-class formative assessments. Ms. Quinn explains, “So they (the district) were using it more as a summative benchmark, but we were using it as more of a formative [assessment].” Although the main purpose of a benchmark is to inform school and district administrators of student progress toward goals which are aligned to the state summative tests, the teachers at Nolan Reid Middle School analyzed the assessment data for each of their students and then used it to plan and implement targeted intervention. Black, et al. (2004) clarified that assessments designed for summative use could also be used for formative ones, as long as teachers applied formative strategies to them. Because students needed to take the benchmark on paper, the district followed a complicated process for creating, administering, and scoring the assessment that took two weeks or more. The district first ordered the benchmark, indicating the standards that needed to be assessed. Then Case 21 created the benchmark and sent it to the district. After administering it, the district sent the students’ assessments back to Case 21 for scoring, and then the scores were sent back to the district.

**Mrs. Quinn: Data analysis and planning.** When Ms. Quinn and her colleagues receive the benchmark scores from the district, which are given by the district several times during the school year, they examined them very carefully and then continued to review the results throughout the quarter. Studying the benchmark data was helpful to Ms. Quinn.
I would use that data and I would look at it and [the] Case 21 [assessment] actually helped me figure out how I am going to use the last 5-6 weeks of school to remediate, to reteach and to enrich those kids that don't need help in reading.

It became clear, as Ms. Quinn talked in her pre-observation interview, that after only a few months of teaching, she had learned the importance of asking the right kind of questions to elicit the information she needed for assessment.

One thing I did a lot of at the beginning of the year that I realize was not as helpful was giving [students] an exit ticket that was just asking them a very simple question and it wasn’t really getting them to show what they learned. [For example], if I ask my students, ‘What is the main idea?’ [or] ‘What does ‘main idea’ mean?’ and if they can rattle off what main idea means, that's great. But it's another thing if they can be given a paragraph or a passage and I tell them, ‘Tell me in your own words,’ or ‘Write down in your own words what the main idea of this passage or paragraph is.’

She realized that asking students to identify terms was very different from asking them to explain the application of the term to the reading passage.

When asked in the post-observation interview how she determines whether she needs to stop and reteach a concept or to move on, Ms. Quinn explained,

It depends on the skill that's being assessed. If it’s main idea or theme or inference . . . those are things that if we don't get [them] we can't really go on, but . . . that's not going to show up as much . . . Let's say we take a quiz and a lot of kids didn't do well on it and its on something as important as main idea . . . then we need to stop, and we need to really go over that quiz and I need to reassess them.
This constant assessment of what constitutes essential knowledge for students is especially necessary given the fact that Ms. Quinn must teach both Social Studies and English Language Arts during her ninety minutes of class time. She explained in her pre-observation interview:

Unfortunately, I haven't been able to focus much on language standards - I've taught them all, but I don't have time to focus on language much because I don't have time to do writing . . . reading is such a struggle.

As her experience has grown, so has her ability to recognize how long to spend on certain skills, she went on:

At the beginning of the year . . . I didn’t know how long to spend on one thing so when April came, I realized that, okay there are some things I need to go through quickly, but I am going to . . . make the most of the group review time without making it monotonous for them. But next year I’m really going to spend more time at the beginning of the year going over some of the skills . . . like main idea and theme - that is something you're teaching all year long and so, you know, hopefully by the end of the first nine weeks they’re going to be able to identify the theme of the passage. . . Now I know to make sure that all of those skills are being used all the time.

In her post-observation interview, Ms. Quinn reiterated that out of the 68 students she was teaching, Ms. Quinn estimated that ten were actually reading on grade level, and that this concern greatly influenced her planning and instruction.

They come to class with all this extra stuff from home and they're already behind. Language Arts has historically been their most challenging subject, and it’s very
difficult to get them interested in things sometimes. [Therefore], one of the biggest challenges I face is trying to scaffold for them because if I have to give them a text that is not on their level, we have to spend a lot more time with that text [and] I don’t necessarily have the time.

In the post interview, she reflected that a program such as Achieve 3000, which provides the same informational texts on multiple reading levels, would be very helpful, but, “We don’t have the technology [to access it].”

To prepare for targeted intervention and instruction, Ms. Quinn had divided the benchmark into skill sets and then recorded each student’s score on each of the skill sets. She was then able to quickly group students according to their proficiency as she moved from skill set to skill set in remediation.

_Mrs. Quinn: Targeted intervention and instruction._ The observation of Ms. Quinn’s classroom took place in late spring. For this reason, she was spending most class periods doing remediation and review. As she described below, she had learned by experience about the recursive nature of English Language Arts and therefore already knew some of the changes she would make the following year.

Next year, I’m really going to spend more time at the beginning of the year going over some of the skills [that keep coming up]. Now I know to make sure that all of those skills are being used all the time [so we don’t have to spend so much time on review]. I have a better idea of how you can’t teach [ELA] standard-by-standard [because] they are going to forget [what they learned earlier].
In the meantime, she reported, she uses her classroom’s group-friendly set-up to make the most of her targeted intervention. In fact, the only times students in Ms. Quinn’s class were not grouped by proficiency was during a unit test or something similar. “If it was just a regular class day, sometimes they’d be in pairs, sometimes they’d be in groups of three or groups of four and it would always be leveled,” she said.

At the time of the observation, however, these skills did need reviewing and re-teaching, and she had planned carefully to incorporate them in her targeted instruction. The week of the observation, Ms. Quinn was working with her students on informational text. On Monday, she had invited students who had made fours or fives on that section of the benchmark to sit at any table they chose, but they could sit at the same table. Once they were seated, students who had scored threes were invited to choose their tables, but there could be no more than two “three’s” at one able. The lowest scoring students sat in the semi-circle of seats at the front of the room. She described in the post-observation interview that she liked the fact that students saw that they could control where they sat to some extent, and therefore, worked hard to maintain that privilege. Ms. Quinn would then choose a standards master or leader from among the “fours and fives,” who would lead the group in the reading activity for the day. Students’ names were posted on the wall chart next to their scores on the most recent benchmark.

That’s where I've got all their names and they can look up there and see where they are and if they forget they can look up there and sometimes, it serves as a way to motivate them, like, 'I don't want to be a two anymore.'
During the pre-observation interview, Ms. Quinn explained that the activity was a continuation of review and targeted intervention in reading informational text. Ms. Quinn had already reviewed reading literature with students, and had worked to fill the gaps they had in understanding. During the post-observation interview, she said,

I think that the students do a lot better on informational text than literature because literature requires a lot of connecting to what you’re reading, but [for] informational text, the answer is right there – not all the time, but it is much more cut and dried . . . it’s a struggle to get them to appreciate literature.

The day of the observation, Ms. Quinn greeted students as they entered and informed certain students they would be sitting at the front. Other students found their places in groups according to the seating protocol Ms. Quinn used. Ms. Quinn began class with a mini-lesson and a short video on women’s suffrage that gave students background and context for the reading they would be doing in class. Then she went over ideas and vocabulary presented in the reading passage and reminded them of a strategy for answering questions they had worked on the day before. This strategy was to deconstruct the question before trying to answer it. Within groups, students could help each other. I examined the handout and found that it was one of two paired passages on women’s suffrage. Each passage was about a page in length and was followed by eight to ten questions, as well as some questions about both passages. Students received only the first passage during the class I observed. Although most of the questions were multiple choice, they were designed to send students back to the passage in order to find the answers. Most students read the passage independently, but they asked one another questions about problem words as needed.
Although there were some disruptions due to students who were off-task, most students worked to complete the assignment with little redirection from Ms. Quinn, who led the group of students in the semi-circle to do the same. However, making sure fifth graders stay on task while working in groups can be a challenge, and in the post-observation interview, Ms. Quinn talked about her frustration at having to stop and redirect students.

I feel really bad for the ones and twos who are up there with me because sometimes we only get through 2-3 questions. but it’s good to sit there because if we . . . really take our time and we look at it and see why we got it wrong and what we need to improve on I think it's going to help them more than just sitting in a group and [then]get behind or . . . lost.

With the students at the front, Miss Quinn worked with them to practice the question deconstruction strategy. She made a point of ensuring that the students knew what the question was asking them by asking questions like, “What is this question asking? What does it mean to infer?”

At the conclusion of class, students turned in their papers to Ms. Quinn. In the post-observation interview, Ms. Quinn explained that she would check over the responses of the students who had been in groups to be sure their answers were accurate, and that they had explained “their thought process to me in a way that help their answers make sense.” In this way, she could assess their understanding of the reading and their use of the question deconstruction strategy. The following day, they would review the passage and questions as a class. Then they would read a second of the two paired passages on women’s suffrage that also had questions. The second passage also included questions on both passages.
After completing these two passages and reviewing them with the class, Ms. Quinn explained in the post-observation interview, she planned to give them another set of paired passages and compare their performance on the first pair to that on the second, “to see how well they've improved.”

**Mrs. Quinn: The nature of formative assessment.** In the pre-observation interview, Ms. Quinn said, “I think a lot of people confuse the word formative with formal, and I think that's where the discrepancy is.” She acknowledged that at the beginning of the year, “formative assessment was difficult for me because I wasn't really sure what that looked like.” Her understanding grew as the year progressed. “My understanding is that formative assessment is a means of assessing a student, not necessarily for mastery but what they got from a particular lesson or what they need to work on before they are formally assessed for mastery,” and her use of formative assessment grew as well. “As I’ve grown throughout the year. I've learned different ways to collect data that I don't necessarily put into the computer [grade book] but that I can look at and say, okay this is how they're doing.” Ms. Quinn demonstrated this new skill as she led her small group through the lesson I observed. Although she did not plan to grade the performance of the students she worked with at the front, she made notes on her own passage of things she wanted to target the next day.

**Mrs. Kilpatrick: Experience with OFAPs.** Like Ms. Quinn, Mrs. Kilpatrick had experience with both Schoolnet and Case 21 benchmarks. Mrs. Kilpatrick often printed off quizzes from ClassScape to use with her lessons in science.
The tests are not always very accurate, and the kids would get mad because I would give them a unit test that also [covered the same material, and would ask] “Why do we have to have a unit test when all the stuff is on the benchmark?”

The problem, she said, is that the benchmarks did not have questions for material she covered in class, so she created her own unit assessments. Ms. Quinn also printed off formative assessment quizzes from Schoolnet to use “every day.” Since she was printing them, she copied them first into a Word document because “it wastes so much paper the way it's laid out, so usually I'll copy it and put it in a word document and make the font smaller.”

She scored these quizzes by hand or ran them through the Scantron scanner. She also regularly used her iPads to have her students play Kahoot, a game-based response system which allows students to compete in answering timed questions on a variety of topics. Every team which answers correctly during the time period earns a point.

At the time of the study, Mrs. Kilpatrick was in her second year at Nolan Reid, and she had been able to scrounge a second desktop computer and acquire, through a generous donation, five iPads for classroom use. These extra computers, combined with her laptop and the original desktop, gave her a total of seven, and enabled her to use them to review and reinforce vocabulary and content knowledge for group competition in Kahoot at the end of class and to give her information as to how well students had understood the lesson objectives. She also used other game-type systems that can be pulled up on iPads or cell phones to motivate students and allow them to practice.

Mrs. Kilpatrick liked Schoolnet in that “it is helpful is getting down to what is important in standards – how in depth are we supposed to go?” It reviewed the kinds of concepts that the
students would see on the End of Grade test in science, and allowed her to so do some
assessment in reading for information.

Mrs. Kilpatrick: Data analysis and planning. When this study took place, Mrs. Kilpatrick had also first used Schoolnet and then Case 21 benchmarks to analyze student progress toward reading goals. Since she taught science and was also responsible for the ELA Common Core State Standards for reading for information, she evaluated both areas. The biggest concern for her students was, she said, “Always – a lack vocabulary [knowledge].” She expressed concern about what a disadvantage a lack of vocabulary development would be for them in every area of their education. The cause of this lack, she reported in her post-observation interview, was not always clear. She asked, “Where’s the disconnect? Why is he such a low reader when his mom is graduating from nursing school?” Whatever the cause, she chose to make vocabulary the main reading focus in her classes.

Mrs. Kilpatrick conferenced with her students every two weeks to keep them on track and to communicate goals and progress toward those goals. In addition to tracking their quiz grades themselves, she also had her students keep track on an index card of anything on which they needed help during a unit of study. She had large charts on her walls showing student benchmark scores. “Parents are fine with [posting student scores], and it does create competition,” she said. Although the goal for everyone is at least 80% proficiency, she explained that the kids are well aware that it is most important that they continue to move up in reading and in science.

When asked in the pre-observation interview how she decided what skills to remediate and when to move on, she indicated that her decision was a judgment call.
You can’t just reteach to everyone – it’s a huge a waste time, but if not a single kid understood it then I know it's my fault. If it's just one or two [who understand], that's not going to work out, and I know I did something wrong. We go back and I will go back and I won't reteach it in the same way because obviously it didn't work, so there’s no point in doing it again [the same way]. Depending on how big a deal it is - like if it's an essential concept they have to know, then I'll reteach it and I'll spend a day on it. I'll go back to the most basic, and once we get halfway or a quarter of the way, and I can tell that half the kids get it and they need move on. Then they will start splitting up and [I’ll say] ‘You are here and this is what you do,’ and then I will be able to tell who does not get it at all.

Analyzing data for Mrs. Kilpatrick was not limited to the actual assessments she does, but also to everything she and her students do in class. In her post-observation interview, she explained, “For instance you see that kids were weak in something and you think, Oh, well, this next time when I teach them a new concept, I'm going to do it another way.” Mrs. Kilpatrick also explained knowing her students is critical to teaching them. When her students did not understand a concept,

I'll think about how I introduced it. Did I try to link it to what they already knew? Like a lot of the times, I'll link it to something they already understand. . . Because if you don't link it to something they already understand, nothing sticks. And sometimes you just get involved, and you’re just going and going, and this, and this, and you don't think of the kinds of things you have to do. You could have all these great ideas and that's awesome, but if they can't relate it to information they already
have, then they're not going to understand at all. That's why we were talking about breeding dogs yesterday [in talking about inherited traits], because breeding dogs is understood here, so they all get that.

She added further that putting concepts into context was important in all areas of her teaching. For example, she realized that students were having a very hard time with charts and graphs.

For most kids, that is really hard for them - interpreting data is really tough, so that's what we are doing next week. We're spending a whole day just interpreting data from questions that have a purpose, not just worksheets.

Reflecting on her students and her efforts to meet their needs, she explained in her post-observation interview that she analyzed how much she required of her students in the classroom. “Last year I was trying to do way too much in a class period, and I realized it was not going to work because they were not receiving the information,” she explained. Now, she said, the class period seems to be fast-paced because she has revised what she expects students to do in one class period, and now, the “classes are like ‘let’s go, let’s go!’”

Mrs. Kilpatrick: Targeted intervention and instruction. At the time of the observation, Mrs. Kilpatrick was working on skills she identified as problem areas on the third quarter benchmark. “I didn’t [remediate] the first benchmark, but the second benchmark I did. It was only 1-15 questions, so it was pretty easy to remediate.” She told me in the pre-observation interview that she now uses remediation regularly. On the day of the observation, students answered context clue questions based on hints from a short video they would watch on inherited traits. One of the weaknesses Mrs. Kilpatrick reported that her
students share is a lack of background knowledge about the world outside their county. Therefore, the context clues not only included words, but pictures as well. After going over the questions and watching the video, students reported to their assigned tables or desks.

Mrs. Kilpatrick gave each student two papers, printed on both sides. On each side of one paper were sixteen different pictures of traits dogs could inherit, such as tail length and coat color. The other paper contained short paragraphs followed by questions the students were to answer about inherited traits. This paper also included boxes students were to fill in.

Students worked together in their groups to complete the reading and the questions. During the observation, students used a chart on the wall to help them in their reading. This chart was a reading strategy, called RUBIES (R-Read and reread everything; U-Underline the question; B-Bracket the key words; I – Identify the key concept or pattern; E- Eliminate the wrong answers; S – Select the best answer) Mrs. Kilpatrick used to help students access texts and answer questions about them successfully. Also, like Ms. Quinn, Mrs. Kilpatrick required students to explain their answers to the questions they answered. In her post-observation interview, Mrs. Kilpatrick talked about realizing the need for using these and other strategies to help students in their reading. She said, “[Recently], their awareness of what they are learning has been a lot better.” When asked about this, she admitted,

Part of it was me, because I realized we were not reading to learn; we were just reading [to read]. At the beginning of the year we were reading passages, and we were doing one a week about whatever we were learning. We were going through it and annotating it, but we weren’t really reading it and taking notes. But then in February and March it was more like putting it into their hands.
She also began to ask questions as the students were reading, such as “What are you understanding from this text? What are they trying to teach you? What is this about? What’s the point? Why are we reading it?” Using these strategies, students began to think about what they were reading and became both more successful and more engaged.

In addition to using a variety of reading strategies and questioning techniques, like Ms. Quinn, Mrs. Kilpatrick used independent groups working together on an assignment while she worked with struggling students at the front. Before she and Ms. Quinn began using the seating arrangement and grouping they used during the study,

Usually [intervention] was in even smaller groups, and one thing I would do was like a lunch bunch, which was during lunch, and certain kids would have a date with me. We would sit and go through questions - it was usually just one question [per lunch period].

Analyzing reading data helped Mrs. Kilpatrick recognize the need for reading strategies and questioning techniques. Using her grouping protocol, she was able to conduct targeted intervention with struggling students while continuing to build skills of her other students.

*Mrs. Kilpatrick: The nature of formative assessment.* The use of a seating protocol to enable Mrs. Kilpatrick to more easily target misunderstandings and missing concepts was a practice that most students enjoyed. They liked the idea of making choices, sitting with their friends, and working in groups. Understandably, some students balked at having to sit with the teacher. Others, however, placed themselves at the front of the room.
In February they wouldn’t have any idea why they were grouped together. They wouldn't have really known that, but now they have more awareness [about] why they are in this group or why it would probably be good for them to be here.

Mrs. Kilpatrick, like Ms. Quinn, made no secret of why students need to be in the front. It was clear from the wall charts, the seating arrangements, and the explanations she used with her students that the purpose of school for them is to become proficient on the fifth grade standards. Some students needed more support than others, and she was there to help them be successful.

**Mrs. Jackson.** At the time of this study, Mrs. Jackson taught at Turtle Creek Middle School, an older one-story building with open-air passageways built around three central courtyards. Turtle Creek is located in a small suburban city in the central part of the state. It is a traditional middle school, serving grades six through eight and organized by teacher teams. During the time of the study, there were 740 students, of which almost half are Caucasian, with African-American and Hispanic students equally representing the other half. Class size averages twenty-three students. Turtle Creek furnishes all students with laptops which they carry with them throughout the day.

Mrs. Jackson’s classroom had a triple window on one side with a door to the hall on an adjacent wall. Thirty desks were arranged in rows varying in length from three to six to allow room for a teacher’s desk, file cabinets, bookshelves, two worktables, and other classroom amenities. Posters and reminders covered the walls, and assignments for the week were written on the double chalk-board at the front of the room. Between the chalk-board and student desks was a stool, a small table, and an overhead projector.
Mrs. Jackson was teaching two classes of regular ELA and two classes of advanced ELA when this study occurred. Each of these classes was seventy minutes in length.

*Mrs. Jackson: Experience with OFAPs.* Over the last few years, Mrs. Jackson had used several online formative assessment tools, including ClassScape, Achieve 3000, Quia, and Newela (Table 3). At the time of the study, she was regularly using Quia, Achieve 3000, and ClassScape in her classroom. Each one had strengths and weaknesses, so in using several of these programs, she was able to maximize their usefulness. However, the limitations of the programs, including “the amount of content [they provided], a teacher’s ability to customize based on students’ needs, [and the fact that] reporting functions [were] not always developed in a way that reflects how teachers need to access data were frustrating.

In her pre-observation interview, Mrs. Jackson explained the differences in the programs she used and what she liked about each one. The two she liked best and used most often were Achieve 3000 and ClassScape. Achieve 3000 was a program that provided informational text on multiple Lexile levels. This was helpful for differentiating, since an article on an archeology dig, for example, would be available at five different levels as determined by the Lexile readability scale. Therefore, all students could access the same information about the archeology dig. Achieve 3000 also nested their reading passages in complete lessons, including photographs, charts, and other graphics and a writing component. Mrs. Jackson particularly liked “more high-interest material delivered at a student’s Lexile [reading] level” and that Achieve 3000 was “a major help with differentiation . . . because of the Lexile settings.”
ClassScape, provided free by the state, gave teachers the choice of using prepackaged assessments or creating their own using passages and questions from its custom bank. The prepackaged assessments included two or three reading passages and ten questions that assessed a single standard. “I feel like ClassScape gives me a good picture of how students are performing in relation to the state standardized testing [and it] has a full range of literary selections,” she said in the pre-observation interview, but “not enough passages are available.” She was also frustrated that “you cannot access questions based on the high/medium/low designations that are developed as part of the ClassScape item-writing process” which made differentiation more difficult.

Because, like Mrs. Williams, Mrs. Jackson written and reviewed questions for ClassScape for several years, she had some insights about the development process. “I know the expectations for item writing intimately,” and writing and reviewing items for ClassScape had, she felt, helped “me understand how to ask good questions and to better incorporate the Common Core State Standards,” she said. In fact, in her post-observation interview, Mrs. Jackson explained that she had conducted several professional development sessions to train other teachers in asking good critical thinking questions and matching questions to the Common Core State standards based on the training she had received through her work with ClassScape.

Mrs. Jackson also used Quia, a program which allows teachers to enter their own passages, questions, and answers, and then design quizzes, games, or class activities using the tools in the program. Mrs. Jackson found Quia helpful when she wanted to assess students using a specific reading passage; for example, one from her class anthology.
In Mrs. Jackson’s pre-observation interview, she said, “I am seeing improvement in students’ Lexile levels (as measured through Achieve 3000) as well as improvements in performance on EOG-type questions through the use of ClassScape.” In addition, the immediate feedback from some programs allows her “use them to have students adjust their own understandings via correction activities.” As a result of using the programs, she continued, she can also “can make more specific adjustments, particularly in reteaching concepts as needed or knowing what terms students see in questions and don’t seem to understand.“ Mrs. Jackson is “seeing progress in most of my students, but the ones who are lacking progress are the same ones who tend to also neglect every other assignment.”

**Mrs. Jackson: Data analysis and planning.** When asked how she analyzed data and plans resulting remediation and instruction, Mrs. Jackson said, “It depends on what I am using [the assessment] for and what point we are at in the year. I do look for patterns in wrong answers to see what misunderstandings need to be addressed.” In fact, looking for patterns was a strategy Mrs. Jackson used for in all aspects of her teaching and interaction with students. “I look at trends of what students are missing and what wrong answers are showing up commonly as well as what they are doing well. I base subsequent teaching and assessments on this [information],” she said.

In addition to examining the data from the OFAP assessments and determining the areas in which the students were weakest, thinking about what caused students to perform poorly was key to Mrs. Jackson’s formative assessment strategy. A few years before, Mrs. Jackson reported,
I kind of did an experiment. When we had a longer class block, I gave [my students] a full-length practice test and I analyzed the pattern of their right and wrong answers . . . and I found that [on] the first third of the test, most kids did really well, not a lot missed; middle third, [not bad]; last third, [they bombed]. So then I gave them another one and totally switched out the order of the types of passages - same thing happened. So I thought, well, it’s not the passages. . . So my tact this year is, “What can I do to change their mental habits?”

During this analysis of students’ “mental habits,” she reported, “I’d been doing single selections review [data analysis] for every single answer they chose. You know, for eight of ten answers, you can do that.” She wondered how she could help students think more deeply about what the questions were actually asking and to form new habits in answering questions. For the next assessment, she decided to require her students to justify their answers to assessment questions.

And I was very specific when I gave them directions. I said, you cannot tell me “I guessed or “idk” [I don’t know]. . . I gave them examples of what I expected and what I was looking for and I gave them credit - say for eight questions they could get 16 points. So I said half the credit is going to come from doing the explanations. Even if your answer is wrong but you've attempted an explanation and it’s not one of those junk answers, you’ll get a point. I’ll give it to you. You’ll get a point. You won’t get a point for the answer, but you’ll at least get a point for that. And the first time I did it was kind of blah. But the second one they did was a lot better. And I thought, maybe they’ll buy into this idea.
After practicing this method for several single selections, Mrs. Jackson decided to transition to the next step:

[I told them to] go through the process - you don’t have to write [the justification] out . . . you try to do the process mentally and if you get the answer right, you don’t have to write [the explanation] out. If you get it wrong, then you have to figure out the correct answer from the ones remaining and explain your choice and I’ll give you credit based on whether or not you are [correct] - not just if the answer is right but for the explanation as well. So that went pretty well.

When she had considered the practice students had in justifying their answers, Mrs. Jackson’s next step in helping students form better “mental habits” was to remove still more of the scaffolding she had originally put in place. In her post-observation interview, she said that, in planning the activity I observed,

I decided to go with ‘here’s the correct answer,’ but I want them to - I’m most concerned that with the ones they missed, that they try to figure out the justification, because then that will help them correct their thought processes.

My observation of Mrs. Jackson’s two classes took place in the late spring. Therefore, she was cognizant of the fact that students would soon take the state-administered summative assessment in reading. At this point, she decided to combine assessment, analysis, and feedback into one activity. In planning the activity, she reported that she had four goals, summarized below:

• To give students more practice on reading both informational and literary passages.
• To give students the opportunity to practice for the strenuous summative reading test, which required that they read and answer multiple-choice questions for numerous long passages over several hours.

• To give students immediate feedback on their performance.

• To help students practice critical analysis of questions and ensure that they understood why one answer was the best one.

Although students at Turtle Creek Middle School had school-supplied laptops, the screens were small, so they must scroll to access an entire reading passage online. Since the selections and questions are shown on the same screen, and students must refer directly to the text to answer the questions, “it’s time-consuming and frustrating for the [them],” Mrs. Jackson explained. Turtle Creek has desktop computers in the media center but, as the number of online learning and assessment programs available to students grow, so does the demand for the media center computers. For this reason, Mrs. Jackson often prints off her assessments.

The assessment Mrs. Jackson used for my observation was one she created using the custom assessment tool in ClassScape, which allowed her to choose the standards she wished to assess and then reading passages and questions that were designed to assess them. At this point she could have had her students complete the assessment she created online. However, to ensure that students could easily see the entire passages as well as the questions, she created a pdf document. She made the pdf with the passages and questions available to her students via their laptops and in print form with room for answer justifications. Her plan was that students would come to her to check their answers, and that in this way, she could not
only immediately see how each student was performing, but she could also provide feedback, encouragement, and scaffolding, as needed.

*Mrs. Jackson: Targeted intervention and instruction.* Before introducing the assessment activity students would be working on during my observation, Mrs. Jackson collected homework and reviewed vocabulary often used in assessment. She gave each student a packet of twelve passages, or selections, with 93 questions that she had downloaded and printed from ClassScape. She explained that students would read the assessments on their computers and then answer the questions on their answer sheets. Mrs. Jackson assured students that they (she and the students) would take their time through the assessment, and that she had allowed several days to complete it. She went on to explain that one of her goals was for students to practice long periods of reading passages and answering questions and that another was to help them understand and practice how to think through the questions. As students completed one passage and its accompanying questions, they were to raise their hands and Mrs. Jackson would come to their desks to check their answers. If the answers were correct, students would move on to the next passage. If one or more answers were incorrect, Mrs. Jackson would give the student the correct answer(s), but the student would need to explain on the handout why the correct answer(s) was the best one. She emphasized that this process was extremely important, and they must be specific in their explanations. They would then raise their hands again, and Mrs. Jackson would check to see that their justification(s) demonstrated understanding. Although Mrs. Jackson told students that they could do the passages in any order they wished, most began with the first one and worked
through them in the order they were listed on the handout. Within a few minutes, student hands went up to indicate that they were ready to have their first questions checked.

When students had explained why the correct response was the best answer, they went on to the next passage. Mrs. Jackson checked these justifications when she checked the newly completed items and gave immediate feedback to students by confirming their reasoning or by conferring with them on how they might rethink the question or where they might look in the passage. Thus she was able to immediately correct their misconceptions and redirect them as needed.

*Mrs. Jackson: The nature of formative assessment.* The cyclical nature of formative assessment (Greenstein, 2010) is obvious in Mrs. Jackson’s activity. I observed that although the assessment was one which would take several days to complete, students received individual, immediate feedback on each selection and its questions as they were completed. Simultaneously, Mrs. Jackson conducted data analysis as she checked each student’s work and noted whole-class patterns in student misunderstanding. After checking only a few student answers on the first selection, she was able to tell that one question was posing a problem for several students. Not only were they missing the same question, but they were also choosing the same incorrect response. I observed that Mrs. Jackson noted this on her copy of the assessment. During the post-observation interview, she explained that she wanted to think about why students were consistently missing that question. “At first I thought it was because that question was more abstract than others, but then I realized that it was because they were oversimplifying [and] choosing an answer that was only partly true,” she said. As the class continued, Mrs. Jackson found more items in which students tended to
choose the same incorrect response. She realized that students were having trouble with summarization. I saw that she also noted this problem on her copy of the assessment. Later, she said, “We’ve done a lot on summary, [but] they are better at generating [summaries] than they are at analyzing the [summaries given as] answer choices.”

During the post-observation interview, while describing her students’ difficulty with certain questions, Mrs. Jackson expressed surprise that vocabulary was proving a challenge for some. “The most common questions I got asked were about vocabulary” she said. She went on to explain the amount of work she and students had done on vocabulary, studying roots, prefixes, and suffixes, and how they could identify the word’s part of speech.

In this last class, I was actually asked what ‘pity’ was [because it was a word students were asked to define in the activity]. I was really surprised when that question came up because I thought that one they would have known. But you never know. If they don’t have those skills and that knowledge they have no way to answer the question. They get hung up on it and I think some of the more advanced kids think, too, ‘If I don’t know exactly what it is, I can’t handle it.’ And I think, no, you can. Let’s act this out a little bit and find . . . the percent of the word that you do know. I’m hoping that since they’re doing [the activity] in small chunks that they [will think], ‘Okay, I got all those question but one.’ They’re seeing it as . . . they’re successful.

In the post-observation interview, I asked how Mrs. Jackson thought the activity went. She said that she was generally pleased, saying,

They’re doing pretty well on this. I had a number of kids - two or three kids - that had 3 out of 4 . . . 100 percent correct. I had a couple of my really high kids overthink a
few questions, and I think they can benefit from analyzing the answer choices [for the next passages on the activity] and figuring it out.

The assessment Mrs. Jackson gave her students on the day of the observation was purposely lengthy with 93 questions. As students finished each passage, they asked her to check their answers. Mrs. Jackson gave them the correct answers for any they missed and then students had to explain why that answer was the best one. In answer to the question of whether she would use this kind activity again with classes in future years, Mrs. Jackson said,

I think I would. I think I would. They were focused and I think it helped especially that I was just checking one selection at a time. I’ve tried this approach before, and I’ve always gone back to it because with every class you get a different set of commonalities - people in the class, what they know what they don’t know, and where they get stuck, so I think that idea of looking for the patterns is really helpful, and I will definitely do that again. I think that is the valuable data in data-driven instruction.

As a veteran teacher, Mrs. Jackson was able to combine multiple steps of the formative assessment cycle in one activity. In addition to assessment, Mrs. Jackson practiced data analysis as she checked student work and recognized patterns in their incorrect answer choices. As she checked answers, she also gave immediate feedback and by requiring that students explain why the correct answer was the best choice, she practiced a type of targeted intervention and reassessment. Mrs. Jackson also made several notes to herself about the patterns of incorrect responses students made in order to plan and implement the intervention needed to correct their misunderstandings.
As a result of conversations with the subjects and examination of sample reports and Internet documents, I was able to identify the major components of the main programs used by the teachers in this study. Detailed information about each program can be found in Chapter 2. On the following is a table illustrating the various assessment programs used by the study participants.
Table 5. *Assessment Programs Used by Teacher Participants (2012-2013)*

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Immediate Feedback/ Reports</th>
<th>Frequency, Content, and Administration Determination</th>
<th>Used by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achieve 3000</td>
<td>Yes</td>
<td>Teacher</td>
<td>Mrs. Jackson</td>
</tr>
<tr>
<td>Formative assessment in reading informational text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increase reading proficiency through practice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 21</td>
<td>No</td>
<td>School/district administrator</td>
<td>Ms. Quinn</td>
</tr>
<tr>
<td>Benchmarking</td>
<td></td>
<td></td>
<td>Mrs. Kilpatrick</td>
</tr>
<tr>
<td>ClassScape/ Schoolnet</td>
<td>Yes</td>
<td>Teacher</td>
<td>All four</td>
</tr>
<tr>
<td>Formative assessment in reading literature and informational text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benchmarking</td>
<td>Yes</td>
<td>School/district administrator</td>
<td>Ms. Quinn</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mastery Connect</td>
<td>Yes</td>
<td>Teacher</td>
<td>Ms. Quinn</td>
</tr>
<tr>
<td>Formative assessment in all subjects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Various, including formative assessment, games, practice (per teacher input)</td>
<td>Individual student score only</td>
<td>Teacher</td>
<td>Mrs. Jackson</td>
</tr>
<tr>
<td>Quia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Island</td>
<td>Yes</td>
<td>Student/Teacher</td>
<td>Mrs. Williams</td>
</tr>
<tr>
<td>Test preparation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Emergent Themes

To investigate the impact online formative assessment programs (OFAPs) have on the middle school English Language Arts teachers’ planning and reading instruction data analysis was conducted on three sources of data from each case: interview notes and transcripts, observational notes, and documents. As a result, the data were coded and then analyzed both within each case and across cases, revealing four common, emergent themes.

• Overcoming Obstacles
• Encouraging Metacognition
• Leveraging Experience
• Putting Students First

Overcoming obstacles. A recurring theme in this study was that teachers in this study had to overcome several obstacles to use OFAPs. These ranged from computer and Internet accessibility to lack of communication within the school and district. These obstacles prevented all four teachers from using OFAPs as often as they would have liked.

Computer and internet access. Ironic though it was for a study about the use of online programs, the overwhelming obstacle for all four of these teachers in using formative assessment was the lack of appropriate technology. Even though each student in Mrs. Jackson’s seventh graders at Turtle Creek Middle School and Mrs. Williams’s eighth graders at Edwards had school-furnished laptops, the design of the assessments for reading and the size of the laptop screens made them inadequate for long reading passages. Students had to scroll up and down to be able to access each passage when answering questions. Both teachers wished they could use the desktops in their media centers more often, but were
stymied because the overwhelming demand for computers by other classes who needed them to complete assignments and assessments. Thus Mrs. Williams reported in her pre-observation interview that she did a ClassScape assessment “six or seven times this year,” although she said, “I would do an assessment every two weeks if not more” if she could use the media computers as often as she wished. Her administrators also wanted teachers to administer OFAPs more often, but she explained that not many of the teachers at her school used them regularly, if at all.

For Ms. Quinn and Mrs. Kilpatrick at Nolan Reid Middle School, the lack of computers was critical. Although Mrs. Kilpatrick had five iPads, combined with her laptop and the two desktops in her room, she still had only eight computers for each class of twenty-three fifth-graders. Ms. Quinn fared even less well, with only her laptop and one desktop. Although the school had a computer lab, the computers were not connected to the Internet, so many of the programs Ms. Quinn and Mrs. Kilpatrick would have liked to use were out of reach. For them, completing individual assessments online was impossible, unless students took them over multiple days because of their school’s lack of computers with Internet capability. Even an assessment of twenty minutes would require four days of class time for Ms. Quinn. Asked during the post-observation what she would ask for if she could ask her district for anything at all, Ms. Quinn said,

More technology, because I wouldn't have to use paper. I could use clickers (student response systems) or if there was a computer that students could use [to take the assessment], put their answer[s] in, and email it to me, that would make life so much easier. So technology is the biggest thing - I really think that’s the biggest thing.
Looking toward her second year at Nolan Reid, Ms. Quinn was hopeful. “They say they’re in the budget, but I really don’t know,” she said. Her teammate, Mrs. Kilpatrick, agreed. “. . . finding markers – I can get markers, but finding a computer, ho-ho. Good luck with that.”

Although teachers print their assessments because of the absence of computer access, paper is also at a premium. In her post-observation interview, Mrs. Kilpatrick reflected, “The kids have figured it out. They don’t write on anything [I’ve printed]. They assume everything is a class set.” And this becomes even more of a problem when she must print out assessments.

[Schoolnet] wastes so much paper the way it's laid out [because one passage and one question take up an entire page online], so usually I'll copy it and put it in a Word document and make the font smaller and download the answer key . . . and make the words smaller.

In addition to the lack of computer access, using online systems created other problems, as well. Veteran teachers Mrs. Jackson and Mrs. Williams both had students whose privileges had been revoked due to misuse, whether mistreatment of the computer itself, or, more often, going to forbidden websites. While the consequence was one both teachers supported, it meant that they had to supply those students with paper copies whenever their classes did assessments or other activities using computers. As online use in schools becomes more prevalent, so do the issues with student misuse, and thus, the resulting consequences for teachers who want to use online programs to ensure the success of their students in a world where knowledge and experience with technology is essential.
**OFAPs disconnects and information gaps.** Besides the obstacles that online programs present in terms of computer access, another obstacle for the teachers was the OFAPs themselves. Although Edwards Middle School teacher Mrs. Williams expressed great satisfaction with ClassScape, even she was careful to examine questions to be sure they were clearly worded. At Turtle Creek Middle School, Mrs. Jackson used a variety of OFAPs to cobble together a system that she felt covered what she needed for her seventh graders. In her pre-observation interview, she described the Schoolnet features as “a mess,” and was frustrated that she could not access the difficulty level for questions in ClassScape, even though she knew, as a writer and reviewer of items, that this feature was there. She also wanted more available reading passages and would have liked a written component that would allow students to enter essays in response to reading passages that could then be automatically scored. Mrs. Jackson explained that Achieve 3000 “only focuses on informational text.” In her post-observation interview, she reported that although she liked the fact that students were required to write short essays, or constructed responses, to one question for each Achieve 3000 assessment, she had to score the response herself. However, she was unable to add the constructed response score to the rest of the assessment score in Achieve 3000, and thus did not have a record of the student’s overall performance on that assessment. Overall, she reported in her post-observation interview, although the programs were very helpful for students, she felt that many of the OFAP programs “didn’t work to help teachers.” She explained that she was often frustrated with how the programs’ features, especially reports, worked, and that having to go back and forth between programs was time-consuming.
In reviewing sample reports for Achieve 3000, ClassScape, and Case 21, I saw that each program presented student data in different ways. Achieve 3000 provided reports on how students perform on each assessment as well their growth in Lexiles, a measure of text readability. ClassScape and Schoolnet provided several reports for individual students and whole classes, including proficiency on standards, progress over time, and single assessment performance. Case 21 provided one report which listed the performance of each student on the benchmark as measured by percent of questions answered correctly, their projected proficiency level for the state’s end of grade summative test, and a suggested grade. Quia

For new teachers Ms. Quinn and Mrs. Kilpatrick, having access online for their students would have solved many of the issues they had with Schoolnet and Case 21. Although Schoolnet was free to schools, with free student accounts, neither was aware of its capabilities as a classroom formative assessment tool. Though both had used its quizzes in class, they had not been able to use the automated scoring, the instant feedback and reports, and the student portfolios. Also, there had been some mix-up with the test banks, whether through the district or the program itself, so that the passages and questions they had used were duplicated on benchmarks, negating much of the purpose of the benchmark. Ms. Quinn, who taught both English Language Arts and social studies, was also unaware of the reason the benchmarks had so many passages and questions, and said,

I would use Schoolnet sometimes when I was first teaching - the first nine weeks and the second nine weeks. I would select a passage and print it out for the kids to work on in class and then we'd go over it. Then when it came time for the summative assessment, they’d see the same passage, and the kids did really well. . . That
happened at least twice. So then I would think, well, why use it, because it's not going
to give me accurate data . . . so like I said I love the passages, I loved the questions -
for the most part they're aligned - but just the way it was being used it wasn't good.

Her co-worker, Mrs. Kilpatrick, too, was frustrated with the disconnect between the
Schoolnet or Case 21 benchmarks given periodically by the district and what she was
teaching.

The tests are not always very accurate, and the kids would get mad because I would
give them a unit test that also [covered the material]. They would say ‘Why do we
have to have a unit test when all the stuff is on the benchmark?’ and I [would say],
‘It's not on there.’

Still another concern for Ms. Quinn about Schoolnet was the lack of questions on the
language standards; “400 or more items to choose from and nothing for language,”
Ms. Quinn mused. However, she would have liked being able to use it for formative
assessment.

Had our school known and had I known as a teacher that Schoolnet was supposed to
be used for formative assessment and I should be reserving time at the beginning of
class - like ten, fifteen minutes to do that - then it would have been very helpful.

Since Schoolnet was available free to every student, I asked Mrs. Kilpatrick if she
ever used the IPads for Schoolnet. She said, “I don’t know how the kids get into it.” I also
asked her if she had ever used clickers with Schoolnet. She explained that she was unaware
that Schoolnet could be used with clickers, and said, “We should have known this.” She also
described the difficulty in borrowing items from their media center.
I can usually get things . . . but it’s like pulling teeth . . . and one minute later [the media specialist] is at the door to get it back . . . to keep track of it. It’s not going to get stolen. I’ve had five IPads in here since March and I have never had anything stolen. So in the meantime, no one gets to use [media materials].

Overall, both Ms. Quinn and Mrs. Kilpatrick were frustrated by their lack of knowledge about the programs available to them and the lack of training they had received on their use. Mrs. Kilpatrick said, “There hasn't been anything [to teach us about Schoolnet except] there was a [state-provided] wiki. People get upset at training because we're asking all these questions, and I want to say, ‘Chill out!’ [because we need this information].”

**Poverty and culture.** Finally, for at least three of the teachers, the poverty and/or culture of their school settings proved challenging. Mrs. Williams expressed frustration that her English Language Learners could not access assessments that were on their reading levels. Her school had a growing Hispanic population and students for whom English is a second language are required to take the End of Grade test in reading after one year in the United States. Having assessments on their reading levels would allow her to gradually increase their skills. For Ms. Quinn and Mrs. Kilpatrick, their students’ lack of knowledge of the “outside world,” their lack of vocabulary, and their overall low reading proficiency make every lesson more difficult.

However, In spite of the lack of computers, computer access, support materials, accurate and effective programs, and the lack of training and information, these four teachers used OFAPs because they saw the value in assessing students, analyzing data, giving feedback, creating and implementing targeted intervention, and then starting all over again.
Totally absent from any of the conversations I had with them was any mention of these activities taking up too much time. In fact, none of the teachers ever discussed the time it took to download, resize, print, or score the OFAPs. Instead, they were focused on how to make them as effective as possible for their students.

**Encouraging metacognition.** Throughout this study, it was obvious that the needs of students were the main concern of all four teacher participants. To help students to reach their full potential, students must become metacognitive (Pintrich, 2002). One of the effects of formative assessment is an increase in students’ metacognition (Heritage, 2007), but all four of the teacher participants in this study consciously worked at developing these skills in their students. According to Pintrich (2002), there are at least four characteristics of metacognition:

- knowledge of general strategies that might be used for different tasks
- knowledge of the conditions under which these strategies might be used
- knowledge of the extent to which the strategies are effective
- knowledge of self (p. 219).

Each of the teachers in this study used methods to encourage metacognition and self-regulation in students that went beyond giving them feedback on their assessment performance. Mrs. Williams’s emphasis throughout the formative assessment cycle was to help students understand how to analyze questions and answers themselves by examining the text, making connections to other texts and their own experience, and eliminating answers that for one reason or another were not the best answers. This kind of self-talk helps students move from being learners dependent on teachers for feedback and direction to learners who
are independent. In her pre-observation interview, Mrs. Williams reported that during assessments of her eighth graders, she circulated throughout the class, checking on students’ work and giving them immediate feedback and support:

And then they show me their score and I say okay, which ones did you miss?’ and ‘I need you to read that question again.’ Then they do their analysis and write [the analysis] down so I can look at it again. It takes longer but it’s so much better . . . and what I have them do [is] to go in and see the wrong answer and the right answer . . . I have them take their time and work on it but then . . . where the real learning takes place is when they go back and [determine] why the right answer is the right answer. ‘What did you miss about the question?’ [I] have them do an analysis. ‘What did I miss? Why did I miss this?’ And I have them try to figure out why. ‘Oh, I didn’t read the question carefully enough.’

Here she helped students move toward both the first and second of Pintrich’s (2002) metacognitive characteristics, which strategies can be used and when each fits best. Once assessments were completed, and she was going over the scored assessments with students, she again emphasized analysis of questions and answers, saying, “Take a look at the ones you missed. Were you surprised at any of the ones you missed? Take a look to see if these were the ones you had trouble with when you were taking it.” Thus she guided students to begin using the strategies of self-talk and questioning on their own. In addition, she continually asked such questions to help them analyze their own thinking:

- Why is A correct?
- Why is it C?
• What helped you in finding the right answer?
• What thought process did you learn?
• What word is so important in D?
• What’s hard about this question? What’s harder about the answers?
• When people got it wrong, they picked A – why is A wrong?

At the end of class, I observed Mrs. Williams invite students to respond orally to the activity they had just completed: “Reflections on this – steps you used, processes - what worked?” She also had them write these reflections in their journals, which she described in her post-observation interview.

They keep track of their scores throughout the year – and major grades – and they do a reflection on how they are doing with ClassScape – and what they did well – why one test may have been harder than others and what types of questions they missed and all that. [They do an overall reflection] right after each assessment.

Using these strategies helps students learn the kinds of questions they must use to ask themselves in order to become metacognitive learners. Here, Mrs. Williams helped students learn how effective their strategies were working, the third characteristic, knowledge of the extent to which the strategies are effective

Ms. Quinn also worked with her fifth-graders on strategies for answering questions, the first characteristic on the Pintrich list. The day of the observation, she worked with students at the front of the room on how to deconstruct a question so that they could fully understand what it was asking. After class, she planned to look over the work of the students working in groups independently.
I'm going to determine if the kids who are [at the tables] were going to be able to answer the question correctly, use the strategy correctly and explain their answer – their thought process to me – in a way that helps their answer make sense, and I'm going to grade theirs for accuracy.

The characteristics of metacognition also apply to behavior. For example, in Ms. Quinn’s classroom, the overarching goal for students is to move up from whatever score they received on their latest benchmark. In order to do this, they had to understand the goal and then recognize what steps to take to achieve that goal. In her post-observation interview, Ms. Quinn pointed to the chart on her classroom wall.

That’s where I've got all their names and they can look up there and see where they are and if they forget they can look up there and sometimes, it serves as a way to motivate them, like, "I don't want to be a two anymore."

This recognition that the student can exert some control over a situation is a beginning of Pintrich’s (2002) fourth characteristic, a knowledge of self. At first, Ms. Quinn seated students at tables based on a specific plan of balancing strong students with weaker ones.

In the beginning I was not letting them choose their seats because I wanted to see how well they worked with their groups. When I saw after a couple of weeks that they were doing pretty good with the seats I was giving them, I gave them the option to start sitting where they want[ed] to sit. I think that encourages them to do better because they are like, ‘Oh, I can control where I go.’ I hope they make good wise choice. If they don't make a good choice I'll sit them with somebody else. If they're really not working well with anybody I'll give them the option to sit by themselves or
I'll just have to tell them they have go sit by [themselves] and cool down. In the beginning it was consistent but then about the third week of remediation I started to let them choose - so now it's been about three weeks that I've been letting them choose their seat.

Her reasoning was that by giving students choices, they begin to take responsibility for the consequences of their choices. However, she also made it plain to them that she was actively working to support their growth. When she conferenced with students about their progress on the latest benchmark, she explained:

[I] will individually call students back and tell them what they made overall. [For example, ‘So your overall projected level is a 3. So you’re doing really well in technology, science, and social studies passages, but we need to boost you in fiction and poetry.’

The use of the words, “We need to boost you” was important imagery, as it indicated at least two people working together. She further supported students in their growing knowledge of self through encouragement. At the beginning of the class observation, Ms. Quinn greeted students at the door and sent the ones who were struggling most to the front to work directly with her. Upon being told he was to sit in the semi-circle, one young boy balked, saying he wanted to sit at a table with his friends. Ms. Quinn put her hand on his shoulder and leaned down to look into his eyes, and said, “But I really want to give you extra support, and then you’ll be ready to sit at one of the tables.” The communication of concern and care for him, along with the overall classroom culture of both acceptance of one another’s differences and
of working together for the improvement of all, helped this young man join the semi-circle ready to work.

Mrs. Kilpatrick also helped students toward metacognition in several ways. She taught students general reading strategies, such as RUBIES (read and reread, underline the questions, bracket the key words, identify the key concept or patter, eliminate the wrong answers, and select the best answer) and required that they use RUBIES whenever they read informational text. Strategies such as RUBIES enabled students to learn what to do on their own, thereby moving them toward metacognition (Heller, 1986). Requiring students to “always explain their answers” was another strategy that she, along with the other three teachers, used to help students become more metacognitive. As they became accustomed to having to explain their answers to their teachers, they learned the strategy of always asking themselves why one answer was the best choice.

Mrs. Kilpatrick also helped her fifth grade students begin learning about self by showing them how their behavior affected their work, which resulted in limited choices - in not being able to either sit with their friends, not being able to choose the table at which they sat, having to sit up at the front, lack of freedom, having their name beside a one or a two on the wall, letting them choose –either today or tomorrow.

Mrs. Kilpatrick, like Ms. Quinn, made no secret of why students needed to be in the front semicircle to work with her. It was clear from the wall charts, the seating arrangements, and the explanations she used with her students that the purpose of school for them was to become proficient on the fifth grade standards. When Mrs. Kilpatrick began using the semi-circle,
The kids had no clue most of the time. In February they wouldn’t have any idea why they were grouped together. They wouldn't have really known that, but now they have more of an awareness [as to] why they are in this group or why it would probably be good for them to be here. . . . Yesterday I gave two of the kids a choice: ‘Do you want to do this today or tomorrow? You’re going to be doing it no matter what, so do you think you can work with these people [or would you rather wait until tomorrow]?’ That's more behavior than data but sometimes you have to be like that, because some of the kids can’t - it's not going to work out. Just like some adults can't work together, some kids can't be together. But their awareness of what they are learning has been a lot better.

After working in groups for two months, she said, “Now they have more awareness [about] why they are in this group or why it would probably be good for them to be here.” In fact, students were beginning to “One of the boys [came into class] yesterday, and he did NOT understand, and he said, “I need to be in the front. I do not understand.” I [said] “Go [to the front].” She went on to say, “When they can self-identify, that's great.”

This new knowledge of self, another category of metacognition, also helped students understand their strengths. Mrs. Kilpatrick also told the story of a student who had struggled in science.

We have a kid who's in AIG (academically and intellectually gifted), and science is really tough for him. He hasn't done well all year, but he's an AIG kid, and so he’s in the AIG class, but he's not in the group with the AIG kids, ever, because he's not [at that level in science]. Yesterday was the first time he was in that group, and he [said],
“I’m going to explain this,” and he went up there and explained. . . I got my phone and I videoed him and he watched himself, and I said, "See? You understand, you understand! You explained it exactly how I explained it to you, even better in some ways. You know what you are talking about!"

Mrs. Jackson consciously decided to help her students move toward better “mental habits.” She used scaffolding to help her students reach a certain level of metacognition. At first, she required her students to explain every answer to every assessment question. As time went on, however, and they became more skilled at deconstructing questions and analyzing text, she asked them to write an explanation when they were correcting wrong answers. Finally, during the lesson I observed, she removed much of that scaffolding. For that assessment, if students missed a question, she gave them the correct answer but required them to explain why it was the correct one. This concentration on the reason an answer was correct rather than on just getting the correct answer helped students move toward metacognition by practicing the strategies of deconstruction and analysis. “I’m most concerned that with the ones they missed, that they try to figure out the justification, because then that will help them correct their thought processes.” The analysis did not only benefit students who normally struggled. “I had a couple of my really high kids overthink a few questions, and I think they can benefit from analyzing the answer choices and figuring it out.”

A positive side effect of this practice was that the fewer questions students missed the first time, the less work they had to do. Therefore, they began to realize that taking their time
and looking back at the text to ensure that they answered correctly the first time prevented them from having to explain later why the correct answer was the best choice.

One of the strategies Mrs. Jackson taught her students was to look at what they did know to help them figure out what they did not know. Heller (1986) explains that one effective metacognitive strategy is for students to think about what they know and what they don’t know. This strategy is often used with students to introduce new units of study and is more commonly known as the KWL chart. Students list what they Know, what they Want to know, and what they Learned. In the post-observation interview, Mrs. Jackson described how she used this strategy with her students. In teaching vocabulary, she said, she would tell students to look at the part of the word they did know and then use that knowledge to help them figure out the rest of the word. During the observation, Mrs. Jackson expressed to me her hope that students applied this strategy to the long assessment she designed. “I’m hoping that since they’re doing [the activity] in small chunks that they [will think], ‘Okay, I got all those question but one.’ They’re seeing it as . . . they’re successful,” she said.

Formative assessment experts (e.g. Black & Wiliam, 1998a; Heritage, 2007) agree that metacognition is the overall goal of formative assessment. In addition to giving descriptive feedback to students, encouraging them to perform self-assessment is also important to move them toward meta-cognition. The four teachers in this study consciously used strategies such self-analysis, questioning, and using parts-to-whole strategies that moved students toward self-assessment and metacognition.

**Leveraging experience.** Based on the interviews with these four teachers and my observations of their classes, all were cognizant of the formative assessment model and of its
power in helping students learn. In using the OFAPs, each of the teacher participants used all components of the typical formative assessment cycle: setting learning goals, instruction, assessment, data analysis, descriptive feedback, and targeted intervention/instruction (e.g. Black & Wiliam, 1998b; Stiggins, 2002; Greenstein, 2010). However, experience in both teaching and with formative assessment resulted in varying approaches among them.

At the beginning of her first year of teaching, Ms. Quinn was not sure what was meant by formative assessment: “I wasn't really sure what that looked like.” Six months later, during the pre-observation interview, she clearly articulated that “formative assessment is a means of assessing a student, not necessarily for mastery but what they got from a particular lesson or what they need to work on before they are formally assessed for mastery.” She also was able to apply her understanding to the classroom. “I've learned different ways to collect data that I don't necessarily put into the computer [grade book] but that I can look at and say, okay this is how they're doing.”

Ms. Quinn’s frustration with having to use the district benchmark for formative purposes resulted from of a combination of factors, which she explained in the pre-observation interview. First, she did not have access to Internet-connected computers in order to use an online formative assessment program in classroom. This meant she was dependent on a district benchmark, covering four and a half to nine weeks of instruction, for OFAP formative assessment. Second, the district used Schoolnet to design their benchmarks. Schoolnet was originally designed to be used as a formative tool, and as such, each passage could only be used to assess one standard. The intent was to allow teachers to thoroughly assess small instructional units in order to more specifically use targeted intervention. This
meant that benchmarks designed to assess several standards had multiple passages, requiring students to spend extra time reading passages. She said, “The students would have, like 20 passages to read and 30 questions. Knowing my students and knowing the way my students act, they're not going to be able to sit there and go through [such a long assessment].” Also, although Schoolnet had separate question and reading passage banks for benchmarks and formative assessments, a mix-up had evidently occurred, resulting in duplication of the passages which Ms., Quinn printed off for use in her classroom and the ones that appeared on the district benchmark.

By the beginning of the fourth quarter of her first year, Ms. Quinn was able to use the Case 21 benchmark to map out her targeted intervention for the quarter. “[The] Case 21 [assessment] actually helped me figure out how I am going to use the last 5-6 weeks of school to remediate,” she said in her pre-observation interview, because, she explained, she was able to clearly see the skills in which students were weak as well as which students needed more help. Since the Case 21 benchmark was shorter than the Schoolnet ones, she thought it was more representative of her students’ actual proficiency. Also, her experience with benchmarks and her growing understanding of formative assessment allowed her to make more effective use of the Case 21 benchmark data. For this time period, Ms. Quinn used the formative assessment model using a step-by-step method on two levels. First, she used the benchmark to assess, analyze data, give feedback, and to conduct targeted intervention. Reassessment for this first level would be the state summative assessment. However, within this overall formative assessment model, she also applied the model to her everyday plans that grew out of the Case 21 benchmark analysis.
For the first level formative assessment using the Case 21 benchmark data, Ms. Quinn carefully analyzed the Case 21 data and then gave descriptive feedback to her students by way of individual conferences, during which she showed them their scores for each type of reading and explained what the scores mean. Along with percentages of correct answers, Case 21 includes a projection for how well students are expected to perform on the state summative assessment. She posted these projections on the walls of her classroom to inform and motivate students. Since so few of Ms. Quinn’s students were reading on grade level, she planned remediation for all of the reading skills included on the assessment, using her system of group work. For each skill level, some students would almost always be at the front of the room for focused instruction, but some would rotate in or out, depending on their developing proficiency on that particular skill. Because she chose to use informational passages based on the social studies standards, her plans allowed her to use targeted intervention while moving her class forward in social studies.

For her second level of formative assessment using the Case 21 benchmark data, Ms. Quinn used the data on her students’ proficiency in reading informational passages. She then used the article and questions on women’s suffrage to assess the grouped students’ use of the question strategy while giving them practice in using it. For the group at the front, she not only was able to specifically target and remediate, but at the same time, she assessed and analyzed her students’ understanding of the strategy she was teaching them. This ability to combine several components of formative assessment demonstrates Ms. Quinn’s growing skill in formative assessment. “We really take our time and we look at it and see why we got it wrong and what we need to improve on.” Ms. Quinn’s ongoing requirement that students
explain their thought process in answering each question meant that she could easily analyze the kind of interventions she needed to do. The following day she gave written individual feedback as well as whole-class feedback and any necessary remediation. Ms. Quinn then planned to give students a second set of the paired selections and questions and compare these to “to see how well they’ve improved.”

As Ms. Quinn’s experience grew, so did her ability to recognize how long to spend on certain skills. Her experience also helped her in making plans for future instruction, as evidenced in her post-observation interview:

At the beginning of the year . . . I didn’t know how long to spend on one thing . . . But next year I’m really going to spend more time at the beginning of the year going over some of the skills . . . Now I know to make sure that all of those skills are being used all the time . . . it would be pie in the sky for most teachers . . . to get to spring break [and not have to review everything] instead of going back and pick[ing] up those skills.

Throughout the formative assessment process, Ms. Quinn completed the steps in the model one at a time, “Nesting” the models – using both the long-term analysis and intervention at the same time she used the smaller model with her informational text, allowed Ms. Quinn to specifically target student strengths and weaknesses for each type of skill.

Ms. Quinn’s colleague and teammate Mrs. Kilpatrick had many of the same challenges as Ms. Quinn in implementing any type online formative assessment, including lack of enough Internet-connected computers, lack of knowledge of what the various programs could do and their intended use, and using a district benchmark as the OFAP
assessment. However, although Mrs. Kilpatrick was only in her second year of teaching, she had graduated with a degree in education and was licensed in both elementary education and English for speakers of other languages (ESOL) Mrs. Kilpatrick was also frustrated by the content of the district benchmarks from Schoolnet as well as Case 21, explaining in her post-observation interview that the assessments were “always very accurate.” For this reason, she retested students with her own unit tests. At the time of this study, Mrs. Kilpatrick was using most of her class time to review for the state End of Grade test. For that reason, she was using the data she had received from the Case 21benchmark as an overall structure for her lesson plans. As did Ms. Quinn, Mrs. Kilpatrick displayed the projected scores for the End of Grade test on her wall, along with motivational slogans. She analyzed the data and individually met with each student to review their results from the district benchmarks. She used the same grouping protocol as Ms. Quinn for remediation. She too, used the step-by-step formative assessment model, although components of it were seen in all aspects of her teaching.

Unlike Ms. Quinn, Mrs. Kilpatrick used quizzes from ClassScape on a daily basis and every two weeks met with each of her students individually to review their progress. Thus, although she used the Schoolnet and Case 21 benchmarks for a broad view of where her students were, she also used smaller assessments for more detailed information. Her experience with students in her school had helped Mrs. Kilpatrick to understand that how material is presented can make the difference between a lesson’s success and failure. She made this point in her post-observation interview, by explaining why students seemed to understand more in the latter part of the year, she said, “Part of it was me, because I realized
we were not reading to learn . . . but we weren’t really reading it and taking notes.” Like Ms. Quinn, improving her questioning skills also helped make a difference. When a lesson just did not work, Mrs. Kilpatrick reflected on why.

. . . If not a single kid understood it, then I know it's my fault. If it's just one or two [who understand], that's not going to work out, and I know I did something wrong. I will go back and I won't reteach it in the same way because obviously it didn't work, so there’s no point in doing it again [the same way]. In analyzing what went wrong, I'll think about how I introduced it. Did I try to link it to what they already knew? Like a lot of the times, I'll link it to something they already understand. . . Because if you don't link it to something they already understand, nothing sticks.

Mrs. Kilpatrick also used data extensively. She liked the program JumpRope.com to help her track student progress. This program allows teachers to generate reports giving information based specifically on standards. Mrs. Kilpatrick’s students also kept a record of their own quiz data as well as tracking self-identified weaknesses. Students who were missing concepts were pulled into groups of three to five for remediation. She also held lunch-time meetings with students to address a specific skill, often “a graph or data table because for most kids . . . that is really hard for them.”

Just as Ms. Quinn had already begun by the end of her first year to understand how much remediation needed to take place throughout the year, Mrs. Kilpatrick hoped to be more consistent in her use of remediation and targeted intervention.

I want to be better at [creating a] schedule so that the kids know when remediation is coming. . . A lot of times it's kind of disguised so they don't really know what's
happening anyway. But I know the kids thrive on schedules because . . . they do so much better when they know what’s going to happen.

In her second year at Nolan Reid, Mrs. Kilpatrick has been able to use her experience to her advantage in several ways. A generous parent had donated five IPads to her classroom and had also found an extra desktop computer. She was therefore able to use her computers to play short group games to reinforce concepts. Also, Mrs. Kilpatrick’s use of the printed Schoolnet formative assessments gave her immediate information on her students’ progress toward the instructional goals she had set. She was able to use the district benchmark as an overall assessment to guide her long-term goals and then to adjust instruction as she went along, based on her printed quizzes. Thus she has incorporated more of the true formative assessment model into her classroom, in spite of not being able to use the assessments online.

In her tenth year of teaching, Mrs. Williams was able to use several components of the formative assessment model simultaneously. She used data analysis, feedback, and targeted intervention as students were still being assessed.

I know you’re supposed to study the data and then go back . . . But I feel like my most effective time is when they’re taking it and I can work with them one-on-one. If they have trouble while they’re doing it, I usually will go and talk with them, because I think it’s helpful. And then they show me their score and I say okay, which ones did you miss?” and “I need you to read that question again.”

When going over the latest assessment, Mrs. Williams’s experience in assessment was clear as she listened to student answers and their descriptions of their thinking processes. Although she had already conducted data analysis on the assessments from the previous day,
she still assessed students and collected data as she interacted with her students. She gave them immediate and descriptive feedback as she asked probing questions and exploded with, “Right!” when a student made a particularly astute observation or connection. She guided them toward meta-cognition and self-regulation when she required them to explain their reasoning for the way they answered a question. Mrs. Williams’s experience in formative assessment, in teaching, and in working with students was obvious as she used many strategies at once.

Mrs. Jackson, the most experienced teacher, easily integrated the formative assessment components. First, she set four short-term goals for the activity she used during the classroom observation:

- To give students more practice in reading both informational and literary passages.
- To give students the opportunity to practice for the strenuous summative reading test.
- To give students immediate feedback on their performance.
- To help students practice critical analysis of questions and ensure that they understood why one answer was the best one.

Then she planned to assess, analyze data, give feedback and provide targeted intervention/instruction, and then assess again, all in the activity I observed. In giving students directions for the activity, she communicated these goals to them, an important part of the formative assessment model, and one which is necessary for developing student self-regulation and metacognition. Like Mrs. Williams, Mrs. Jackson’s ability to combine these strategies in her interaction with students was a clear indication of her experience. As students completed a single reading selection and questions, Mrs. Jackson checked their
answers, not just to show them what they missed, but also to see for herself how they were doing (assessment and data analysis). For any incorrect answers, students were given the correct answer, but required to explain why it was the correct/best one, which again, she checked (assessment, data analysis, feedback). If the student was still having trouble, she discussed the question with the student (assessment, data analysis, feedback, targeted intervention). Finally, while she was checking individual answers, Mrs. Jackson was watching for patterns of missed questions and misunderstanding (e.g. summarization and overthinking questions) across the class (data analysis) and planning how she would use targeted intervention to remediate these issues. Using her experience in questioning, assessment, feedback, and student metacognition allowed Mrs. Jackson to combine formative assessment model components in an efficient way and also reflects how the model can become an integral part of the instruction.

As each of these teachers applied what they knew and understood about formative assessment, of course their overall experience as educators informed their use of it in their classrooms. Though the more experienced teachers were able to incorporate formative assessment more seamlessly into their instruction, the two newer teachers still effectively applied the principals of formative assessment to increase student learning.

**Putting students first.** No matter the challenges of their jobs, all four of these teacher participants were focused on putting students first. One of the constraints in OFAPS is the popular assumption that these programs are designed to do the job of the teachers and to make it easier for them. These four teachers made it clear that this perception is far from the truth. These teachers were not ever looking for less work, but in fact, were looking for
answers to their questions about how to best help students become proficient in reading. In doing so, they not only researched programs, collaborated with colleagues, and tested theories in class, but they used the process Pope (1999) describes as reflection and refraction. In reflection and refraction, teachers first reflect by carefully looking back on their practice as a whole while asking themselves questions about what worked, what choices were made, and what resulted. Then teachers move on to refraction, in which they turn this reflective lens to challenge their initial reflection to discover new insights, keeping the students at the center of the process. These four teachers used reflection and refraction to discover new ways to be more effective teachers and to help students learn. In doing so, they made students their priority.

First-year teacher Ms. Quinn was constantly seeking out new programs and information that might help her students. In addition to Mastery Connect, a program she learned about from other teachers and which allows teachers to upload their own formative assessment to share with other teachers, which she had heard about from other teachers, she was investigating a Gates McGinnis program (that tests students on books they had read), and SAS Curriculum Pathways (which provides content-rich lessons with short assessments), for the next school year to give her students more resources. At the same time, she had decided that using Accelerated Reader (another program which tests students on books they have read), was not helping her students in the ways she felt was important. Although her school had bought the program, Ms. Quinn thought there was perhaps a program that was a better fit for her students’ needs.
Student needs were a consideration for her teammate Mrs. Kilpatrick when she used rewards to encourage students while enhancing learning. At the end of the class I observed, she used her IPads and Kahoot (a game program gives students a set time, between five seconds and two minutes, to answer questions on a given topic) to reward her students for their attention and work during class and to reinforce concepts they had learned. While students were playing a fun, competitive game, they were also practicing the skills Mrs. Kilpatrick had introduced during class. She understood that time was too important to waste. She also actively searched for programs to enhance her students’ education, and further, both she and Ms. Quinn spent time outside of school to make connections with their students, meeting them on weekends for hikes, participating in 5Ks with them, and attending their sports and community activities. “The kids have to know you care. If they don’t see you [outside of school], they don’t care [about you in school.] They think, ‘Oh, she'll just leave.’”

Ms. Quinn also struggled to overcome her students’ low self-esteem and expressed determination to help them feel good about themselves by helping them to become better readers and learners.

They come in with this preconceived notion that I’m not going to do well, I'm not going to like it, this is boring, so why should I care. So I've had to struggle with that. I still struggle with that . . . , but hopefully, I’ve sparked a love a reading [in] some kids.

Student self-esteem was a concern for Mrs. Kilpatrick as well, yet she recognized that sometimes, their behavior was perhaps due to habit.
A lot of our kids don't think they understand when they do, which is really weird to me. They'll act like they don't know and then they want you to help them. They want you to give it to them and you just have to ask one question and they realize they do know, and it’s fine.

Like Ms. Quinn and Mrs. Kilpatrick, Mrs. Williams also worried about her eighth-graders well-being. “We have students with such issues,” she said in her pre-observation interview. These spill over into class and prevent students from doing their best work, and at times, any work at all. She explained that teachers know these issues have to be dealt with before students can be expected to learn.

Dealing with the whole child is often frustrating for teachers, and it is important to understand that students do not develop on the same schedules. Instead of blaming students or parents for behavior that results from immaturity, or assuming that students would never be able to behave according to a certain standard, Mrs. Kilpatrick wisely accepted this difference in development.

Ms. Quinn and I sort of do a thing with groups, and sometimes I might change it depending on the class. Some classes - they work great in a group. There’re probably only four kids in all of 5th grade [who] cannot work in a group, and they will always be alone because their maturity levels are just not there yet. They cannot do it yet.

Another aspect with which Mrs. Kilpatrick was concerned was the number of changes administrators and teachers often make during the year. Speaking about the routine she and Ms. Quinn had forged with students with the group protocol and others she hoped to establish, Mrs. Kilpatrick explained that students needed routine and predictability:
You can't change every two seconds or [the students] are never going to catch on and get better - you have to give it a chance. And that's the part that new teachers don't understand. If it was really, really bad, then change it, but you want to try a routine [and] give them a chance to figure it out. If it still doesn't work then change it. But don't be like, oh, let's do this, let's do this.

This need for predictability might also apply to the decisions made at the district and school levels. For the first semester, Nolan Reid had a thirty minute, school-wide remedial reading period in which students were asked to read a passage and then answer questions in order to increase their reading skill. When second semester began, this period was changed to 45 minutes and added to the end of the day. Its focus was still reading, but it had less structure. Another major change during the year was changing the frequency of benchmarks from every four and a half weeks to every nine weeks. The district also changed programs, from Schoolnet to Case 21. According to Ms. Quinn, all of these changes were hard on students and teachers alike.

For Mrs. Jackson, one of the great advantages of OFAPs is in the ability to individualize and differentiate. With these programs, she said, “I can do more one-on-one targeting of students’ skills” and she could “customize based on students’ needs.” She liked Achieve 3000 because of its ability to reach students at all Lexile levels and help them improve, as well. She also appreciated the immediate feedback for students so that they know immediately how they did. As do the other three teachers, Mrs. Jackson searches for ways to make sense of the data she gets, not only data on student proficiency, but data she collected through observation and analysis. For example, Mrs. Jackson experimented with
longer assessments that mirrored the length of the state end-of-grade tests to determine whether students scored better on earlier passages because they were easier or from more familiar genres or because students tired as they moved through the assessment. Learning that the decline in student success over the course of the long assessments was due to the assessment length, she was able to help students learn to deal with that problem.

I even shared this with the kids. I [said], ‘This is what happens – I know why - because the passages [are not all that interesting] - they just [are not], and you don’t care and you’re frustrated and you have to get over that. I said, ‘Why don’t you go ahead and name it so you can change it?’

Her scaffolding of students’ justifications for their answers is another way in which Mrs. Jackson puts students first. Originally, students had to explain why an answer was the best choice. After practicing that on several activities and assessments, Mrs. Jackson required students to justify their answer on a corrected question. For the activity/assessment I observed, Mrs. Jackson gave them the correct answer when they had at first gotten the answer wrong, and then asked them to explain why that answer was the best one. Although it would be easy to require students to always explain their answers, Mrs. Jackson carefully planned how to teach students the strategy, practice it each time, use it when they needed it, and then learn to succeed without it. During her directions for the observed activity, Mrs. Jackson assured her students that the activity was not a “gotcha” one. She told them one of her goals was to help them think through each question. She said, “And will I help you? Absolutely. If you get stuck, I will walk you through it.”
While education would seem to be about putting students first, students and their needs can sometimes get overshadowed by the myriad demands on teachers’ time. The four teachers in this study consistently kept the needs of students a priority in their planning and practice. As Ms. Quinn said, “Hopefully I’ve sparked a love a reading [in] some kids.”

**Conclusion**

In their use of online formative assessment programs, these four middle school language arts teachers experienced a variety of issues, including a lack of resources, lack of information, programs that functioned poorly, programs with inadequate or poor content, and programs with tools that did not support teacher or student needs. However, even when all four had to print the assessments out, negating most of the advantages of using OFAPs, they found ways to use them to support student learning. Themes that emerged from the study included overcoming obstacles, such as a lack of needed technology and of information, encouraging metacognition in students through various strategies such as question analysis and justification, leveraging the experience each teacher had to increase the effectiveness and ease in which she was able to incorporate formative assessment, and the ways in which teachers in this study focused on putting students first in their planning and practice.

This chapter presented the data collected and analyzed for this study, which focused on the question: How does the use of an online formative assessment program impact teacher planning and reading instruction in the middle school English Language Arts classroom? Included in the chapter were portraits of each of the four case study participants: Mrs. Williams, Ms. Quinn, Mrs. Kilpatrick, and Mrs. Jackson, as well as the four themes that
emerged from the cross-case analysis of interviews, observations, and documents. The next chapter will provide a discussion of these findings and implications for future research.
CHAPTER 5: Discussion

Introduction

The purpose this study was to investigate teachers used an online formative assessment program (OFAP) in reading instruction. A qualitative methodology was used for this multi-case study focusing on four middle school English Language Arts teachers who used online formative assessment programs to support their reading instruction. Each of the four teachers participating in the study taught reading at the middle school level.

The overarching question guiding this study was: How does online formative assessment in reading impact teachers’ planning and instruction in the middle school English Language Arts classroom? This chapter includes a discussion of the study’s foremost findings as they relate to the research question as well as implications of the findings for the field of English language arts. Suggestions for future research follow the discussion.

This study is unique in that, to date, there is little or no information about the use of online formative assessment programs in the English Language Arts classroom. However, as more and more programs are made available, it is important to understand what their impact is in the classroom. While this was a small study, it indicated advantages and challenges in using OFAPs in the middle school ELA classroom.

The teachers in this study shared several characteristics. All four were challenged to use OFAPs without the kinds of technology support they needed for students to easily access the programs online. Gaps in the information two of the teachers received about available program was another challenge. Characteristic of all four was that they gathered and mined data from the OFAPs but also from listening to and interacting with their students. They
used the data they mined to inform both their planning and instruction, but they also used it to help students become better learners now and lifetime learners in the future. Finally, characteristic of this study’s teacher participants was the manner in which they integrated the use of these online formative assessment programs to further refine their practice, through reflection and refraction and knowledge of and concern for their students.

One of the most common criticisms of online formative assessment programs is that they are intended to take the teacher’s role in assessment. Another is that teachers depend on them for student instruction. Based on the findings of this study, nothing could be further from the truth. Instead, these programs not only enhanced what teachers did with students, but also, teachers adapted the programs to fit the unique needs of their classrooms and their students. Even when, due to the lack of appropriate technology, the teachers were not able to enjoy the immediate scoring, reports, and record-keeping OFAPs offer, they appreciated the accessibility of appropriate passages with questions that solidly aligned to the ELA standards. Each of the four teacher participants used the OFAPs for the intended purpose of freeing them to concentrate on what teachers do best: analyzing student strengths and weaknesses as they relate to curriculum goals, planning and implementing strategies to remediate weaknesses, and teaching students strategies to use immediately and in the future to help them become independent learners.

Discussion of Findings

The technology and information gap. The amount of information that must be collected, disaggregated, and reported in when using assessments to diagnose learning gaps is substantial. Brown, Hinze, and Pelligrino (2008) state that “no individual, whether a
classroom teacher or other user of assessment data, could realistically be expected to handle the information flow, analysis demands, and decision-making burdens involved without technological support” (p. 253). Technology easily enables teachers to collect, analyze, and report student data quickly, accurately, and easily. Programs which assess students and then make the data available to teachers are readily available. Yet the technology and information gap prevents these programs from being used to their potential.

Assessing students by computer has been possible for many years. It has become the norm in many businesses and science-based arenas, but for several reasons, including a shortage of the required technology and concerns about Internet access, such programs have only recently seen widespread use in public schools. In 1999, the National Research Council supported the use of computer software for assessments, pointing out its cost effectiveness for scoring and its ability to provide immediate student feedback. At that time, online programs were already being utilized in schools for both teaching and for assessment.

Technology can help with the design, administration, and scoring of performance assessment and can help reduce the burden on teachers when more formative assessment is added to the instructional process. [For example], computer software can facilitate administration, scoring, and reporting of concept map assessments, reduce scoring costs, and provide helpful feedback on student proficiency in specific cognitive areas (p. 14).

Several years ago, the North Carolina Department of Public Instruction (NCDPI) announced that by the year 2014 all state summative assessments would be administered to students online. The change to computer testing would be cost-effective while offering
immediate scores and reports to schools. Students with special needs would be more easily and effectively served. For example, students who needed assessments read aloud could use earphones, and students with visual impairments could use one or more adaptive screen features. However, the announcement produced a panic in many school districts as administrators scrambled to find money for computers, extra infrastructure, and training to ready teachers and students. Consider that every student in the state needed access to an Internet-connected computer for up to six multiple-hour tests during a window of time lasting no more than two weeks. Not only would the computers need to be available, but the infrastructure had to support the vast amount of data that would travel back and forth between NCDPI and the classroom. In 2010-2013, I traveled throughout the state to help train teachers in the use of an online program. In virtually every school, in wealthy districts and in poor ones, I heard both teachers and administrators lament their lack of computers and Internet capability for the North Carolina End of Grade tests. Many districts did budget for more computers and some increased their Internet capabilities. However, few were able to upgrade their technology to the necessary level that would be required by the online state-wide tests. The year 2014 came and went, but to no one’s surprise, the move to on-line testing for all students in all North Carolina schools did not happen. At this point, NCDPI has not named a new deadline.

However, it still came as a surprise that all four teachers in this study were hindered in assessing students online because of the lack of technology. Although both teachers at the suburban schools, Mrs. Williams at Edwards Middle School and Mrs. Jackson at Turtle Creek Middle School, were fortunate in that their students had school-provided laptops, the
laptops were not conducive to reading assessments with lengthy passages. In her pre-
observation interview, Mrs. Williams explained that new, required online math programs
made it even more difficult to schedule time in the media center to use the desktop
computers. Mrs. Jackson, in her post-observation interview, also mentioned the difficulty
she had in scheduling time with the media center desktops. As a result, neither teacher used
the OFAPs online as often as they would have liked. This circumstance also meant that
many of the advantages of using an OFAP, such as automatic scoring, immediate feedback,
and online assessment portfolios, were negated.

At Nolan Reid Middle School, Ms. Quinn and Mrs. Kilpatrick would have been
grateful to have the problems described by Mrs. Williams and Mrs. Jackson. Although
Nolan Reid Middle School had a computer lab, there was no working Internet connection.
Mrs. Kilpatrick’s status as a second year teacher at Nolan Reid Middle School and her
fortunate connection with a parent donor had at least provided two desktops and four iPads.
Ms. Quinn, on the other hand, was left to work with one desktop computer of undetermined,
but not new, age. Therefore, their use of OFAPs was limited to offline use. As Mrs.
Kilpatrick said in her post-observation interview, “I can get markers, but finding a computer,
ho-ho - good luck with that.” Ms. Quinn commented, “[I have my] fingers crossed [that we
will have] computers next year - better computers that our kids can use . . . they say it's in the
budget.”

Additionally, the lack of computers at Nolan Reid meant that teachers relied on a
once-a-quarter benchmark for data to inform their instruction. Without readily accessible
computers, they were unable to use OFAPs to administer the frequent, during-instruction assessment that is necessary to be truly formative.

According to the North Carolina School Report Card (2013), at Nolan Reid Middle School, 100% of the school’s classrooms were connected to the Internet. However, although classrooms may be Internet-ready and even connected, many other problems can impede access, including infrastructure, routing, connectivity, and even the district protocol for product use.

Using OFAPs had other challenges that were frustrating for three of the teachers. For Ms. Quinn and Mrs. Kilpatrick at Nolan Reid Middle School, one issue was their lack of knowledge about how Schoolnet, the OFAP their district used, worked and what was available to them. In their post-observations interviews, it was clear that neither knew that Schoolnet, for example, was provided free to students via PowerSchool, the school information system provided by the state. With the seven computers she had been able to collect in her classroom, Mrs. Kilpatrick might have been able to use Schoolnet online to some extent. Mrs. Kilpatrick also expressed frustration in her post-observation interview about the lack of training about Schoolnet. “There hasn’t been anything [except] a wiki.” She also said that more training would have been helpful because “we need to know – here are all the options.”

In her pre-observation interview, Ms. Quinn also described finding passages on the Schoolnet district benchmark that she had recently used on a practice assessment.

I would select a passage and print it out for the kids to work on in class and then we'd go over and then when it came time for the summative assessment, they’d see the
same passage, and the kids did really well . . . Well, that happened at least twice, so then I would think, well, why use it, because it's not going to give me accurate data.

Schoolnet has separate banks for classroom and district benchmark use to prevent this situation, so it is puzzling as to why this happened. The result, however, was that Ms. Quinn stopped using the resource. The informational disconnect among state, district, and teachers created unnecessary roadblocks for both teachers and students. Without the necessary tools and information to access it, even a perfect program has little use.

**Making sense of data.** As data is an integral part of formative assessment, determining how teachers made sense of the data derived from OFAPs was a key component for this study. Hunt and Pellegrino (2002) emphasize that “there is no point in formative assessment by a teacher if the teacher cannot identify, analyze, and respond to the problems of individual students” (p.75).

A personal experience proves the point. The second or third week of my son’s fifth grade year, his teacher sent home a packet of papers showing that my son was struggling with long division and suggested that he be enrolled in a special class for remediation that would replace his physical education class. Since my son had never before struggled with math, I was surprised and puzzled. I studied the papers the teacher sent home, and then each of the problems my son had missed. I soon realized that, rather than having “trouble with math,” my son was making a single error in every problem. I sat down with him and showed him what he was doing wrong. I also did not agree to have him placed in the special class, but wrote a note on the packet explaining what I had found. From then on, there were no problems in math. The teacher had made a critical mistake before deciding that my son did
not understand math and needed to miss an elective so he could get extra help. He had not analyzed the data he had. Had he done so, he would have seen what I saw, and would have also been able to quickly correct the misunderstanding.

In contrast to the circumstance with my son, the four teachers in this study recognized the essential role of data analysis and incorporated into their regular practice. At Nolan Reid Middle School, Ms. Quinn and Mrs. Kilpatrick studied the results of their fifth-grade students’ benchmarks and then divided the skills they needed to address into groups. Then they grouped their students according to their performance on each skill. In this way, they knew which students needed intense targeted intervention on any given skill and which students needed a little more practice. In her analysis of student weaknesses, Mrs. Kilpatrick realized that vocabulary was an area in which all her students struggled, preventing them from understanding what they were reading. Therefore, she made vocabulary a priority in her Science/English Language Arts class. The more experienced teachers, Mrs. Williams and Mrs. Jackson, not only looked at individual performance on assessments, but looked for overall patterns in class performance. As Mrs. Williams said, “You can see the patterns forming,” so she then tried to determine “why it was such a tricky question.” The questions that give students the most trouble are ones she chooses to address with her whole class: “I can pull [a question] up and say, ‘Oh, okay, let’s all talk about this as a class.’” Mrs. Jackson also used patterns to guide future instruction:

I look at trends of what students are missing and what wrong answers are showing up commonly as well as what they are doing well. I base subsequent teaching and assessments on this [information].
During the activity I observed, for example, Mrs. Jackson was able to see almost immediately that students were consistently missing a question on summarization and this question represented a reading skill she needed to work on with them. At first, in reflecting on why students missed the questions, Mrs. Jackson was puzzled, since students had quite a bit of practice during the year on summarization. She realized, however, that students had practiced generating summaries, but needed practice in recognizing good summaries when presented with several choices. “They are better at generating than they are at analyzing the answer choices and actually matching up the answer choices with what is there on the paper.” This kind of analysis results in efficient use of class time, because teachers know not only what skill (summarization, for example) needs more practice but allows teachers to spend time only on the reading sub-skill (recognizing the best summary from several) that needs attention.

Mrs. Jackson not only examined the data, but looked beyond to see what other factors, such as fatigue during extended assessments, might be influencing student performance.

I gave [my students] a full-length practice test, and I analyzed the pattern of their right and wrong answers . . . and I found that [on] the first third of the test, most kids did really well, not a lot missed; middle third, [not bad]; last third, [they bombed]. So then I gave them another one and totally switched out the order of the types of passages - same thing happened.
This fatigue factor prompted Mrs. Jackson to have her students practice marathon assessments so that when they had to take the state End of Grade test, their proficiency scores were less affected by fatigue.

The teachers in this study demonstrated that using data was an ongoing, natural part of their planning, instruction, and pedagogy. They studied data from the OFAPs to inform all aspects of their practice, including setting instructional goals, planning, instruction, assessment, feedback, and targeted intervention. In addition to examining data from the OFAPs, these four teachers also used data they had from observing and interacting with their students. They reflected on the information they received from OFAPs, but they additionally used the data gained from observing and interacting with their students to help them choose methods of instruction that would enable students to reach instructional goals.

**Helping students become better learners.** Though this study focused on the way teachers used OFAPs in their planning and instruction of reading, the purpose of the OFAPs they used was to provide them with information on student progress toward instructional goals in reading. While it would be comforting to assume that these instructional goals were solely goals about reading proficiency, the current educational climate dictates other meanings as well. North Carolina English Language Arts teachers are charged with ensuring that students are proficient on the Common Core State Standards in ELA, on which the state end-of-grade test is based. Currently, the scores from these tests are not only used to make decisions about students’ educational settings, but also count heavily in teachers’ performance evaluations. They can also influence staffing, funding, and student assignments. Given these high stakes, teachers do everything they can to ensure student
proficiency on these standards. Ideally, reading proficiency as measured by the end-of-grade tests would translate to proficiency in the reading standards, which would translate to proficiency in reading itself. Thus, the tests should be excellent assessments which accurately reflect the standards, and the standards should accurately describe what a student must be able to do to be able to read on or above grade level. Unfortunately, these criteria are not always true at the same time.

Even if both were always true, there are other factors to consider, such as whether a diverse group of students can all access a test in the same way. What about learning styles? What about attention-deficit and learning disabilities? How can we ever be sure that each child experiences the same test as his peer? Further, just as one photograph of a child cannot reflect his year, neither can one snapshot reflect a child’s learning in one year (Heritage, 2007). Thus teachers are caught between what they know is good practice, what they see in class daily and over time, and what they see as the high-stakes-test dragon breathing down their necks.

As a result, many teachers develop a sort of hybrid teaching practice. They teach reading, and fervently hope, as Ms. Kilpatrick did, that they’ve “sparked a love a reading [in] some kids.” But they also closely monitor how well their students’ skills match the state standards, and they teach the skills students must master in order to show what they know on the state tests. In the classes I observed, I saw all three elements of this hybrid.

To assist students in taking assessments, the teachers in this study taught students how to navigate through them and how to analyze both text and questions. To activate students’ understanding and learning, they used connections and questioning to make the text
engaging, relevant, and understandable. To help students think more deeply about the text they read, as well as better analyze the questions, these teachers taught students to justify their answers by going back to the text. To help students become independent learners, the teachers taught their students to reflect on their progress and on their learning. The teachers married these lessons in test-taking and reading because students need to be skillful in both.

In explaining the importance of activating students’ prior knowledge for formative assessment, Otero (2006), referenced Vygotsky’s theory that helping students become conscious of concepts they have gained through experience allows them to transfer this understanding across situations. Further, Vygotsky stated that teachers should explore this prior knowledge before making instructional decisions in order to optimize learning. In her post-observation interview, Mrs. Kilpatrick explained why connections were so important to her fifth graders. “I’ll think about how I introduced [the topic or concept]. Did I try to link it to what they already knew? . ..[because] if they can’t relate it to anything then they’re not going to understand at all.” In Mrs. Kilpatrick’s science/ELA classroom, making connections for her students meant that a lesson on inherited traits was introduced by talking about breeding dogs, because that is what students knew. In Mrs. Williams’s eighth grade classroom at Edwards, these connections took the form of personal, political, and literary associations. For example, during the observation, when discussing a question on the assessment having to do with a staircase metaphor, Mrs. Williams helped students connect the message the poet was trying to convey to a Langston Hughes poem they had read earlier in the year that also used a staircase metaphor. In their discussion of other questions,
students made connections to personal experiences, wars, social protests, and other important historical events, well-known people, and other literary and informational texts they had studied throughout the year. Making these connections not only helped students understand the current questions and text, but were another way of helping students on their way to independent thinking.

In discussing the assessment her seventh-grade students took during my observation of her class at Turtle Creek Middle School, Mrs. Jackson expressed surprise over the number of vocabulary words that gave students trouble. Throughout the year, she had worked with students on linking what they already know, such as prefixes, suffixes, roots, and whole words, to words they have never seen. This parts-to-whole strategy is also an important one for students to become independent learners.

Ms. Quinn also touched on the importance of making connections for students at Nolan Reid Middle School. At the beginning of the year, she consistently asked students about main idea. Although they could give her the definition of main idea, she found that on the district benchmarks they were not able to apply that knowledge to identify the main idea in a reading passage. In her post-observation interview, Ms. Quinn reported realizing that she needed to make those connections for them so that they could understand main idea on the deeper level. She modeled finding the main idea and began asking students to explain the main idea of a particular passage. Having learned that how a question is asked is as important as the question, she was able to move students to a deeper level of understanding. The NCTE Guideline “A Call to Action: What We Know About Adolescent Literacy and Ways to Support Teachers in Meeting Students’ Needs (2004) supports the types of
connections these teachers made, including making texts relevant to real life experiences and providing “appropriate, strategic assistance” for students who are not making meaning with text (p. 3).

The teachers in this study also recognized that students can focus too much on assessment questions and performance, forgetting that the instructional goal is practice and improvement in reading. One of the ways the four participants addressed this concern was to require their students to explain why they chose the answers they did on multiple-choice questions. This process encouraged students to go back into the text to look for support for their answers. To increase comprehension, according to the NCTE Guideline (2004) on adolescent literacy, teachers need to encourage students to “engage in critical examinations of texts as they dissect, deconstruct, and re-construct” for comprehension (p.3).

Analysis of student explanations was critical for teachers to understand the logic and thinking processes students were using and to check for text understanding. Teachers also found that students were more likely to take their time thinking about both the text and the question and less likely to guess when they had to explain how they made their answer choice. The analysis of these answers also helped teachers conduct their targeted intervention more efficiently. Because students explained their reasoning in making answer choices, teachers were able to correct specific misconceptions rather than depending on their own interpretations of why students missed certain items. In the post-observation interview, Ms. Quinn explained that as she looked over the work students had completed during my observation, “‘I’m going to determine . . . [if] they explain their . . . thought process to me in a way that helps their answer make sense.’” If my son’s teacher had asked him to explain, as
Ms. Quinn did with her students, how he found the answers he did in his long division, his teacher would have been able to quickly see the misconception that occurred in every problem, without even having to study the problem itself.

Not only did this technique of justifying answers work well for data analysis, but it also helped students become more metacognitive. In her pre-observation interview, Mrs. Kilpatrick, teaching at Nolan Reid Middle School, thought that this practice improved her fifth graders’ “awareness of what they are learning.” At Turtle Creek Middle School, Mrs. Jackson first taught her seventh-grade students this process of explaining their answers for each question and then slowly gave more responsibility for the process to the students so that they would internalize that strategy and became responsible for using it on their own. Mrs. Jackson explained in her post-observation interview, “I’m most concerned . . . that they try to figure out the justification, because then that will help them correct their thought processes.”

In Black and Wiliam’s (1998b) seminal article “Inside the Black Box: Raising Standards through Classroom Assessment,” the authors emphasize the importance of dialogue between teacher and student:

Discussions in which pupils are led to talk about their understanding in their own ways are important aids to increasing knowledge and improving understanding. Dialogue with the teacher provides the opportunity for the teacher to respond and to reorient a pupils thinking. . . . What is essential is that any dialogue should evoke thoughtful reflection in which all pupils can be encouraged to take part, for only then can the formative process start to work. In short, the dialogue between pupils and a teacher should be thoughtful, reflective, focused to evoke and explore understanding.
and conducted so that all pupils have an opportunity to think and to express their ideas (pp. 143-144).

The explanation of their answers was one way teachers encouraged this dialogue. They also conducted important dialogue with students during class. Mrs. Williams, teaching at Edwards Middle School, worked to help her eighth graders internalize analysis by asking students to constantly question themselves about the way they answered a question, why a question was difficult, or how the question was crafted. Ms. Kilpatrick also used questioning techniques to help her fifth-graders read with critical and analytic eyes. She asked, “What are you understanding from this text? What are they trying to teach you? What is this about? What’s the point? Why are we reading it?” At Edwards Middle School, Mrs. Williams’s eighth graders used a list of five steps for analysis and self-talk that she gave them early in the year and referred to these steps as they discussed reading passages. As she reviewed the assessment with students during my observation, she continually asked questions: “Take a look at the ones you missed. Were you surprised at any of the ones you missed? Take a look to see if these were the ones you had trouble with when you were taking it.” These questions were focused on encouraging students to become constructivists themselves, creating their own meaning about their learning and thinking (Patton, 2002).

The third process the teacher participants used with their students to help them become better learners was to have students reflect on their progress toward learning goals. At Nolan Reid Middle School, both Ms. Quinn and Mrs. Kilpatrick worked with students to recognize that weaknesses could be overcome. In individual conferences, both teachers worked with students to help them reflect on their weaknesses on specific reading skills as
measured by the Schoolnet and Case 21 benchmarks. Throughout the quarter, as they explained to me, the teachers also encouraged students to think about how the work they were doing in class would help them increase their skills and move them closer to their goals. Mrs. Kilpatrick regularly met with students to help them reflect on their progress toward goals in science and in reading. These conferences included discussion about classwork, homework, and quizzes, as well as student progress toward weekly and quarterly goals. She also had them keep track of quiz grades and to communicate with her in writing any questions or concerns they had about what they were learning. Ms. Quinn and Mrs. Kilpatrick also helped their fifth-graders reflect on the choices they made in choosing the groups with whom they worked. During the class I observed, Ms. Quinn told two students that they needed to think about the choices they were making for where they would sit and decide if sitting with that group had helped them be successful in the past. Reflecting on their thinking was another way these teachers helped students become better learners. In the pre-observation interview, Mrs. Kilpatrick discussed asking her students, “Why do you think that?” Mrs. Williams, at Edwards Middle School, had students write journal reflections of their performance, what they found difficult or easy, and on the assessment itself after each ClassScape (Edwards Middle School’s OFAP) quiz, because “where the real learning takes place is when they go back and [determine] why the right answer is the right answer. “What did you miss about the question?” [I] have them do an analysis. “What did I miss? Why did I miss this?” And I have them try to figure out why. “Oh, I didn’t read the question carefully enough.” This feedback, reflection, and goal-setting are important components in moving students toward independence (Campos & O’Hearn, 2007; Black, et. al, 2004). In fact, in a
study conducted with first and fifth graders, Campos and O’Hearn (2007) found that implementing formative assessment strategies with descriptive feedback and student reflection increased student interest in learning. Further, both students and parents in their study thought that helping students become more knowledgeable about their progress and goals resulted in higher achievement.

**Integrating formative assessment.** During all the stages of the formative assessment cycles observed and discussed in this study, the teacher participants used formative assessment as an integral part of their practice. Rather than stopping class to administer an assessment or provide feedback and targeted instruction, teachers wove formative assessment into their overall instruction. For example, Ms. Quinn and Mrs. Kilpatrick used a seating protocol to allow them to easily target students who needed extra help. They also met with students individually after OFAP assessments to review concerns and progress toward goals. The scores they posted on their walls helped remind students that improvement was possible and expected. Mrs. Williams and Mrs. Jackson integrated several parts of the formative assessment model in one lesson, seamlessly moving from assessment to data analysis to feedback and targeted intervention.

Teachers in this study were able to successfully integrate formative assessment from online programs because of several characteristics they shared. First, they teachers consistently worked to refine their practice. Using reflection and refraction (Pope, 1999), teachers constantly asked questions of themselves about their practice and about their students to improve instruction and student learning. They also used their considerable knowledge of their students to make their practice more effective. Using transparency, they
helped students become partners in the learning process, moving students toward self-regulation, a major goal of formative assessment (Heritage, 2007). These characteristics allowed teachers to move beyond seeing formative assessment as a model for testing and data management extrinsic to instruction to using it as a model for learning, intrinsic to their instruction.

Unsurprisingly, a finding in this study was that experience helped teachers become more effective. However, experience that makes effective teaching is not necessarily experience measured in years. There is an old saying about a veteran teacher who was not very good: “She was a twenty-year veteran, but unfortunately, she had twenty years of first-year experience.” The “veteran” had never grown beyond her novice state even though she had been in the classroom for twenty years. The teachers in this study reflected constantly on their practice, and thus consistently grew more effective.

As a first year teacher, Ms. Quinn constantly looked back on her practice to determine what worked, `what did not, and how she could change it for the next day, unit, or for her second year. In addition to learning how to ask questions to get to deeper understanding, as she explained using her example of main idea, Ms. Quinn’s self-reflection centered on how much time she spent on certain skills. In her post-observation interview, she explained,

At the beginning of the year . . . I didn’t know how long to spend on one thing . . . , but next year I’m really going to spend more time at the beginning of the year going over some of the skills.

Also in her post-observation interview, she expressed concern over the time she was pulled away from working with her students in the front for remediation to deal with
discipline issues. “Sometimes there are discipline issues I have to address. I feel really bad for the 1s and 2s who are up there with me because sometimes we only get through two to three questions.” However, after reflection, she said, she had decided that, “I think it's going to help them more than just sitting in a group [where they will] get behind or [sitting] in the whole class [where they’ll get] lost.”

In her post-observation interview, Mrs. Kilpatrick commented more than once that when students did not understand a concept, she took responsibility for that misunderstanding. “If not a single kid understood it, then I know it's my fault. If it's just one of two, that's not going to work out, and I know I did something wrong,” she said. Similarly, she had learned to link lessons to something students already knew, such as when, in the observation, she began the discussion on inherited traits with examples from breeding dogs, because her students knew about that subject. She also had reflected on routine. “I want to be better at [creating a] schedule so that the kids know when remediation is coming. I know the kids thrive on schedules because they do so much better when they know what's going to happen.”

For Mrs. Williams, reflection was an ongoing practice. She brought this habit of reflection to her students, constantly encouraging them to think about how they interpreted a poem or answered a question. During my observation of her class at Edwards, she asked students “Were you surprised at any of the ones you missed? Take a look to see if these were the ones you had trouble with when you were taking it.” Her focus here was not on how many students missed, but on their reflections on their work. She also asked them “What did you think of these [questions] in general?” The invitation to students to consider these
questions from other viewpoints was powerful in showing how she herself looked at assessments and the work students did. In essence, she was inviting them to use refraction as well as reflection. As Mrs. Williams explained how important she thought it was to use ClassScape regularly, she said,

I would have them do an assessment every – oh, at least every two weeks if not more. I think it’s really worth it. I think if you don’t do it from the beginning of the year you really miss something. . . The problem is you get involved in a novel and you know there’s just one set and the next teacher’s waiting for you to finish [so you don’t use the OFAP as often as you would like].”

Just as Mrs. Williams consistently reflected on her practice, Mrs. Jackson did also. In addition, she reflected on what she was learning from students about her practice. She analyzed student performance on longer assessments and discovered that their performance grew worse as the assessment progressed. She took time to reflect on reasons for this occurrence, and then experimented to find answers. Instead of accepting the worsening of student performance as long assessments progressed, she reflected on how she might address this problem. She wondered, “What can I do to change their mental habits?” It was as a result of this reflection as well as reflection on previous formative assessment data that she decided to require that students write explanations for their answers.

Mrs. Jackson also used reflection with her students during the time of the observation. As she spoke with students about their work on the assessment, she realized that they were missing some of the same questions. She commented to me that she was unsure about the
reason students were missing these questions, and after reflection, decided that these questions were more abstract than others.

This use of reflection by the teachers in this study is supported by Glatthorn’s (1993) findings that teachers constantly think about “student needs, objectives, materials, and likely problems” and use this thinking as “a guide for making crucial decisions (p. 2). Similarly, Pope (1999) uses the metaphor of the mirror to explain teacher reflection. However, she continues this metaphor to explain refraction. In refraction, teachers turn the mirror to reconsider first reactions and to reinterpret what they have noticed. Mrs. Jackson followed a similar process in that she was not satisfied with her initial diagnosis that students were missing the question because it was abstract. The item required that students choose the best summary; since Mrs. Jackson and her students had spent a good deal of time on summary, she assumed that students had proficiency in that skill. However, she refracted by examining the questions and student responses again; looking beyond the fact that the students had practiced summary, she concluded that the difficulty for them was, in fact, summarization. Students could generate summaries but had difficulty differentiating the best summary of a passage from several choices. This refraction process was essential; without it, Mrs. Jackson would have spent time working on abstract questions, when that was not the skill students needed to practice.

In constantly using both reflection and refraction, teachers were able to make both large and minute adjustments in their instruction to best meet student needs. They consistently challenged the assumptions they made as they looked over their students’ work
and over their own practice, and it was this reflection and refraction that made their teaching effective.

In addition to reflection and refraction, teachers used strategies that Clark and Lampert (1986) identify as types of knowledge. Teachers’ interactive knowledge allows them to understand how to achieve their goals through instruction. Their speculative knowledge is that which allows them to make short and long-range decisions but also helps them manage the myriad of decisions made during class, including dealing with student behavior and attention, and what types of questions to ask. These kinds of knowledge are based in large part on teachers’ knowledge of students. Throughout the study, based on the interviews, observations and documents review, the teacher participants demonstrated their knowledge of their students. All four used their knowledge of their students to determine the best kind of assessment to give at a particular time, how to design classroom activities and assessments, how to probe for thoughtful answers, how much support and encouragement to offer, and even how to design their classroom seating.

Beyond noticing that her students were missing certain common questions in the assessment they took during the observation, Mrs. Jackson also was able to see that some of her higher-level seventh-graders were “overthinking” questions, making them seem harder than they were. She commented on this problem in the post-observation interview, and said, “I think they can benefit from analyzing the answer choices and figuring it out. “ Recognizing that students were overthinking a question was important in planning next steps for them. If they missed the questions because of missing concepts, targeted intervention would be to fill in those. If they had missed them due to carelessness, another intervention
was needed. Determining that they were overthinking the questions, Mrs. Jackson was able to prescribe an effective intervention.

For Mrs. Williams at Edwards Middle School, knowing her seventh-grade students was also an important part of her teaching. In her pre-observation interview, she said she was concerned about several students in her classes who had problems of one sort or the other over the year. These included family and personal problems as well as those that stemmed from school. She worried about how these students were dealing with the problems and how their problems affected their learning. During the observation of her class, she encouraged her students to talk about personal concerns as they made connections to reading passages, questions, or discussions. One student, for example, discussed the impact that one school family’s loss of a family member had on the entire school. This concern for her students and acknowledgement of her students’ needs to have their own concerns recognized is supported by several researchers and psychological theorists, including and Dewey (1916) and Garbarino (1982), who emphasize the important role in education a student’s personal life has. In fact, according to Dewey (1916), “it is the teacher’s obligation to know each child intimately,” Sullivan (1966). Mrs. Williams further demonstrated her knowledge of her students in her post-observation interview when she explained that for an assessment, “I’ll give them ten questions, because that’s about how many . . . my slower readers [can finish], and ten is a good number because usually they get most of them right.”

The importance of knowing students and their needs was also significant for Ms. Quinn and Mrs. Kilpatrick. They knew when they volunteered for Teach for America that their students would need more than most, and they pushed themselves to meet that
challenge. Through a process of first reflection on student behavior and then refraction, Mrs. Kilpatrick attained a kind of wisdom about her students that greatly influenced the way she and her students interacted. Students at Nolan Reid Middle School are used to teachers’ leaving. Because Nolan Reid Middle School is in a rural area, it is difficult to find and attract teachers. Some stay a year or two. Some stay less than a year. During her post-observation interview, Mrs. Kilpatrick explained how important it was to become part of the community. Fifth-graders in Ms. Quinn and Mrs. Kilpatrick’s classrooms already knew that teachers who do not become a part of that community will not stay. Therefore, both teachers spent part of their free time making community connections. They met with students and parents to go on hikes, they participated in community runs, and attended their students’ sporting events. Without that investment in the lives of their students outside the classroom, students did not invest themselves in their teachers or classroom.

Another aspect of knowing students is understanding that they mature at different stages. Maslow’s (1968) health and growth hierarchy illustrated that humans must complete one stage before being able to move to the next and that that certain needs must be met before children can advance to the next stage. In classrooms, students are often at different stages on this hierarchy. Both Ms. Quinn and Mrs. Kilpatrick spoke in their post-observation interviews about dealing with students who were not able to work in groups. Rather than labeling these students as behavior problems, both teachers recognized that working in groups is a skill for which some of their students were not yet ready. They were prepared, therefore, to allow students to make the choice of sitting by themselves or to make the choice
for them because, as Mrs. Kilpatrick said, “their maturity level is just not there year, and they cannot [work well in a group] yet.”

While they recognized student differences in maturity and learning, teachers in this study realized the power of sharing this knowledge with students. One of the most important tenets of formative assessment is the involvement of students in learning. The eventual goal of formative assessment and education in general, is to enable students to take charge of their own learning; therefore, they must become partners with teachers (Heritage, 2007). The teachers in this study used transparency as one strategy for moving students toward partnerships in learning. This transparency was evident in the relationships the teachers developed with students. The teachers were open with students about their goals, about why certain students needed individualized support, and about the fact that assessments can be boring. Ms. Quinn and Mrs. Kilpatrick used a grouping protocol at Nolan Reid Middle School that enabled them to spend more time with students who were struggling the most, while giving others the opportunity to benefit from more proficient peers. This grouping protocol followed Vygotsky’s (1978) theory of the ZPD, which is that the difference between what a child knows and can do independently, and what that child can learn must be supported or 'scaffolded', with help from an adult or peer. Even though she had to cover both English Language Arts standards and social studies standards, Ms. Quinn was committed to moving as slowly as necessary with her small group in order for them to understand the concept she was teaching, even when she worried that they were not covering enough material. In this process, she demonstrated what Carroll (1963), Bloom (1968) and others
established about student learning: students who are allowed to learn at their own pace (i.e. given enough time) can learn as well as their peers to whom learning comes more easily.

In Daniels’s (2002) model for literature circles, students adopt specific roles in groups of varying skill levels to ensure good discussion and group understanding. Ms. Quinn and Mrs. Kilpatrick adapted this model in their grouping protocol to allow more able students to support the understanding of their peers. In her post-observation interview, Ms. Quinn explained that at each table is a standards master, wearing an identifying lanyard, a student who has demonstrated skill in whatever standard is under study and thus can assist other students in the group.

Ms. Quinn and Mrs. Kilpatrick also posted student scores on their classroom walls to let students know exactly where they stood and to remind them that the goal for each student was to improve. Ms. Quinn and Mrs. Kilpatrick worked to create a school culture for their fifth graders that was accepting of individual differences but was focused on improvement. After several weeks, according to Mrs. Kilpatrick, students who were pulled to the front for intense remediation began to understand “why it would probably be good for them to be here.”

This openness about what is traditionally private information might seem unwise to some. However, at Nolan Reid Middle School, teachers created a culture in which students recognized that a score was just where they started and that finishing at a proficient level is what counts. That is not to say that either teacher lacked sensitivity. For example, during my observation of her ELA/social studies class, Ms. Quinn made it clear to one fifth-grader who did not want to be at the front that she wanted him to get extra support directly from her. In
her post-observation interview, Ms. Quinn’s colleague, Mrs. Kilpatrick, spoke fervently about the need of her students to believe that she cared enough about them to become a part of their community outside of school. Because of the relationship of trust and mutual respect Ms. Quinn and Mrs. Kilpatrick built with their students, their transparency in grouping and scores worked.

Mrs. Williams also used transparency to help her students understand that teacher and student were working together as a team. During the observation, in which she and the students reviewed a previous assessment, she often took on the role of a cheerleader, clearly thrilled when students made difficult connections or analyses. She also informed students of commonly missed questions, went on to give them the most popular incorrect answer, and then asked them to explain why it was popular. Other problem questions were projected on the whiteboard for discussion. When a student admitted to not having completed an assignment, Mrs. Williams asked, “Well, do you think you should?” Putting this decision back onto the student made it clear that her job was not as a task-master but as a coach. Given this atmosphere, students felt confident in claiming the growth they were making. Several times during the observation, Mrs. Williams reminded students to look back over the text to find support, or proof that the answer they chose was correct. Near the end of the class period, one student sheepishly admitted that each time he had looked back in the text and found proof, his answer was correct.

When Mrs. Jackson initially did the experiment in extended assessments, she wanted to determine whether students scored lower on the end of the test due to boring passages or to the length of the assessment. Although her research confirmed that the lower scores were
generally due to the length, she reported acknowledging to students that the passages often were boring, and then said, “they just [are], and [so] you don’t care, and you’re frustrated, and you have to get over that. [So] why don’t you go ahead and name it so you can change it?” Allowing acknowledgement of the imperfections inherent in the assessments helped to develop the atmosphere in which learning could best take place.

Although the integration of the formative assessment model these teachers were able to achieve was not dependent on using an OFAP, the fact that OFAPs are more formal in nature, and may best be described by what Heritage (2007) identifies as embedded assessments, or those that are more formal in nature, may have resulted in a more conscious effort to integrate all stages. Noteworthy, however, is that even without the kind of technology they needed to properly access the programs, the teachers in this study were willing to adapt the use of the OFAPs to their circumstances because they considered OFAPs a valuable tool.

**Problems for OFAP Implementation in the English Language Arts Classroom**

This study focused on how middle school English Language Arts teachers used online formative assessment programs (OFAPs) and the data they produced in their planning and reading instruction. Results of the study revealed that teachers understand formative assessment and are eager to use the tools OFAPs provide. However, lack of appropriate technology, including classroom computers, adequate Internet access, and screens large enough for students to access long reading passages were constraints to using them often and consistently.
According to Brown, Hinze, and Pelligrino (2008), technological support is imperative for teachers to be able to handle the “significant amounts of information [that] must be collected, interpreted, and reported” for diagnostic assessments (p. 253). Bell and Cowie (2001) acknowledge that formative assessment is a complicated and multi-faceted task. The formative assessment model requires that assessment be embedded in the instruction and used frequently. Essential to using OFAPs then, is easy and immediate access to Internet-capable computers. Assessments for reading pose the additional challenge of including long reading passages that may not be easily viewed on laptops or tablets. Since effective assessments in reading require that students look back at the text to analyze, synthesize, find support, or determine how words or language are used, students must be able to look back at parts of the text as well as to the text as a whole. Schools who have invested in laptops or tablets as a way to stretch technology dollars may be impeding the ability of students to fully access the assessments because of the smaller screens laptops and tablets provide. For other schools, such as Nolan Reid Middle School, there may not be enough Internet-ready computers available to allow for frequent and within-instruction assessment. Teachers at schools lacking adequate technology must then choose between assessing less, at least with OFAPs, and thus negating the advantages to students of frequent assessment; printing the assessments on paper, thereby losing the advantages of immediate scoring, feedback, and online progress portfolios these programs offer; or not using OFAPs at all.

A second important point is that sometimes the subjects chose to use paper and pencil. One reason was that because the state still uses paper and pencil summative tests, teachers wanted to prepare their students for these kinds of tests. However, paper and pencil
also have advantages in that students can “get into” the text in a way that is difficult, if not impossible, with online assessments. Three of the four teachers in the study indicated that this factor was one consideration when planning a formative assessment. Ms. Quinn spoke about the necessity of students’ being able to highlight, underline, and write notes on a print copy in order to better understand the text, to reflect on it, and to use text analysis strategies they had been taught. Online programs now provide tools such as highlighting and/or underlining in addition to others that are computer-specific, such as enlarging text, changing background color, and read-aloud capabilities. Thus, while it seems that students may benefit from a combination of both the tactile advantages of paper and pencil and the digital ones of online programs, the investment of teachers’ time to print assessments, score them, and record scores must be considered.

In her post-observation interview, Mrs. Kilpatrick also expressed concern over the amount of paper it takes to print assessments. “[The assessments waste] so much paper the way [they are] laid out so usually I'll copy it and put it in a word document and make the font smaller.” Another concern with primarily using paper and pencil is that students live in a technology-based world and must learn the skills necessary to live in that world. However, until schools are able to provide adequate computer and Internet access and the state summative assessments move completely to online formats, English Language Arts teachers, along with their students, will be caught in a dual world of paper, pencil, and computer, and students will need to be proficient in both formats of assessments.

A third problem is that for some teachers, there may be a disconnect between the state, district, and teachers about available programs, how to access them and how to use them.
Providing teachers with tools to support student learning is only effective if teachers have been given the appropriate information on how to use them. At Nolan Reid Middle School, both Ms. Quinn and Mrs. Kilpatrick were frustrated with the lack of information and training they had received on Schoolnet, the OFAP that was provided free to all teachers and students in the state. Ms. Quinn did not know, for example, that each passage assessed only one reading standard so that teachers could assess students in the small increments best suited for formative assessment. The design of the program was based on two formative assessment concepts: Breaking learning down into smaller “bites” allows teachers to determine exactly where the gap in student understanding occurred and to be able to effectively target the gap in subsequent instruction. At the same time, breaking instructional goals down to smaller increments, or sub-goals, is more likely to result in success in closing the gap. Sadler (1989), for example, explains that a student must feel that the goal of closing the gap is attainable, yet must also see the gap as significant enough to warrant the effort to close it. Without this critical information about the design of the program, however, Ms. Quinn was unable to use the program as it was intended. Although Mrs. Kilpatrick was aware that Nolan Reid Middle School owned a student response system, or clickers, she reported that it was difficult to check them out, and Ms. Quinn was unaware that the school owned one. A student response system connects clickers, or small hand-held devices similar to remote controls, to one computer program. Students can each respond to a question shown on a screen and the computer stores their answers and can even import them into programs like Schoolnet. Using a student response system would allow both teachers to assess students online with
only one computer and gain back the online assessment advantages of immediate scoring, feedback, and student portfolios.

Another disconnect in information for Mrs. Kilpatrick was that she did not know how to make Schoolnet available to students online. Though the state used state-issued student identification numbers that were also used for student lunch accounts and other programs, this had not been communicated clearly to teachers at Nolan Reid Middle School. Though Internet-accessible computer access was limited, Mrs. Kilpatrick might have used her five iPads and two desktops to assess students online had she had this information. Also, although a wiki was made available to teachers, Mrs. Kilpatrick expressed frustration that during the training session other teachers at her school were impatient with questions others asked, and Mrs. Kilpatrick thus felt that she was not able to get full benefit of the training.

State and local administrators must be cognizant that some teachers, especially those who are new to the classroom, need support that goes beyond simply introducing a new program. For example, referencing student lunch numbers as logins might not be clear enough to teachers who are new to the classroom and are not familiar with what seems to be standard information and processes. Care must be taken to make sure that all teachers understand the programs available to them and how to use them. This may mean that several smaller training sessions are made available over the course of the first quarter, rather than during the first or second week of school, so that new teachers, who are overwhelmed with information onslaught, can better assimilate new knowledge. Further, it is the responsibility of school administrators to ensure that teachers are aware of the materials and tools available to them. As the availability of technological tools increases, it is essential that mentors, instructional
resource teachers, media specialists, and or school administrators make all teachers fully aware of how to access these tools and ensure that they are readily available to teachers.

Finally, although there are a plethora of programs offering instructional support, whether as online formative assessment, test preparation, games, or instructional activity programs, teachers must use care along with their experience and expertise in deciding which of these programs has merit for their classrooms, which serve a valuable purpose, and even, which parts of a program are worthwhile. In their efforts to reach students and fill learning gaps, teachers can become overwhelmed with the number and variety of programs available. Although experienced teachers such as Mrs. Jackson can assess the benefits of multiple programs and then use what she needs from each, other, less experienced teachers may not have the ability to determine quality programs that will significantly add to their instructional program. Providing complete information on system-provided programs, such as Schoolnet, ClassScape, and Case 21, is one way administrators can ensure that teachers are able to make the most of fewer programs. Districts might also consider adding program reviews and recommendations to their websites for convenient teacher access. Technology specialists also can help provide information and recommendations for teachers. However, administrators and other personnel, as well as teachers must be vigilant not to allow classrooms to become “program heavy,” remembering that the best instructional tool in a classroom is still the teacher.

**Policy and Budget Implications**

The sum of these problems in implementation and the concerns about quality of the assessments is disquieting, given the expenditure of time and money on these programs.
While Schoolnet and ClassScape are now free to teachers, the costs of development and maintenance are not insignificant, and come out of the state budget, which is ultimately funded by taxpayers. Other programs, such as Case 21 and Achieve 3000, are purchased by the schools or districts. It is difficult to get a clear picture of the cost per student of these programs, since prices are closely-held secrets, and are adjusted depending on the number of students in the school/district, the number of years in the contract, and other factors. However, Olson (2004) reported that it was estimated that these companies would bring in $243 in 2004 alone, and that that number would continue to increase. At Turtle Creek, Mrs. Jackson used multiple programs for various assessment information, and both she and Mrs. Williams had trouble accessing their media center computers due to the number of online programs other content areas were using. At Nolan Reid, a school that struggled with resources, administrators chose to pay for their third-quarter benchmark rather than use the free ones available through Schoolnet. Given that there have not been sufficient studies to show that OFAPs are beneficial in either increasing scores or in increasing proficiency, it is unclear whether or not the significant investment of time and resources is worthwhile, especially when also considering the challenges created by inadequate technology to use the programs. The conveniences of immediate scoring, feedback, and reports as well as storage of student assessment results were not readily available to the teachers due to the challenges they faced with computer access, and this deficiency caused them to spend precious time preparing and printing the assessments and then scoring and analyzing the results.

Even so, Mrs. Williams and Mrs. Jackson seemed to think that the investment was a good one. During my observation, Mrs. Williams told her class that although she did not
have data to support her opinion, she strongly believed that using ClassScape the previous year had resulted in an increase in her students’ scores on the end-of-grade test. Regular use of formative assessment has been shown to increase student achievement (e.g. Stiggins & Dufour, 2009; Bakula, 2010), but currently there is not enough empirical evidence showing that OFAPs do so. Mrs. Williams and Mrs. Jackson also appreciated the pre-made questions, their rigor, especially in ClassScape and Achieve 3000, and the ability to easily adjust assessments to student needs. They also liked the fact that the questions taught them how to create better questions of their own. Seeing how the standards were translated into questions was another factor that both Mrs. Williams and Mrs. Jackson appreciated. Mrs. Jackson had even used the knowledge she gained in working with OFAPs to train other teachers in translating standards into questions.

The answer is not as clear for Ms. Quinn and Mrs. Kilpatrick. Ms. Quinn also liked the rigorous questions as well as the passages provided in Schoolnet. However, although both teachers worked hard to make the quarterly benchmark meaningful, it is unclear whether students benefitted as much from targeted instruction on one long assessment as they would have on shorter, teacher-made, within-instruction, paper/pencil assessments administered frequently throughout the quarter, following a research-based formative assessment model. Both teachers indicated a clear understanding of formative assessment, but the lack of clear information about the programs and program features available to them left them to use a benchmark, essentially a summative assessment, as a formative one, and the district’s intense emphasis on test scores resulted in a period of nine weeks spent on remediation, focused completely on students’ moving from one level to another, rather than on real learning.
The situation at Nolan Reid is indicative of how today’s teachers are caught between policy makers who have dictated that the end-of-grade tests will significantly impact students, teachers, and schools; assessment companies which are pushing their products as the pathways for raising student scores, if not proficiency; and doing what they signed up to do – teach students to love literature and reading.

Formative assessment experts (e.g. Black & Wiliam, 1998b; Stiggins, 2002; Heritage, 2007) claim that to be effective, formative assessment take place frequently as an integral part of ongoing instruction. The fact that the four teachers in this study did not use OFAPs on a frequent, regular basis was due, in large part, to lack of computer access. In spite of these drawbacks, all four teachers spent time in analyzing the data they gained from OFAP assessments, provided specific student feedback, based plans for instruction on the data the assessments provided, and used targeted intervention to remediate gaps.

This study also revealed that experienced teachers could use more than one formative assessment step at a time. For example, Mrs. Williams combined feedback and targeted intervention into one lesson, and Mrs. Jackson combined assessment, data analysis, feedback, targeted intervention, and reassessment into one classroom activity. Also of note was that the two teacher participants at Nolan Reid Middle School used a benchmark assessment, usually considered summative in nature, as a formative one, and applied the formative assessment model in using it. The teachers in this study all had challenges to overcome, and yet they adapted to the special circumstances of their settings and technology assets to incorporate formative assessment in to their reading instruction. They were also cognizant of using formative assessment to move students toward metacognition.
Beyond the classroom are issues that must be addressed at the state level. State administrators and policy-makers must understand the challenges schools in North Carolina face in providing online assessments of any kind. Online formative assessment has been available for many years. In fact, ClassScape, developed at the request of NCDPI, is more than twelve years old. Yet it is clear that at least some public schools in North Carolina still do not have the technology to support online assessment, whether for formative or summative purposes. Even if the state allows either paper/pencil or online assessments, consideration must also be given to whether schools with adequate online access have any advantage over those without it. Is it fair, for example, to compare students who are assessed biweekly with OFAPs with students who are unable to use OFAPs? Before more money is invested in developing online tests, priority should be given to ensuring that every school can assess students effectively online. This will require significant and ongoing investment in computers, but also essential are Internet access, infrastructure, and maintenance. Until this happens, there is little point in further development of state online assessments.

**Suggestions for Future Research**

This study was limited to the investigation of how four middle school English Language Arts teachers used online formative assessment for reading instruction in one formative assessment cycle. To date, most studies of online formative assessment have focused on their uses in undergraduate and distance learning or in specific careers, such as medicine. Further investigation in the field of English Language Arts instruction might look more in depth at how English Language Arts teachers use online formative assessment
programs for reading instruction over a full semester or school year, and the effects of using frequent during-instruction assessment using these programs.

Also of interest would be a comparative study of the growth in reading proficiency of students who are regularly assessed using an OFAP and students who are assessed on a similar schedule without an OFAP. This study might consider the speed and convenience of online programs as opposed to the more individualized teacher-created assessments and whether the advantages of the immediate scoring and feedback the OFAPs provide make a real difference to teachers’ instruction and student learning.

Of interest, too, is the difference between reading text in print and reading it online. Although some research has been done in this area (Topping, 1998) and more is ongoing, to date, comparing outcomes on formative assessment has not been studied.

Still another avenue of research is the policies guiding education today. For example, the summative testing required by the federal government, with consequences for districts, schools, and students, have been made even more high-stakes by North Carolina’s decision to use the scores in teacher evaluation. This can result in classrooms in which testing is the major concern. Schools are eager to invest large sums in programs, including OFAPs, that will ensure the success of their students on the summative tests. Do policy-makers understand how their decisions affect daily instruction? This situation begs the question as to what the English Language Arts classroom will look like in ten years if this trend continues. Will students learn about literature and read for pleasure, or will their time be spent practicing how to take tests?
Finally, while there are many OFAPs available, most of the claims of the programs’ validity come from the companies which create them. School officials often choose one program over the other based on cost or on the number of extra features made available rather than any real knowledge of the program’s effectiveness. Therefore, a study of interest would be one which compared several OFAPs and student proficiency as a result of using one or the other.

**Conclusion**

The purpose of this study was to determine the impact of using an online formative assessment program (OFAP) for reading on the planning and instruction of middle school English Language Arts (ELA) teachers. When I began this investigative journey, I thought I had few, if any, assumptions about this study and its question of how teachers used OFAPs in their planning or instruction. I realized during my research that in fact, I did have some, and it also surprised me to learn that I held certain assumptions about the teacher participants, about the programs they used, and about the settings in which they used them. I assumed, for example, that my study’s participants would all be sixth, seventh, or eighth grade teachers who had a dedicated class period for ELA instruction. I assumed that the teacher participants would have regular access to computers and the Internet. I assumed that schools in North Carolina commonly used OFAPs school wide, especially since one OFAP was available free from the North Carolina Department of Public Instruction (NCDPI). I assumed that all participants would regularly use one OFAP, either supplied by the state or purchased by the district. I assumed that teachers would administer an assessment on day and analyze
data, give feedback and provide targeted intervention on subsequent days. All of these assumptions proved to be false.

Since this study was concerned with OFAPs, I assumed that the teacher participants would have access to computers and the Internet, allowing them to regularly use online formative assessment programs. I learned that while some teachers were fortunate enough to have laptops for each of their students, the small screens made it difficult for students to access the passage and the question at the same time. The desktop computers in the media centers at Mrs. Williams’s and Mrs. Jackson’s schools were in such high demand that both teachers found it difficult to reserve them for OFAP use. Ms. Quinn and Mrs. Kilpatrick did not have computers with which to assess their students online. Although the school had a computer lab, there was no working Internet connection.

I assumed that my study’s participants would all be sixth, seventh, or eighth grade teachers, yet two of the participants were fifth graded teachers who taught at a middle school. I had experience with middle schools that included fifth graders, but they are uncommon, so I had not expected to include one in my study. Nolan Reid Middle School includes grades five thought eight, a broad spread for a middle school. Its small student body required that three teachers serve all fifth graders, so two teachers share responsibility for English Language Arts.

Another of my assumptions was that the subjects would have a dedicated class period for ELA instruction. Because of their school’s small size, Ms. Quinn and Mrs. Kilpatrick also had to teach two subjects in their seventy-minute class period. Given the overwhelming amount of content in ELA, this teaching assignment would be a challenge for any teacher.
The two teacher participants at Nolan Reid Middle School accepted this challenge and adapted.

Because Schoolnet is now available free to all students and teachers through NCDPI, I assumed that all participants would regularly use one OFAP, either Schoolnet or another one purchased by the district. I learned that some teachers were not aware of the availability of Schoolnet as a formative tool. I also learned that use of OFAPs varies within schools, and is not always school-wide.

I also assumed that teachers would use only one OFAP, primarily due to the constraints of time. I learned that the teachers in this study often used more than one online program for formative assessment purposes, and that the programs they used were not always designed or marketed as OFAPs. However, the subjects of this study understood formative assessment, and they were able to adapt it to fit their needs and the needs of their students.

I assumed that teachers would administer an assessment one day and analyze data, give feedback and provide targeted intervention on subsequent days. I learned that for Ms. Quinn and Mrs. Kilpatrick, the administration of the assessment, a benchmark, sometimes took place weeks before the targeted intervention. Mrs. Williams and Mrs. Jackson both combined parts of the formative assessment cycle, so that assessment, data analysis, feedback, and targeted intervention could take place simultaneously.

The most important lesson I learned is that these teachers were using what they had and adapting to make it work. I learned that these teachers were committed to formative assessment, but had to adjust what they knew to be its ideal implementation to circumstances beyond their control. Ms. Quinn and Mrs. Kilpatrick would have liked to assess their
students with an OFAP on a regular, formative basis. Without the technology tools they needed to do so, they adapted what they did have and used the district benchmark, as well as their own assessments, for the data on student performance they needed. Mrs. Williams and Mrs. Jackson did not give as many online formative assessments as they might have liked due to the small screens on the student laptops and the competition for the media center computers. They, too, adapted, by printing assessments on paper and combining online delivery with printed copies. The teacher participants were also aware that reading online is different than reading on paper, and also, since the state summative assessment is still administered on paper, they felt an obligation to their students to provide practice for that medium.

It is clear from this study that teachers do not expect OFAPs to relieve them from work. They do not expect OFAPS to be the only assessment tool they use, nor do they expect OFAPs to teach reading for them. These four teachers were instead constantly looking for ways to make learning more meaningful and accessible for their students. They were constantly analyzing data, both written and verbal and both formal and informal. They were continuously refining their practice through data analysis and through reflection and refraction. Throughout the study, these four teachers proved to be constructivists (Merriam, 2009), creating meaning with their students from the text they read and the questions about the text that they answered. They worked hard to help their students move beyond the role of receivers of information to that of self-motivated learners and thinkers. More research needs to be conducted on how OFAPs can support instruction and on their effectiveness, as well as which programs are worthwhile. However, based on this study, it is clear that OFAPs can be
helpful as a formative assessment tool in the English Language Arts classroom, offering a
variety of features to allow teachers more time to work directly with students to improve
their skill in reading. However, as is the case with much of educational research we end with
yet another question: Will the present policies in education allow OFAPs to be used as
instructional tools or will they too, become one more instrument in the sacred quest for
testing excellence?
References


Beach, R. & Swiss, T. (2010). “Literary theories and teaching of English language arts.” In D. Lapp & D. Fisher (Eds.), Handbook of research on teaching the English language arts: Co-sponsored by the international reading association and the national council


Hayes, A. M. (2013, November).*Online formative assessment.* PowerPoint presented at the National Council of Teachers of English Annual Convention, Boston, MA.


Stiggins, R. (2009). Assessment for learning in upper elementary grades: students learn more when they use assessments to evaluate their own learning and they have greater feelings of efficacy about their academic abilities. *The Phi Delta Kappan, 90*(6), 419.


Appendix A

North Carolina State University
INFORMED CONSENT FORM for RESEARCH

Title of Study: FROM ASSESSMENT TO INSTRUCTION: THE IMPACT OF ONLINE FORMATIVE ASSESSMENT IN READING ON TEACHERS’ PLANNING AND INSTRUCTION IN THE MIDDLE SCHOOL ENGLISH LANGUAGE ARTS CLASSROOM
Principal Investigator: Ann Hayes
Faculty Sponsor: Dr. Carol Pope

What are some general things you should know about research studies?
You are being asked to take part in a research study. Your participation in this study is voluntary. You have the right to be a part of this study, to choose not to participate or to stop participating at any time without penalty. The purpose of research studies is to gain a better understanding of a certain topic or issue. You are not guaranteed any personal benefits from being in a study. Research studies also may pose risks to those that participate. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something in this form it is your right to ask me for clarification or more information. Please print a copy of this consent form for your records. If at any time you have questions about your participation, do not hesitate to contact Ann M. Hayes.

What is the purpose of this study?
The purpose of the study is to discover how the use of data from an online formative assessment program impacts teacher planning and instruction for reading instruction in the middle school English Language Arts classroom. Since so many schools use this type of assessment, it is important to know whether it makes a difference in teacher instruction.

What will happen if you take part in the study?
You will grant me permission to observe you teaching a follow-up lesson based on student data from an online formative assessment tool.

You will be asked to participate in two audio-recorded interviews. These interviews will take place before and after the observation.

You will also provide the following documents: instructional objectives, germane notes, and instructional handouts for the observed lessons.

You will be asked to review the transcripts of the audio recordings and the notes I have taken to ensure accuracy.

Risks
There is minimal risk to you for participating in this study.
Benefits
To date, the impact of online formative assessment on teacher instruction in reading has not been studied. Since these tools are becoming common-place in the classroom, it is important for teachers, administrators, state and federal policy-makers, as well as pre-service teacher program administrators to understand their impact on actual instruction. Information from this study might influence professional development decisions, the types of formative assessment tools selected by districts, and the uses other teachers make of them. You may also benefit from the self-reflection and sharing of insights during the study, and may gain a deeper understanding of your own pedagogy.

Confidentiality
The information in the study records will be kept confidential to the full extent allowed by law. Data will be stored securely in my computer in a password-protected file or in a locked file cabinet in my home, to which only I have access. No reference will be made in oral or written reports which could link you to the study. Your school and location will not be identified.

Compensation
There is no compensation for participating in this study, other than my sincere gratitude for helping me in my research.

Participation
Participation in this study is not a requirement of your employment at Lee County Schools, and your participation or lack thereof, will not affect your job.

What if you have questions about this study?
If you have questions at any time about the study or the procedures, you may contact the researcher, Ann Hayes, at akhayes@ncsu.edu or 919-601-6278.

What if you have questions about your rights as a research participant?
If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Deb Paxton, Regulatory Compliance Administrator, Box 7514, NCSU Campus (919/515-4514).

Consent To Participate
“I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may choose not to participate or to stop participating at any time without penalty or loss of benefits to which I am otherwise entitled.”

Subject's signature__________________________________Date ___________________

Investigator's signature________________________________Date ___________________
Appendix B: Interview Protocol: Pre-Observation

Date of Interview:

Place of Interview:

Teacher Pseudonym:

1. What is your teaching experience?

2. What is your teaching assignment this year?

3. What is your experience with formative assessments (both formal, such as the online programs, and informal, such as exit tickets, quizzes, questioning, etc.)?

4. What kinds of formative assessment have you used in your classroom in the past?

5. What do you consider successful formative assessment? In other words, can you describe what a successful formative assessment model looks like?

6. What kinds of information do you need from a formative assessment?

7. Please describe the online program that you are currently using.

8. Have you ever used another online formative assessment program? If so, how did it differ?

9. How do you feel about the specific program you are using?

10. What are the advantages? How has it helped students, and how has it helped you in planning/teaching?

11. Disadvantages? Are there problems with the program? If so, what are they? Are there things you would change?

12. Are there intrinsic limits to an online program? If so, what are they?
13. What information have you been able to glean from your formative assessments?


15. In what ways do you give feedback to students after a formative assessment?

16. Do you adjust your instruction in any way after assessing the students in a formative way?

17. How do you use the data from the formative assessment to guide your planning and instruction?

18. Have you seen any changes in your students as a result of using an online formative assessment program? If so, what are they?

19. Have any of these changed as a result of your using an online formative assessment system? If so, please explain in what way(s).
   a. Short-term planning
   b. Long-term planning
   c. Whole group instruction Small group instruction - Differentiation - Individual instruction –
   d. Feedback to students (type, method, frequency) -
   e. Feedback to parents (type, method, frequency) - Assessments (type, method, frequency) –
   f. Thinking about assessment/grading/feedback/data –

20. Do you plan to use handouts or other documents or notes tomorrow? If so, what?

21. Is there anything I haven’t asked that you wish to tell me about formative assessment, the online program, your class, or your concerns?

22. Do you have any questions about the observation, data collection, or any other questions related to the study?
Appendix C: Interview Protocol: Pre-Observation

Date of Interview:

Place of Interview:

Teacher Pseudonym:

1. How did the modified instruction go? Was it what you expected?

2. Would you change anything if you were doing it again? If so, how?

3. Why did you choose these strategies?

4. Does the data for this assessment and/or the lesson in any way change your ideas of how you will approach your next unit of study?

5. Would you change your original instruction (prior to the assessment)? If so how?

6. How successful do you think instructional modification based on formative assessment data is?

7. Has your instruction changed as a result of the online formative assessment program and the data it generates? If so, how and why?

8. Does the data from one assessment affect how you approach future instruction that is not associated with that assessment? If so, how?

9. Overall, what are the benefits of using an online formative assessment program?

10. What are the disadvantages?

11. What do you wish the online formative assessment program could do that it does not do?

12. Is there anything you wish to add to the information you have given me?