2024

A Program Evaluation Of I-Ready Implementation In A North Carolina Middle School

Juliana Irene Owen Thompson
William & Mary - School of Education, jiothompson08@gmail.com

A PROGRAM EVALUATION OF i-READY IMPLEMENTATION IN A NORTH CAROLINA MIDDLE SCHOOL

A Dissertation

Presented to

The Faculty of the School of Education

The College of William and Mary in Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

By

Juliana Irene Owen Thompson

A PROGRAM EVALUATION OF i-READY IMPLEMENTATION IN A NORTH CAROLINA MIDDLE SCHOOL

By

Juliana Irene Owen Thompson

Approved December 8, 2023, by

STEVEN M. CONSTANTINO, Ed.D.
Committee Member

STEPHEN S. BAKER, Ed.D.
Committee Member

MARGARET E. CONSTANTINO, Ph.D.
Chairperson of Doctoral Committee

Dedication

This dissertation is dedicated with love and gratitude to my family for their constant encouragement, steadfast support, and joyful optimism.

Acknowledgments

Thank you to my devoted friends and family who encouraged me to "keep on keeping on" to finish this journey. You believed in me when I doubted myself, steadied me when life threw curveballs, and inspired me to remember why the work we do every day matters.

To my professors and dissertation committee, thank you for pouring into me and guiding me throughout this process. Thank you Dr. P. Constantino, Dr. S. Constantino, and Dr. Baker for your willingness to serve on my committee. Dr. P. Constantino, thank you for always answering when I called with questions or dilemmas, for your mentorship, and for your friendship.

To the educators who participated in this program evaluation, thank you for your insights, which helped me write this dissertation and will continue to guide our implementation work for years to come. Your work makes a difference in the lives of our students.

To my parents, Sam and Deborah Owen, thank you for teaching me what is most important about life: loving God and loving each other. Not only did you live your faith in Jesus and pass it on to me, but you taught me how to work hard each day and never give up. You taught me how stand up for what's right – for who's right – for ALL the children. And you taught me that we never know how many days we have left, so make the best of each one, fight the good fight, and keep having fun.

To my husband, Scott Thompson, thank you. You inspired me. You always encouraged me to follow my dreams. You supported me, so I could earn my doctorate. You pushed me to do my best and think through all angles of any situation. Most importantly, you gave me the most amazing gift possible: our son, Alex Thompson. Alex, you bring so much joy, love, and humor to our world, and I am so very proud of you. Scott and Alex, I will never be able to thank you enough for your support, encouragement, and love that made this journey possible.

Table of Contents

List of Tables

List of Figures

Abstract

Aurelius School District used Curriculum Associates' i-Ready program, which was philosophically aligned to the science of reading and included i-Ready Diagnostic, i-Ready Instruction, and Teacher Toolbox, to improve Cole Middle School students' reading achievement. Because of the high monetary cost of i-Ready, the school district was considering changing or ending this program. This pragmatic, mixed methods design, CIPP (Stufflebeam, 2003) program evaluation analyzed the fidelity of implementation of the i-Ready program through descriptive statistics (frequencies); assessed the effectiveness of the i-Ready program in increasing student achievement (on the Diagnostic, NC Check-Ins, and NC EOG) through descriptive statistics (mean, standard deviation, and range), Repeated Measures ANOVA, and correlations analyses; and determined stakeholders' perceptions of the value of the program to inform recommendations to continue, modify, or discontinue usage of the i-Ready program (through surveys and interviews prior to qualitative review using open and axial coding). Based on the review of related literature, the i-Ready program was a coherent and adaptive assessment system and an effective instructional intervention designed to increase adolescent literacy, which occurred in grade levels where the program was implemented with fidelity. Because it provided staff and students with timely, understandable, and helpful instruction and data that helped students grow, recover unfinished learning, and improve their achievement, the i-Ready program was valuable and worth its cost. As the implications for positive social change, it is recommended that the i-Ready program continues in its current form with a few adjustments to motivate and encourage teachers and students to result in increased usage and engagement.

A PROGRAM EVALUATION OF i-READY IMPLEMENTATION IN A NORTH CAROLINA MIDDLE SCHOOL

CHAPTER 1

INTRODUCTION

This study was a program evaluation of i-Ready implementation in one middle school in a North Carolina school district. I sought to determine fidelity of usage of diagnostic assessments, personalized instruction, and teacher resources in the implementation of the i-Ready program. Additionally, I assessed the effectiveness of the i-Ready program in increasing student achievement and determined stakeholders' perceptions of the program to inform recommendations to continue, modify, or discontinue usage of the i-Ready program.

Background

What is one of our greatest educational responsibilities as educators, parents, and a collective society? The answer is simple to say, yet complex to do—that is to ensure that every child learns to read well. In a keynote speech to the American Library Association, then-U.S. Senator Barack Obama (2005) offered an assessment of the power and importance of developing reading skills that holds true today: "Reading is the gateway skill that makes all other learning possible, from complex word problems and the meaning of our history to scientific discovery and technological proficiency" (p. 50). Reading allows people to acquire knowledge, engage culturally, participate in democracy, and succeed economically (Castles et al., 2018), and it cultivates empathy, social imagination, and motivation (Guthrie et al., 1996; Johnson et al., 2013; Lysaker et al., 2011).

Sadly, too many children and later, adults, are not learning to read proficiently. It is estimated that 1 of every 5 people globally are completely illiterate, and another 3 billion people

struggle to read and write at a basic level (World Literacy Foundation, 2018). According to the most recent National Assessment of Educational Progress (NAEP), only 33% of fourth graders and 29% of eighth graders (National Center for Education Statistics [NCES], 2022), and only 37% of 12th graders in the United States are reading proficiently (NCES, 2019). The direct economic costs of global illiteracy are estimated at over $1 trillion dollars annually, but indirect costs of increased crime and welfare dependency, decreased health outcomes, disruptions in social cohesion, and loss of personal opportunities and self-worth are enormous societal burdens (World Literacy Foundation, 2018).

Because of the importance of literacy, it is no surprise that a substantial body of research has been conducted on how children learn to read, what effective components of reading instruction are, and when and where to intervene when children have reading difficulties. Although educators understand the importance of literacy, with some believing teaching students to read is their most fundamental responsibility (Moats, 2020), they disagree about how to teach reading and what resources are most effective for students, especially for non-proficient adolescent readers at the secondary level. Not all pedagogical approaches, programs, and products are supported by sound research (e.g., What Works Clearinghouse), and because schools face the need to teach many standards with a finite amount of time and money, it is important that what they choose is high-quality and effective. While teaching foundational reading skills is not part of grade-level standards after elementary school, teaching those grade-level standards without addressing essential literacy skills creates inequity, does not provide equal opportunities for all students to succeed, will not close achievement gaps, and will lead to academic failure (Kazakoff, 2021).

One idea to meet the reading challenge is to use a blended learning approach with technology personalizing what non-proficient adolescent readers need to close their individual reading gaps and become proficient readers. One such commercial program that is touted as fulfilling this need is i-Ready, a reading program used widely in schools across the United States (Curriculum Associates, 2021d). The goal of this program evaluation in one North Carolina middle school was to evaluate the fidelity of i-Ready implementation, determine i-Ready's impact on student achievement, and investigate stakeholders' perceptions of i-Ready to determine if the i-Ready program is worth its cost in terms of improved reading skills for students.

Program Description

North Carolina's NAEP scores mirror the national scores with only 32% of fourth graders and 26% of eighth graders reading proficiently, and these scores dropped from the 2019 results (NCES, 2022). Unfortunately, Aurelius School District (ASD) [pseudonym] students were performing lower than the state average on the state's End of Grade (EOG) reading tests. For the last 5 years (2016-2021), fewer than half of ASD students in Grades 4-8 have demonstrated reading proficiency. As of 2020-21, only 35.7% of students in Grades 4-8 were reading proficiently. Widespread low reading proficiency in Grade 4 is an indicator that foundational reading instruction, including fluency and automaticity, should be examined (Chall & Jacobs, 2003). Until 2021-22, ASD espoused a balanced literacy philosophy emphasizing tenets of whole language and teacher autonomy over foundational skills and standardized curriculum. Anecdotal and standardized test data indicated that inequities have emerged among classrooms as teachers naturally had different years of experience and levels of efficacy and varying participation in professional development opportunities. A lack of systematic reading curriculum, instruction,

and assessment can create inequitable opportunities for students in learning foundational reading skills, as they are completely dependent on each assigned teacher's abilities to locate and facilitate their own materials. Lengthy school closures from the COVID-19 pandemic also exacerbated inequities in students learning foundational reading skills. ASD was able to purchase the i-Ready reading program using the Elementary and Secondary Schools Emergency Relief (ESSER) funds, which were designed to address COVID-19 recovery needs, to close gaps in reading proficiency.

Context

This program evaluation was focused on the implementation of i-Ready in the middle school of a small public school district with approximately 350 staff members, 2,800 students, one pre-kindergarten center, two elementary schools, one middle school, two high schools and one alternative learning center. In 2021, student achievement declined significantly with only 35% of students proficient in reading in Grades 3-8 (see Appendix A). ASD is in a small town in a rural area. During 2022-23 at ASD's Cole Middle School, 61% of students were from economically disadvantaged families; 27% were chronically absent; 57% were White; 28% were Black; 8% were Hispanic; 4% were two or more races; and 3% were Asian. Cole Middle had 14% beginning teachers. Student achievement declined significantly in 2021 with only 31% proficient in math and 39% proficient in reading in Grades 6-8 (see Appendix A for more district data).

ASD began implementation of Curriculum Associates' i-Ready program as a comprehensive resource to improve student achievement in K-8 reading and mathematics in May 2021. In addition to collaboration in robust professional learning communities (PLCs) using consistent pacing guides, staff needed standard-specific, North Carolina End of Grade (NC

EOG)-aligned student diagnostic data to inform instructional decisions and to personalize instruction throughout the school year. Implementing i-Ready, however, came with a hefty price tag. In 2021-22, ASD paid $73,093.50 for the Diagnostic, Instruction, Teacher Toolbox, and Professional Development, which equated to $41.29 per student; that cost increased for 2022-23 to $94,382.22, which equated to $53.30 per student.

Description of the Program

The i-Ready suite included i-Ready Diagnostic (i.e., universal screeners and benchmarks); i-Ready Instruction (i.e., programs primarily implemented online with personalized lessons based on students' performance on the i-Ready Diagnostic); and Teacher Toolbox (i.e., whole class, small group, and individual lesson plans for teachers to address specific gaps in student learning). Both i-Ready Instruction and Teacher Toolbox integrated with the results from i-Ready Diagnostic to provide a seamless connection between student data and instruction and provided normative- and criterion-referenced data (Curriculum Associates, 2021b).

Research findings as presented by the developer of i-Ready claimed that i-Ready was an effective online supplemental program in reading and mathematics because (a) students using i-Ready Instruction as intended showed greater learning (i.e., 46% more in reading and 38% more in math) than those not using i-Ready; (b) on average, students in special populations using i-Ready Instruction as intended showed greater learning than those not using i-Ready; and (c) students in Grades 1-8 made positive, statistically significant gains after controlling for prior achievement, indicating that i-Ready was an evidence-based intervention that met Every Student Succeeds Act (ESSA) Level 3 Promising Evidence Standards (Curriculum Associates, 2020,

2021c).[1]. Meeting Level 3 meant the evidence was supported by one or more well-designed and well-implemented correlational studies and had statistical controls for selection bias (U.S. Department of Education, n.d.).

The underlying theory of action for the program (see Table 1) was that the stated priorities emerging from the situation could be addressed (and external factors overcome) if the inputs engaged in the outputs to achieve the outcomes. Inputs, or resources, were the district's investments of money, instructional time, technology, teaching staff, specialized staff and administrators, and consultants. Outputs included the district's participation in the following activities: professional development, professional learning communities, pacing guides, i-Ready Diagnostic, and i-Ready Instruction. Outcomes included increased staff understanding and efficiency; increased data-based decision-making with districtwide alignment, coherence, and consistency in reading curriculum, instruction, assessment, and intervention; improved i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOG scores; increased student reading proficiency; and eliminating COVID-19 learning loss.

[1] This research was funded and published by Curriculum Associates, the publisher of i-Ready; thus, the findings should be read with caution due to the lack of independent verification of the findings.

Table 1

Theory of Action for the i-Ready Program in Aurelius School District (ASD)

Activity	Theory of Action
i-Ready Diagnostic	If ASD teaching staff uses ASD-issued technology to administer the i-Ready Diagnostic to students three times a year, then ASD students will have instruction that is personalized to their specific needs.
i-Ready Instruction	If ASD teaching staff facilitates ASD students' use of i-Ready Instruction as intended (i.e., use personalized online program 30-49 minutes per week, pass 70% or more lessons, and have face-to-face, teacher-led, small group interventions from Teacher Toolbox), then ASD students will have improved i-Ready Diagnostic performance and improved NC Check-Ins/NC Check-Ins 2.0 and NC EOG achievement.
i-Ready Program	If ASD implements the i-Ready program with fidelity from 2021 to 2024, then ASD should eliminate students' COVID-19 learning loss by finishing previously unfinished learning in reading and improving student achievement to pre-pandemic levels (or higher) and ensure students are reading on grade level.

Note. NC = North Carolina; EOG = End of Grade assessment

Overview of the Evaluation Approach

The guiding paradigm for this evaluation was pragmatic, which was responsive to the needs of stakeholders but mainly focused on usefulness in improving the program as it was implemented in the stakeholders' specific context (Mertens & Wilson, 2019). The evaluation results should be meaningful and useful to the stakeholders who commissioned the evaluation. Consequently, the program evaluation was designed to investigate the fidelity of i-Ready implementation, to determine if i-Ready is meeting ASD's goal of increasing student reading achievement, and to establish if the program seems beneficial to specific stakeholders using it.

Program Evaluation Model

The meso-level, linear program logic model describes the initial full-scale i-Ready implementation plan in K-8 reading. Threads of related program elements flow horizontally across the logic model, and some participation and outcome elements merge into larger boxes

across the logic model. The program evaluation of this study focused on elements of the logic model that are highlighted in yellow.

This program evaluation followed Stufflebeam's (2003) Context, Input, Process, and Product (CIPP) Evaluation Model. Mertens and Wilson (2019) noted that CIPP is "a time-tested approach to use-focused evaluation" (p. 104). The approach recognizes the need to consider stakeholder input and stakeholders' need for information.

Context. In context evaluation, needs, problems, assets, and opportunities are evaluated to define goals, priorities, and desired outcomes (Mertens & Wilson, 2019; Stufflebeam, 2003). Figure 1 includes a brief identification of the current ASD situation, priorities, assumptions, and external factors existing prior to program implementation. To assist visually, all aspects of context appear in blue boxes.

Situation. At the beginning of program implementation, ASD's reading situation was troubling due to misalignment of curriculum, instruction, and assessment. Data indicated that many ASD students were not reading proficiently and that reading proficiency on state summative assessments had decreased. The COVID-19 pandemic exacerbated unfinished learning and increased staff and student absences.

Priorities. In implementing i-Ready, ASD prioritized improving its reading situation by creating a coherent assessment system with personalized instruction. The district desired to ensure all students were reading proficiently and all COVID-19 learning gaps were closed. ASD identified that fidelity of implementation was essential.

Assumptions. To implement a new system, staff needed effective, on-going training. ASD assumed that professional development provided on i-Ready would be robust, effective, and targeted to its needs and that professional learning communities would support the

professional development. In initially deciding to implement i-Ready, ASD leadership assumed that i-Ready was internally aligned, would be used as intended, and was aligned to NC Check-Ins/NC Check-Ins 2.0 and NC EOG assessments.

External Factors. There were other contextual factors that may influence reading outcomes. For example, North Carolina has mandated that all K-5 teachers participate in Language Essentials for Teachers of Reading and Spelling, commonly known as LETRS (Moats & Toulman, 2022) professional development. Also, ASD implemented *Wonders 2023*, a new reading curriculum in K-5. Positive reading gains could be attributed to either LETRS or *Wonders*. Negative reading gains could also be related to COVID-19-related absences of staff of students.

Input. In input evaluation, alternative approaches, competing action plans, participant characteristics, staffing plans, and budgets are evaluated to determine feasibility and potential cost-effectiveness, choose among competing plans, write funding proposals, allocate resources, assign staff, and schedule work (Mertens & Wilson, 2019; Stufflebeam, 2003). Figure 1 includes a more detailed description of the implementation plan through inputs or investments. To assist visually, all aspects of input appear in pale yellow boxes. Investments in the logic model include i-Ready consultant staff, ASD specialized staff and administrators, ASD teaching staff, ASD technology, instructional time, and money.

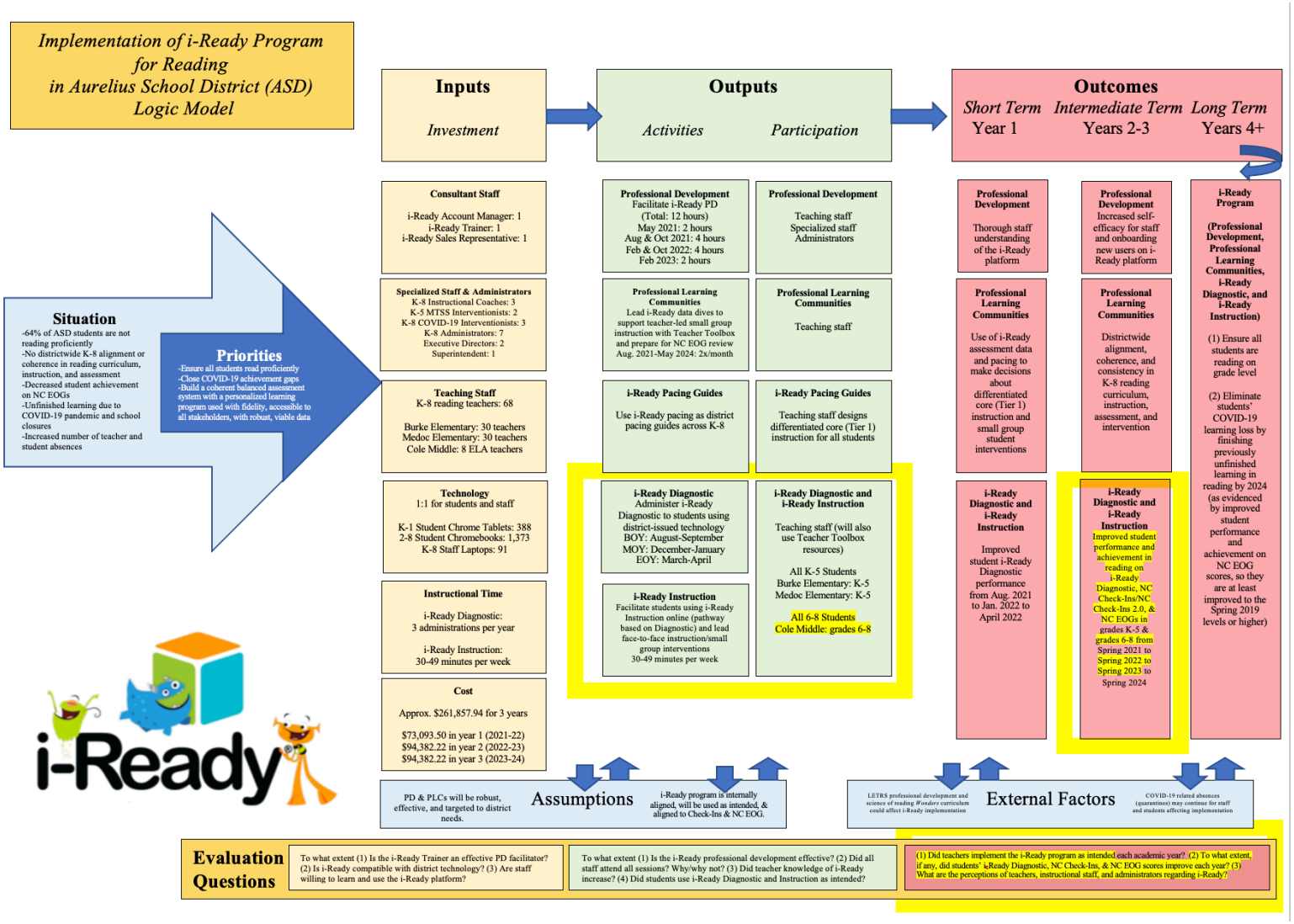
Process. In process evaluation, implementation of plans is evaluated to help staff make needed revisions in activities, to judge performance, and to interpret outcomes (Mertens & Wilson, 2019; Stufflebeam, 2003). Figure 1 includes a more detailed description of the implementation plan through outputs or activities and participants. To assist visually, all aspects of process appear in light green boxes.

This program evaluation was situated mainly as a process evaluation, which is also sometimes referred to as an implementation evaluation (Mertens & Wilson, 2019). The evaluation used mixed methods with quantitative review of Year 2 student usage data (processes) and achievement outcomes (products) to provide a periodic progress report on program implementation as well as quantitative and qualitative analysis of stakeholder perception data (processes) to obtain their assessments of the program's progress. ASD used the process evaluation findings to strengthen staff activities and strengthen the program design.

Product. In product evaluation, identification, and assessment of both short-term and long-term as well as intended and unintended outcomes are evaluated to help staff keep focused on achieving desired outcomes and gauge the success of the program in addressing needs (Mertens & Wilson, 2019; Stufflebeam, 2003). Figure 1 includes a more detailed description of the specific and measurable short-term outcomes (Year 1), intermediate-term outcomes (Years 2-3), and long-term outcomes (Years 4 and ongoing) as measured by criteria indicated in the evaluation questions. To assist visually, all aspects of process appear in pink boxes. Although the program evaluation was situated mainly as a process evaluation, it also was a product evaluation as it examined some intermediate-term outcomes. The theory of action (see Table 1) for this model also indicated that short-term outcomes lead to intermediate-term outcomes lead to long-term outcomes. Figure 1 shows the logic model for ASD's program implementation of i-Ready.

Figure 1

Aurelius School District Program Implementation of i-Ready Logic Model



Purpose of the Evaluation

In September 2024, ESSER funds expire, and ASD had to determine whether to terminate i-Ready or other programs to keep i-Ready. This i-Ready program evaluation sought to determine the degree to which teachers implemented i-Ready as intended by its developers (i.e., all students participated in the i-Ready Diagnostic three times per year; all students participated in i-Ready Instruction 30-49 minutes per week; and all students passed at least 70% of lessons in i-Ready Instruction); the amount students grew on the i-Ready Diagnostic, NC Check Ins/NC Check Ins 2.0, and NC EOG; and the value teachers and other staff found in i-Ready. The i-Ready program would be valuable and possibly worth its cost if it provided staff and students with timely, understandable, and helpful instruction and data that help students grow, recover unfinished learning, and improve their achievement. Results of the program evaluation were intended to provide feedback to stakeholders regarding the program's continuation, modifications, or termination.

Focus of the Evaluation

The two main areas of focus of this program evaluation were process and product. For process, the evaluation relied on quantitative and qualitative data to consider the fidelity of activities by participants and perceived benefit of the program for stakeholders using it as an instructional resource. For product, the evaluation examined student outcomes after participation in the program. Outcomes that were the focus of this evaluation were measured by formative and summative assessments which are highlighted in yellow in the logic model (see Figure 1).

Evaluation Questions

The theory of action for the i-Ready program presumed that staff engagement in professional development and professional learning communities, usage of i-Ready pacing

guides, administration of the i-Ready Diagnostic as intended, and facilitation of student usage of i-Ready Instruction as intended would lead to improved student performance and achievement on the i-Ready Diagnostic and NC EOGs, ensure all students are reading on grade level, and eliminate students' COVID-19 learning loss by finishing previously unfinished learning in reading by 2024. This program evaluation addressed the following questions that were designed to assess the fidelity of implementation (process evaluation) and attainment of intermediate term outcomes (product evaluation):

1. To what extent did teachers in the selected school implement with fidelity the i-Ready program (i.e., i-Ready Diagnostic and i-Ready Instruction) in the 2022-23 academic year?
2. To what extent, if any, did student reading proficiency improve during the 2022-23 academic year as measured by the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOGs?
3. What were the perceptions of teachers, instructional staff, and administrators regarding i-Ready in terms of guiding program elements such as the following: (a) the value of student data generated by the i-Ready Diagnostic, (b) the value of the i-Ready Instruction as an instructional resource, and (c) the extent to which i-Ready is having an impact on student achievement?

Definitions of Terms

Curriculum Associates: An educational resources company that owns i-Ready Learning, Teacher Toolbox, Magnetic Reading, Ready, and PHONICS for Reading (Curriculum Associates, 2023a).

Fidelity: The degree to which teachers and others implement programs *as intended* by the program developers (Dusenbury et al., 2004).

i-Ready Diagnostic: An adaptive assessment available for reading and mathematics in Grades K-8, recommended to be administered at three points during the school year (fall, winter, and spring), and designed to provide teachers with actionable insight into student needs. It offers a complete picture of student performance and growth, eliminating the need for multiple redundant tests. By adapting to student responses and assessing a broad range of skills—including skills above and below a student's chronological grade—the i-Ready Diagnostic pinpoints student ability level, identifies the specific skills students need to learn to accelerate their growth, and charts a personalized learning path for each student (Curriculum Associates, 2021b).

i-Ready Instruction: Personalized instruction delivering online lessons intended to motivate students on their paths to proficiency and growth. Driven by insights from the i-Ready Diagnostic, i-Ready's online K–8 lessons in reading and mathematics provide tailored instruction that meets students where they are in their journey and encourages them as they develop new skills. It is intended for students to use 30–49 minutes per week per subject (Curriculum Associates, 2023a).

Literacy: The ability to identify, understand, interpret, create, communicate, and compute, using written materials associated with varying contexts. Literacy involves a continuum of learning in enabling individuals to achieve their goals, to develop their knowledge and potential, and to participate fully in their community and wider society (UNESCO Institute for Statistics, 2023).

Personalized Learning: A variety of educational programs, learning experiences, instructional approaches, and academic-support strategies that are intended to address the distinct learning needs, interests, aspirations, or cultural backgrounds of individual students and provide customized learning experiences for each student (Great Schools Partnership, 2015).

Teacher Toolbox: A digital collection from which teachers can draw relevant, standards-based resources in an easy-to-navigate platform that provides a complete set of instructional resources to differentiate Grades K-8 reading and mathematics instruction for students above, at, and below grade level as recommended by i-Ready Diagnostic results (Curriculum Associates, 2023a).

CHAPTER 2

REVIEW OF RELATED LITERATURE

The most important responsibility of K-12 educational institutions is teaching students to read well because reading affects all other academic achievement and is associated with social, emotional, economic, and physical health (Moats, 2020). Though various philosophies of teaching reading exist, decades of research have shown curriculum grounded in the science of reading to be more effective (National Institute of Child Health and Development [NICHD], 2000). When curriculum aligned to the science of reading is coupled with coherent assessment systems and personalized interventions, student achievement, especially for non-proficient adolescent readers, increases (Kazakoff, 2021). Although there is much research knowledge about learning to read, it has not translated into increased public understanding (Castles et al., 2018), and it is difficult to make data-based decisions on effective reading programs and practices because often there is either no evidence of effectiveness or the evidence is of poor quality (Petscher et al., 2020). One area of needed research is independent studies of i-Ready's efficacy in improving middle school students' reading achievement.

The purpose of this review of related literature is to describe research illustrating philosophies of teaching reading, coherent and adaptive assessment systems, and personalized interventions and to explore how a specific reading program can be evaluated at the school level. Fidelity of implementation is also discussed as an important component of program evaluation. The literature review concludes with a description of i-Ready for reading, the focus of this program evaluation.

Philosophies of Teaching Reading

For decades, passionate debates about how to teach students to read have been described as the "reading wars" (Castles et al., 2018; Chall, 1967; Kim, 2008; Yaden et al., 2021). Two beginning reading approaches faced off in head-to-head combat: phonics and whole language. In recent years, newer approaches have combined elements of phonics and/or whole language as well as emphasized other aspects of literacy instruction.

Phonics

In the phonics approach, a skills-based process, educators explicitly and systematically teach children to sound out words based on how they are spelled (Chall, 1967; Flesch, 1955). Students start by learning the alphabetic code, which is the relationship between sounds and letters, and progress to decoding words in isolation before reading words in a longer text (Chall, 1967). Teachers use lists of words and decodable texts (i.e., books with predictable patterns and simple vocabulary) with beginning readers until students are more adept with their phonics skills, when teachers progress to discussing word meanings and gradually release students to reading stories with more meaningful plots (Chall, 1967). In 2000, the National Reading Panel declared that phonics instruction was crucial to teaching beginning readers (NICHD, 2000). Critics believe that limiting beginning readers to phonemes and then to decodable texts restricts their prosody and fluency development (Lorimor-Easley & Reed, 2019).

Ehri et al. (2001) conducted a quantitative meta-analysis to evaluate the effectiveness of systematic phonics instruction on learning to read as compared to unsystematic or no-phonics instruction. Researchers found a statistically significant difference between mean scores of the experimental and the control groups as reported using Cohen's *d* values. The overall effect of phonics instruction on learning to read was moderate ($d = 0.41$), but effects continued after

instruction ended. When phonics instruction began earlier, effects were larger (d = 0.55) than after first grade (d = 0.27). Phonics instruction helped decoding, word reading, spelling, and text comprehension in many readers. Researchers used 66 treatment-control comparisons from 38 experiments. They examined methodological strengths and weaknesses of the studies including random assignment (37% of the studies) versus nonequivalent groups, presence versus absence of pre-treatment differences between the phonics and control groups (five of the studies had pre-treatment differences), and large versus small sample sizes (comparisons ranged from 20 to 320 students). Neither random assignment nor sample size differentially influenced mean effect sizes; however, pretreatment differences did influence the effect size, which suggests that weaker designs may have hindered the detection of effects.

Systematic phonics instruction helped children learn to read better than all forms of control group instruction, including whole language, and the researchers indicated it should be implemented as part of literacy programs for beginning reading as well as remediation for reading difficulties; however, the studies they used neither fully described the features included as part of the phonics instruction, nor included information about the full program used to teach reading. Lastly, the meta-analysis only included published studies, which could mean that the pool of studies is biased and unrepresentative of a population of mostly unpublished studies finding no effects.

Research converges to show that explicit and systematic phonics instruction to help beginning readers understand how letters are linked to sounds to form letter-sound correspondences and spelling patterns and to help them learn to apply their knowledge as their read is valuable and should be an essential part of reading programs (NICHD, 2000). Phonics, however, is a means to an end, rather than an end itself, and should be integrated with other

reading instruction in phonemic awareness, fluency, and comprehension strategies to create a complete reading program (NICHD, 2000). More research is needed to define how many months or years a phonics program should continue as well as how to balance standardizing the phonics instructional sequence with also maintaining teacher interest and motivation and differentiation for students.

Whole Language

In the whole language approach, a meaning-based process that focuses on students learning entire words, educators create a literacy-rich, motivating, and engaging environment for children and emphasize meaning of texts through experiences over the sounds of the letters (Goodman, 1967; Reyhner, 2020; Smith, 1971). Consequently, researchers do not advocate for breaking words into decodable parts; instead, they explain that reading is a "psycholinguistic guessing game" in which readers pick up graphic cues to words, form perceptual images, and read by a selection process consistently using long-term and short-term memory (Goodman, 1967). Teachers expose students to authentic literature regardless of students' ability to read the words on the pages of the book as they construct their own personal meanings for the text using their prior knowledge to interpret the meaning of what they read (Reyhner, 2020). Whole language instruction depends on teachers having a strong base in theory and research and using their knowledge to make practical decisions and plan instruction (Goodman, 1989). Critics believe that without decoding and encoding skills, students practice merely compensatory strategies (e.g., relying on picture cues) and miss out on valuable instructional time (Lorimor-Easley & Reed, 2019).

Hattie (2009), who synthesized 138 influences across 800 meta-analyses and ranked their effect sizes related to student achievement from very positive to very negative effects with a

desired effect of 0.40 or greater, found the whole language approach to have only a 0.06 effect size (as opposed to phonics instruction with a 0.60 effect size). According to this calculation, whole language instruction has no real impact on student learning. Hattie (2009) described effect sizes between 0.0 and 0.15 as "what students could probably achieve if there was no schooling (and is estimated from the findings in countries with no or limited schooling)" (p. 20). Hattie (2009) also stated that "whole language programs have negligible effects on learning to read—be it on word recognition or on comprehension" (p. 138). Recently, Hattie's findings on whole language, which derived from four meta-analyses that involved a total of 64 studies, have been questioned on the grounds that he used an incomplete definition of whole language in locating the studies, included one meta-analysis that should not have been included since it used studies that pre-dated the whole language movement, averaged the effect sizes of the four meta-analyses (even though the two most recent meta-analyses had diametrically opposed findings at 0.65 and -0.65), and potentially was biased against whole language in favor of systematic phonics and direct instruction (O'Connor, 2020).

Whole language is based on the premise that learning to read comes naturally to children in the same way as learning to speak. Research on whole language does not have scientific support for its effectiveness. While whole language focuses on teaching meaning and encourages students to read more, it emphasizes cueing systems that teach students to use context clues to guess the meaning of words and de-emphasizes scientifically proven skills-based approaches such as phonics.

Balanced Literacy

Over time, the pendulum swung back and forth between approaches favoring phonics and approaches favoring whole language (NICHD, 2000) until a compromise appeared when

practitioners blended these concepts into a balanced approach, which combines skills-based and meaning-centered instruction with the goal of having students read books they enjoy as quickly as possible (Hollingsworth, 2023). Balanced literacy is a philosophical orientation blending instruction and support on phonics, whole language, and captivating literature with student choice and student engagement (Carroll, 1997; Fountas & Pinnell, 1996; Pressley, 2005).

Verhoeven and Snow (2001) believe that enthusiasm for reading is a prerequisite to learning how to read, so nurturing the joy of literacy is an integral part of literacy programs. Balanced literacy instruction incorporates shared reading (i.e., teachers read aloud and ask students questions about the text); guided reading (i.e., students read texts at their current ability level and discuss them with the teacher in small groups of other students at the same ability level); and independent reading (i.e., students choose books to read on their own; Lorimor-Easley & Reed, 2019). This approach has been embraced by many colleges of education (Hollingsworth, 2023). According to EdWeek Research Center (2020), 72% of elementary teachers say they embrace balanced literacy, and their schools are using it; however, the approach is nebulously defined and criticized for not including enough explicit, systematic phonics instruction. Reading comprehension is hampered when students possess a weak foundation of decoding and can lead to students learning how to guess words instead of how to sound them out (Gough & Tunmer, 1986; Hollingsworth, 2023).

While balanced literacy was popularized in the mid-1990s to describe a middle ground between whole language and phonics in terms of text selection, instruction, skills, and strategies, Fisher et al. (2023) point out that the term has become vaguer. Their descriptive study polled 25 veteran teachers to identify commonalities among their definitions and instructional applications. Interviews with each of the teachers were conducted by two researchers using an interview

protocol, and a third researcher participated in coding and analyzing transcripts using a constant comparative method with theme identification and writing results. The 25 teachers identified five different aspects they consider as they plan literacy instruction for their students: balancing reading and writing, phonics and comprehension, informational texts and narrative texts, direct instruction with dialogic approaches, as well as whole class and small groups. Of these, only one, balancing whole class and small groups, was identified by all teachers. A major limitation of this study was the self-reporting nature of the data. Fisher et al. were looking for teachers' understanding, which was represented by the data in the findings, but they did not validate the examples they provided with classroom visits and observations.

Intended to be an integrative approach taking the best of both phonics instruction and whole language, balanced literacy was embraced by teachers and colleges of education. As of 2020, it appeared that most teachers and schools were using this approach (EdWeek Research Center, 2020). From the mid-1990s until 2023, research has shown that teachers' understanding of the definition and major concepts has become vaguer, with some suggesting that balanced literacy is a recasting of whole language disguised with new terminology, which research showed was ineffective (Hattie, 2009; Thomas B. Fordham Foundation, 2023).

Science of Reading

In response to statistics showing merely one-third of fourth graders are reading proficiently (NCES, 2022), parents of children with dyslexia pushing for more systematic instruction (Hollingsworth, 2023), and a growing number of NAACP chapters describing literacy as a civil rights issue (Hollingsworth, 2023), the "reading wars" (Pearson, 2004) have resurfaced with balanced literacy criticized in favor of approaches aligned to the science of reading. Although there is not an agreed upon specific definition, in general, the science of reading

combines objective investigation and reliable evidence about how humans learn to read, what happens in the brain during reading, and how students should be taught to enable skillful reading (Cheatham & Allor, 2012; Goodwin & Jiménez, 2020; Gough & Tunmer, 1986; Hoover & Tunmer, 2020; Liberman et al., 1989; Metsala & David, 2016; Moats & Tolman, 2022; Molenaar & Roda, 2008; Scarborough, 2001). The International Literacy Association has defined the science of reading as "a corpus of objective investigation and accumulation of reliable evidence about how humans learn to read and how reading should be taught" (Goodwin & Jiménez, 2020). Research has shown that people can innately pick up spoken language, but they do not have an innate ability to read, so teachers should systematically teach students how to read (Liberman et al., 1989).

Shanahan (2020) explained that "science of reading" is a term that has been used for more than 200 years, especially to refer to the pronunciation and decoding of words. It was first mentioned pedagogically during the 1830s, but use of the term waxed and waned since its recent upsurge beginning in 2018. The well-documented contentiousness of the reading field continues with some embracing and some opposing the science of reading (Shanahan, 2020). Thomas (2022) advised that policymakers should be wary of overstatements and oversimplifications, attend to other influences on measurable student reading achievement (e.g., socioeconomics, teacher expertise and autonomy, and teaching and learning conditions), recognize that student-centered is supported by research but also difficult to implement, and shift policy away from one-size-fits-all mandates and move to support for student needs and teacher-informed reform.

Instruction aligned to the science of reading emphasizes a phonics-rich, skill-driven, intentionally designed, and systematic approach to foundational word recognition and language comprehension (Cheatham & Allor, 2012; Goodwin & Jiménez, 2020; Gough & Tunmer, 1986;

Hoover & Tunmer, 2020; Liberman et al., 1989; Metsala & David, 2016; Moats & Tolman, 2022; Molenaar & Roda, 2008; Scarborough, 2001). Research emphasizes the importance of phonological awareness (i.e., teaching students how words are broken into sounds) and why learning sounds in words and phonemes and matching sounds to graphemes are crucial for learning to read (Metsala & David, 2016; National Early Literacy Panel, 2008; NICHD, 2000). Additionally, explicit, and systematic phonics instruction that develops automaticity has proven to be the most beneficial for students (Molenaar & Roda, 2008; National Early Literacy Panel, 2008; NICHD, 2000). Cheatham and Allor (2012) also point out the benefits of having students apply new phonics skills by practicing reading words in connected texts.

Both the National Early Literacy Panel (2008) and the National Reading Panel (NICHD, 2000) conducted meta-analyses to establish what was effective in learning to read and found the concepts of the science of reading to be effective. National Early Literacy Panel (2008) found that six variables (i.e., alphabet knowledge, phonological awareness, rapid automatic naming of letters or digits, rapid automatic naming of objects or colors, writing or writing name, and phonological memory) representing early literacy skills or precursor literacy skills had medium to large predictive relationships with later measures of literacy development and maintained their predictive power even when the role of other variables, such as IQ or socioeconomic status were accounted for. An additional five early literacy skills (i.e., concepts about print, print knowledge, reading readiness, oral language, and visual processing) were also moderately correlated with at least one measure of later literacy achievement, but either they did not maintain predictive power when other variables were accounted for, or they had not yet been evaluated by researchers. The studies were categorized into five analytical categories: code-focused interventions, shared-reading interventions, parent and home programs, preschool and kindergarten programs, and

language-enhancement interventions. Code-focused interventions, which are designed to teach children skills related to cracking the alphabetic code (often including phonological awareness instruction), reported statistically significant and moderate to large effects across a broad spectrum of early literacy outcomes and demonstrated positive effects directly on children's conventional literacy skills.

Similarly, the National Reading Panel, whose reading experts evaluated the evidence for how to best teach reading, chose eight categories of teaching tasks and conducted a meta-analysis of 38 studies involving 66 controlled experiments from 1970 through 2000. Their results showed curriculum grounded in the science of reading with five components of reading instruction (i.e., phonemic awareness, phonics, fluency, vocabulary, and comprehension) to be more effective than other approaches to teaching reading (NICHD, 2000). Of these, the two that seemed most effective were phonemic awareness and phonics. Studies showed higher levels of phonemic awareness in kindergarten and first grade were predictors of stronger reading skills later and students receiving systematic phonics instruction, especially when started before first grade, scored higher on sounding out words, word reading, spelling, and comprehension (NICHD, 2000).

The major limitation of meta-analyses is the quality of the original studies that are being combined. Studies have varying degrees of weakness in implementation and reporting, different study-design features, and various demographic and environmental characteristics of participants. Not all studies on the topic could be included. Meta-analyses can provide clues about what influences the effectiveness of an intervention but cannot provide a definitive answer (National Early Literacy Panel, 2008).

In 2021, the General Assembly of North Carolina passed Session Law 2021-8, which mandated that literacy instruction must be aligned with the science of reading; this legislation defined science of reading to mean evidence-based reading instruction practices that address the acquisition of language, phonological and phonemic awareness, phonics, spelling, fluency, vocabulary, oral language, and comprehension that can be differentiated to meet the needs of individual students. Reutzel (2022) also notes that curriculum and instruction utilized for science of reading instruction should be research-tested, not merely research-based.

Structured Literacy. Structured literacy is a term that unifies multiple evidence-based programs; applies the science of reading to classroom instruction; often uses a scripted approach; and describes explicit and systematic teaching that focuses on phonological awareness, word recognition, phonics and decoding, spelling, and syntax at the sentence and paragraph levels (Lorimor-Easley & Reed, 2019). Two important frameworks for understanding and identifying Structured Literacy are Gough and Tunmer's Simple View of Reading (1986) and Scarborough's (2001) Reading Rope Model.

LETRS. Many states, including North Carolina through Session Law 2021-8, are transforming the way they teach reading by moving to approaches aligned with the science of reading, and they are mandating staff professional development using one specific program: Language Essentials for Teachers of Reading and Spelling, commonly known as LETRS (Moats & Toulman, 2022). Using a speech-to-print approach, LETRS explains what literary skills must be taught, why they must be taught, and how to plan to teach them to students. Aligned to Gough and Tunmer's (1986) Simple View of Reading, LETRS helps teachers learn how to teach and assess phonemic awareness (i.e., knowledge of the sounds used in the English language); phonics (i.e., how sounds represent letters that can create words); morphology (i.e., parts of words that

create meaning); fluency, spelling, vocabulary (and other spoken language abilities); and comprehension. LETRS also helps teachers create language-rich classrooms, build connections between reading and writing, diagnose reading problems, and differentiate reading instruction (Moats & Toulman, 2022).

Although there might not yet be one specific definition of the science of reading, the concept includes brain research and evidence about how people learn to read and how teachers should teach beginning reading. Gough and Tunmer's (1986) Simple View of Reading and Scarborough's (2001) Reading Rope provide a foundation for the science of reading. Research has shown that curriculum and staff development programs (such as LETRS [Moats & Toulman, 2022]) grounded in the science of reading with five components of reading instruction (i.e., phonemic awareness, phonics, fluency, vocabulary, and comprehension) are more effective than other approaches to teaching reading (NICHD, 2000).

Summary of Philosophies of Teaching Reading

While there are various philosophies of teaching reading, explicit and systematic phonics instruction, which helps students learn how letters are linked to sounds to form letter-sound correspondences and spelling patterns, should be an essential part of any beginning reading program (Ehri et al., 2001; NICHD, 2000); however, it should be integrated with other reading instruction to create a complete reading program (NICHD, 2000). Research showed that whole language, based on the premise that learning to read comes naturally to children in the same way as learning to speak so teachers should focus on meaning and encourage students to read more, was not an effective strategy in teaching students to read (Hattie, 2009). Intended to combine the best of phonics and whole language, balanced literacy was embraced by teachers and colleges of education, but research has not proven its effectiveness (Fisher et al., 2023), and some suggest

that balanced literacy is a repackaging of whole language disguised with new terminology (Thomas B. Fordham Foundation, 2023), which research showed was ineffective (Hattie, 2009). While there may not yet be one specific definition of the science of reading, the concept includes brain research and evidence about how people learn to read as well as how teachers should teach beginning reading. Research has shown that curriculum grounded in the science of reading with five components of reading instruction (i.e., phonemic awareness, phonics, fluency, vocabulary, and comprehension) is more effective than other approaches to teaching reading (NICHD, 2000).

Adolescent Literacy

Adolescent literacy is a complex concept that includes comprehending different kinds of texts, mastering new vocabulary, and sharing ideas with others (Kamil et al., 2008). Too many students demonstrate reading difficulties continuing into middle and high school. Data from the 2022 NAEP in reading indicate that only 31% of Grade 8 students are reading proficiently and able to comprehend text at their grade level (NCES, 2022). These scores declined 3 points (from 263 to 260), which show that 2022's results were the same as 1992's results. This decline is one of the largest observed in a single assessment cycle, and it likely was affected by the COVID-19 pandemic learning disruptions. What constitutes reading skills also increases to incorporate not only foundational reading skills but also content-area-specific skills (e.g., reading to learn in subjects that present ideas and content in different methods, building knowledge by comprehension of different kinds of texts, mastering new vocabulary, and sharing ideas with others). The literacy achievement gap continues to widen every year: as non-proficient readers continue through the grades reading infrequently, they miss out on increasingly more background knowledge, skills, and specialized vocabulary, which are important for understanding the schools' core curriculum (Alvermann, 2005). Without continuous literacy

instruction, non-proficient sixth graders will likely never catch up with their proficient peers (Kamil et al., 2008). It is evident that there is a need for improved literacy instruction for adolescent students; however, instruction in reading as a formal part of school curriculum usually decreases as students move out of elementary school (Kamil et al., 2008). Also, Black and Wiliam (1998) found evidence from many studies that learners' beliefs about their own capacity as learners can affect their achievement.

Adolescent Aliteracy

While it is imperative that educators ensure students have the basic skills necessary to read, and therefore are not *illiterate*, Verhoeven and Snow (2001) point out that *aliteracy*, which they define as the condition of being able to read but choosing not to do so, is also a challenge, especially as students transition to middle school. Adolescent rebellion is typical of the teenage years, and the ultimate rebellion in a reading or English language arts classroom is aliteracy, so teachers must focus on motivating and engaging students (Dredger, 2013). Aliteracy presents a specific set of challenges to teachers of students in middle school primarily around low self-efficacy and motivation for reading (Alvermann, 2005). Middle school students are often more oppositional than elementary school students (Strambler & Weinstein, 2010), and they may hold more diffused identities than high school students have (Kroger et al., 2010). Mikulecky (1978) argues that aliteracy occurs when functional or basic literacy skills and testing and exams are over-emphasized and voluntary reading and reading for pleasure are deemphasized.

The joy of reading is never discovered if there is no instructional time remaining for reading activities that lead to true literacy (Decker, 1986). George (2021) notes other influences contributing to a rise in aliteracy including a decline in the emphasis placed on reading and writing in K-12 schooling, a lack of reading role models, a lack of interesting books in school

libraries, and the fact that many students prefer video games and the internet over books. Decker (1986) advocates for strategies to prevent aliteracy including making reading relevant, connecting reading and writing, and breaking the cycle through effective teacher preparation. Guthrie et al. (2013) modeled the interrelationships of reading instruction, motivation, engagement, and achievement in two contexts using data from 1,159 seventh grade students and found that Concept-Oriented Reading Instruction, which emphasizes support for reading motivation, reading engagement, and cognitive strategies for reading informational text, was associated with positive changes in motivation, engagement, and achievement relative to traditional reading/language arts instruction. Limitations of the study include that few causal claims can be made about the effectiveness of Concept-Oriented Reading Instruction because it had a substantial advantage for increasing multiple outcomes, the study design's quasi-experimental nature, the restriction to informational text reading, and the population studied was a mixture of African American and European American students, so results may differ in different populations.

Strategies for Addressing Adolescent Literacy

Ivey and Fisher (2006) believe a whole-school approach, instead of isolated teaching strategies and textbooks, is necessary to improve the literacy skills of secondary school students. They describe four recommendations based in the whole language philosophy for all content area teachers to improve adolescents' literacy skills:

1. Expose students in all subjects to literacy-rich experiences.
2. Use curriculum materials that draw students into literacy experiences that they find relevant and meaningful.

3. Focus instruction on themes, big ideas, and essential questions that span the subject areas.
4. Provide time every day for students to read independently.

Kamil et al. (2008) created a report for the Institute of Education Science's National Center for Education Evaluation and Regional Assistance on effective classroom and intervention practices to improve adolescent literacy. Their five recommendations to support non-proficient adolescent readers are evidence-based, aligned to the science of reading philosophy, and derive from rigorous studies:

1. Provide explicit vocabulary instruction.
2. Provide direct and explicit comprehension strategy instruction.
3. Provide opportunities for extended discussion of text meaning and interpretation.
4. Increase student motivation and engagement in literacy learning.
5. Make available intensive and individualized interventions for struggling readers that can be provide by trained specialists.

Providing explicit vocabulary instruction, providing direct and explicit comprehension strategy instruction, and making available intensive and individualized interventions for struggling readers that can be provided by trained specialists had strong evidence to support their effectiveness. Heller and Greenleaf (2007) advise teachers to focus on the cognitive, social, and personal strengths students bring with them from home, instead of looking only at students' deficiencies in reading and writing.

Secondary Teachers' Self-Efficacy in Teaching Non-Proficient Adolescent Readers

Bandura (1997) describes self-efficacy as the "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p. 2), and teachers'

perceived self-efficacy determines their behaviors, which directly affect learning outcomes of students (Bandura, 1993; Lotter et al., 2018; Neugebauer et al., 2019). According to Bandura's (1986) Social Cognitive Theory, different circumstances affect self-efficacy in response to various situations, and desired outcomes and individual reactions are based on individuals' personal motivations, emotions, and environments. At the secondary level, the responsibility for teaching reading skills often seems to belong to no one specifically (Heller & Greenleaf, 2007). Few middle school teachers have had formal training in providing reading instruction, and many secondary teachers report feeling unprepared to help students improve their reading proficiency, or they do not think that teaching reading skills in their content-area courses is their responsibility (Heller & Greenleaf, 2007; Kamil et al., 2008). Even when teachers have a sense of efficacy about teaching in their specific content area, they may have a low sense of efficacy about teaching reading students they perceive as struggling or unmotivated (Fine et al., 2011).

In summary, adolescent literacy includes a wider range of issues than teaching beginning reading to younger students. Instructional standards require students to read to learn, and adolescents who are not proficient readers face more difficulties every year. Motivating adolescents to read becomes a challenge for teachers as adolescent aliteracy, which is being able to read but choosing not to do so (Verhoeven & Snow, 2001), increases. Research shows that supporting reading motivation, reading engagement, and cognitive strategies for reading informational text is beneficial for adolescent literacy (Guthrie et al., 2013), and secondary teachers need assistance in teaching reading skills to nonproficient adolescent students.

Coherent and Adaptive Assessment Systems

Educational assessment aims to determine how well students are learning and can be used to assist learning, to measure individual achievement, and to evaluate programs (Pellegrino et al.,

2001). Assessments can provide educators with information about students' strengths and growth opportunities, which can inform instructional decisions, and they can also serve accountability purposes by providing public stakeholders with information about the effectiveness of educational services. It is important for educators to develop productive relationships between both the formative (i.e., assessment for learning) and summative (i.e., assessment of learning) functions of classroom assessment, so their judgments inform and align with formal, external assessment of students, which increases the validity of all assessments (Black & Wiliam, 2018). Pellegrino et al. (2001) believed that students should learn something important from every assessment situation and that information should ultimately help teachers improve learning. Ideally, students would be able to become their own assessors and use assessments for new learning (i.e., assessment as learning).

Formative Assessments

Formative assessments are informal assessments that teachers use to assess student understanding and to differentiate instruction, and they should inform both teachers and students about progress being made on assessed skills (McGlynn & Kelly, 2017). One specific type of formative assessment is the diagnostic assessment, which is a pre-assessment used to collect data on what students already know prior to instruction. Wiliam and Thompson (2007) provided a framework for formative assessment identifying three processes: establishing where learners are in their learning, establishing where they are going, and establishing how to get there. Typically, teachers do not grade formative assessments (e.g., short quizzes, entrance and exit tickets, and classroom questioning) because they are used as practice on assessed skills. McGlynn and Kelly (2017) advise teachers to ensure their formative assessment methods are not only informative, but also engaging, to maximize student participation and performance. Pellegrino et al. (2001)

found that assessments should focus on making students' thinking visible to both their teachers and them (e.g., determining schemas, metacognitive skills, and problem-solving strategies that students possess) as well as provide timely and informative feedback to students in real-time. Once teachers have data from formative assessments, they can group students with similar levels of understanding together, which allows teachers to provide targeted instruction and scaffolding and effectively meet each student at his or her level of need without providing more or less work (McGlynn & Kelly, 2017).

Andrade and Heritage (2018) showed that formative assessment can enhance learning, achievement, and academic self-regulation. Jerald (2001), who investigated 4,577 high-poverty, high-minority, and high performing schools, found the effective use of daily classroom assessment is key to promoting high levels of student achievement in traditionally low-performing schools, but a limitation is that one should not use data from the study to make comparisons across states concerning the performance of schools. Black and Wiliam (1998) conducted a review of the literature on classroom formative assessment and noted that several studies showed firm evidence that strong, frequent feedback to students about their learning yielded substantial learning gains. Furthermore, Black and Wiliam highlighted that formative assessment is not well understood by teachers and is often weak in practice and student-involved formative assessment helps low achievers more than other students, which reduces the range of achievement and raises overall achievement. Achievement gaps can be reduced if classroom assessments exhibit the characteristics listed in Table 2.

Table 2

Research-Based Classroom Assessment Strategies to Reduce Achievement Gaps

Strategy	Researchers
Focus on clear purposes.	(Black & Wiliam, 1998, 2018; Stiggins & Chappuis, 2005)
Provide accurate reflections of achievement.	(Stiggins & Chappuis, 2005)
Provide students with continuous access to descriptive feedback on improvement in their work.	(Black & Wiliam, 1998; Bloom, 1984; Pellegrino et al., 2001; Rodriguez, 2004; Stiggins & Chappuis, 2005)
Bring students into the classroom assessment process.	(Black & Wiliam, 1998; Meisels et al., 2003; Stiggins & Chappuis, 2005)

Summative Assessments

Summative assessments occur at the conclusion of a defined instructional period (e.g., the end of unit of study, a specific course, a semester, a program, or a year of school for the purpose of evaluating whether students have learned the material they have been taught by comparing it against a standard or benchmark and certifying and reporting learning that has been achieved (Andrade & Heritage, 2018). Because of the time in which they are given, summative assessments often do not provide teachers with information that allows for planning instruction or informing interventions, and thus they are evaluative, rather than diagnostic. Often, they are high stakes, which means they are heavily weighted, and doing poorly carries sanctions. On assessments used for state and federal accountability, results are used to make important decisions about students, teachers, schools, and districts. These assessments, which include grading students, have gained a reputation for having unintended, possibly destructive, consequences for learning and motivation (Andrade & Heritage, 2018).

Coherence in Educational Assessment

Reducing variation in student learning and closing achievement gaps require clarity and consistency among all instructional components, which include learning goals or standards, followed by instruction with feedback and correctives, that through evaluation (hopefully) results in competent learners (Guskey, 2005). Shepard et al. (2017) argue that in designing assessment systems, state and local leaders should integrate coherent assessment, meaning multiple approaches are used to collect formative and summative information about student learning, with rich curriculum and effective instruction in equity-focused research on learning.

Cohesion is crucial because educational assessment does not exist in isolation, and only through alignment with curriculum and instruction can assessment support student learning. Furthermore, assessments should assess what really matters for student learning (e.g., core content and key skills), rather than superficial or simple facts (Pellegrino et al., 2001). When educators design coherent systems in which curriculum, instruction, and assessment effectively work together toward the same valuable, compelling, and reachable goals, students move closer to shared definitions of what they ought to know and be able to do (Shepard et al., 2017). Effective reading instruction should be predicated upon the systematic collection of valid, reliable, and meaningful assessment data using multiple assessments or test scores (Gambrell et al., 2011; Reutzel, 2022). There is no one assessment that can give teachers all the information they need to make a strong educational decision, and using multiple assessments helps teachers understand where and why a student is struggling.

Coherent assessments drive systematic instruction, which has three components: a defined scope of skills and strategies, skills that are taught in a planned sequence, and instruction that includes a periodic review and reteaching cycle (Reutzel, 2022). Reading skills and

strategies are taught through explicit instruction, which begins with explaining a specific learning objective, is followed by teacher modeling of cognitive processes in meeting the learning objective, and moves to students practicing and finally becoming independent users of the skill (Reutzel, 2022).

Adaptivity in Educational Assessment

Assessments should characterize student achievement in multiple aspects of proficiency; chart students' progress over time; include multiple paths of performance; and model performance at various levels (e.g., students, groups, classes, schools, and states; Pellegrino et al., 2001). Some kinds of assessments are more efficient and effective at meeting these goals than others. Two fundamental types of assessment design are fixed form and adaptive.

Fixed-Form Assessments. Many assessments, especially high-stakes tests, are fixed-form assessments, so all test takers receive the same questions usually in the same order. Fixed-form assessments efficiently measure proficiency on grade-level skills, include items based on a prior design, and have a narrow scope. Computer technology has provided some more options including randomized item order to discourage cheating, immediate scoring, and paper resource savings (Center on Standards and Assessment Implementation, 2019). Unfortunately, while fixed-form assessments tend to work well for students whose performance is on or near the level of the assessment, they do not show as easily exactly what material students on the further ends of the spectrum have or have not mastered. A fixed-form assessment cannot show how much a high achieving student knows or how far a struggling student is behind, which makes it more difficult to determine what instruction is needed to close achievement gaps among students (Davis, 2012).

Computer-Adaptive Assessments. While technology will not improve educational assessment on its own, it can assist with creating more robust assessment practices, including the use of adaptive assessments, which leverage advanced technology to provide an in-depth, personalized assessment of each student and track student growth continuously and consistently over the entirety of a student's school career (Curriculum Associates, n.d.). Computer adaptive testing begins with a large pool of questions and then selects individual questions for students, depending on their responses as they go along in the testing experience. If the student test taker answers questions correctly, the questions increase in difficulty. If the student test taker answers questions incorrectly, the questions decrease in difficulty (Center on Standards and Assessment Implementation, 2019). Students' scores on computer-adaptive assessments depend on the number of items they answered correctly and the difficulty of the items presented, rather than how many total correct or incorrect answers (Davis, 2012). Various components are needed to implement computer-adaptive assessments including a large pool of questions that are calibrated to a common measurement scale, a mechanism to select new questions based on students' earlier responses, a process to score students' responses, a process to terminate the test, and reports that share students' scores with their instructional needs (Center on Standards and Assessment Implementation, 2019).

Limitations of Computer-Adaptive Assessments. Because adaptive assessments require technology, a fundamental limitation to using them is a lack of technology and/or network infrastructure (i.e., a device for each student and sufficient internet connectivity). They also require extensive item banks as there must be enough test questions to cover all skills taught in school over a range of difficulty levels; these questions must be kept secure as well as relevant to student test takers. Another limitation of computer-adaptive assessments is that students cannot

skip questions to return to later or perform a post-test review prior to submission; instead, test questions must be answered as they appear (Center on Standards and Assessment Implementation, 2019).

Benefits of Computer-Adaptive Assessments. Computer-adaptive assessments can be used for a wide variety of purposes including formative assessments and summative assessments. They have many benefits including maximizing information on student performance (including performance described based on multiple grade levels), promoting accurate measurement of a growth across a student's school career, and helping administrators make long-term decisions and measure impact (Curriculum Associates, n.d.). Other benefits include item selection adapting to performance, improved efficiency, test security and item bank maintenance, and immediately delivered results (Center on Standards and Assessment Implementation, 2019).

Rasch Model. Rasch (1980) created a model used to measure underlying traits such as ability or attitude and indicates the probability of an individual getting an item correct on a test. The model is created from actual data of items (progressing in difficulty) and people's responses (progressing in number correct indicating the test takers' ability). The Rasch Model differs from Item Response Theory, which fits a model to data.

In summary, coherent and adaptive assessments involve multiple approaches to collect formative and summative information and create an in-depth, personalized assessment of each student that informs rich curriculum and effective instruction in equity-focused research on learning (Curriculum Associates, n.d.; Shepard et al., 2017). By tracking student growth continuously and consistently over the entirety of a student's school career, they provide meaningful data for teachers and administrators to use to identify what students know and can do prior to instruction, how much they are learning throughout instruction, and when and how to

differentiate instruction to meet students' specific needs. Coherent and adaptive assessments are key to aligning curriculum, instruction, assessments, and interventions.

Effective Instructional Interventions

When students are struggling with academic content, teachers need to intervene quickly with extra help. Students can receive additional help through remediation, yet this approach usually involves waiting for the students to fail before help is provided. Once provided, too often remediation is limited to workbook or worksheet activities emphasizing skills development (without engaging in the reading of connected text) facilitated by paraprofessionals or other non-certified staff. Rather than remediation, Neal and Kelly (2002) advocate for early or late intervention, which takes a proactive, catch-up stance toward students' specific learning needs to enable acceleration to become on-grade level with their peers. Interventions should be research-based and systematically delivered by a trained professional based on a targeted area of need to students in small groups or individually. They also describe six characteristics of successful late intervention programs that promote acceleration of learning performance:

1. Consider individual student needs.
2. Implement an apprenticeship model of teaching and learning.
3. Select appropriate materials.
4. Establish a focus on accelerative instruction.
5. Consider the role of fluent responding.
6. Provide for affirmation of success.

These characteristics illustrate an instructional framework for an intervention approach for adolescent non-proficient students.

Intensive Interventions

For some students, these interventions will still not result in successful learning, and more intensive interventions are needed. The National Center on Intensive Intervention (2013) defines Data-Based Individualization as a research-based process, characterized by increased intensity and individualization, for individualizing and intensifying interventions through the systematic use of assessment data, validated interventions, and research-based adaptation strategies. Data-Based Individualization is typically implemented within the context of a multi-tiered intervention framework, such as North Carolina's Multi-Tiered Systems of Support (MTSS), and it supports students who need intensive and individualized support for whom core and supplemental instruction are not working. The process begins when an intervention team decides a student needs a more intensive and individualized version of the intervention program he or she is receiving, and interventionists use progress monitoring data and diagnostic assessment to assess the student's response to the intervention and determine when adjustments are needed and to individualize instruction on a student-by-student basis. It is comprised of five steps: (a) secondary intervention program delivered with greater intensity, (b) progress monitoring, (c) diagnostic assessment, (d) adaptation, and (e) continued progress monitoring, with adaptations occurring whenever needed to ensure adequate progress.

Blended Learning Interventions

Blended learning interventions are intended to help non-proficient students master skills through a combination of teacher-led instruction and student completion of online activities using digital technology (Horn & Staker, 2011; Macaruso et al., 2020; Pytash & O'Byrne, 2018). At least some level of student control over where, when, and how to work is evident in blended learning. Literacy engagement is connected to students having active roles in their own

development (Verhoeven & Snow, 2001). Online activities can deliver personalized instructional content, which facilitates student control, but programs should be scaffolded sequentially to promote mastery and build confidence (Hirsh-Pasek et al., 2015). These activities provide real-time performance data to teachers to provide individualized instruction (Macaruso et al., 2020).

Mastery Learning

Bloom, who believed in using classroom assessments as learning tools and following with feedback and corrective procedures, created a specific instructional strategy known as "mastery learning" (Bloom, 1971). In this strategy, teachers organize concepts and skills into 1- or 2-week instructional units, provide initial instruction, and administer a brief formative assessment to give students feedback on their learning. Along with each formative assessment are individualized corrective activities for students to use in correcting any learning difficulties, so students only work on concepts they have not yet mastered. Once corrective activities are completed, students take a second, parallel formative assessment to determine if the correctives worked and to give students a second chance at mastery, which would increase their motivation. If students performed well on the first assessment and did not need correctives, then they would be provided with enrichment or extension activities (e.g., self-selected activities, special projects or reports, academic games, or other complex, problem-solving tasks) to broaden their learning experiences. Feedback, corrective, and enrichment procedures are crucial to mastery learning. Bloom (1984) found that mastery learning and one-on-one tutoring are more beneficial than conventional instruction, and it is important because the tutoring process demonstrates that most students have the potential to reach a high level of learning.

Motivation and Engagement

Motivation, or the desire, reason, or predisposition to become involved in a task or activity, and engagement, or the degree to which a student processes text deeply using active strategies, thought processes, and prior knowledge, play a key role in adolescents' academic literacy development (Alvermann, 2005; Guthrie & Wigfield, 2000; Kamil et al., 2008). Explanations vary as to the cause but decreases in intrinsic reading motivation have been noted as students, especially struggling students, move from elementary to middle school (Gottfried, 1985; Harter et al., 1992). Student motivation is highest when they receive feedback that is informational but not controlling (Kamil et al., 2008), and literacy, thinking, and motivation are not easily separated (Verhoeven & Snow, 2001). Guthrie and Wigfield (2000) conclude that instructional approaches including direct instruction, guided reading, and scaffolding do not directly affect student reading outcomes such as time spent reading independently, achievement on standardized tests, performance assessments, or beliefs about reading. Instead, the level of student engagement is the mediating factor through which classroom instruction influences reading outcomes, so teachers should focus on the degree to which students engage or disengage over time in reading-related tasks. Kamil et al. (2008) note that motivation and engagement are widely recognized as important moderators for learning, but they find there is limited scientific evidence linking these factors directly to student reading achievement.

Strategies to Increase Motivation and Engagement. Kamil et al. (2008) examined two experiments and one quasi-experimental study without major flaws to internal validity other than a lack of demonstrated baseline equivalence, three studies of weaker design, six experimental and quasi-experimental studies with low external validity, and two meta-analyses to support their

finding of moderate evidence regarding motivation and engagement. They provide four recommendations to increase student motivation and engagement in literacy learning:

1. Establish meaningful and engaging content learning goals around the essential ideas of a discipline as well as around the specific learning processes used to access those ideas.
2. Provide a positive learning environment that promotes student autonomy in learning.
3. Make literacy experiences more relevant to student interests, everyday life, or important current events.
4. Build classroom conditions to promote higher reading engagement and conceptual learning through such strategies as goal setting, self-directed learning, and collaborative learning.

Gamification. Action video games, which are highly engaging and intrinsically motivating to students, are a specific kind of intervention that has been found effective to improve reading rate and fluency via enhanced visual attention for students with dyslexia; however, the few studies that examined reading accuracy or comprehension outcomes following action video game training reported no improvement, which suggests more research is needed (Carioti et al., 2021; Łuniewska et al., 2018; Peters et al., 2019). Online interventions and video games can share other features. For example, Hirsh-Pasek et al. (2015) found that built into the application design itself, scaffolding toward a learning goal can take various forms, ranging from hint systems that provide supportive background knowledge, to curriculum-leveling strategies that provide more or less challenging options during a play session, to sophisticated adaptive

learning systems that model relevant behavior, understanding, and opportunities for each child in order to prescribe personalized sequences of learning experiences.

Effective instructional interventions close learning gaps for students. They take different forms and might be provided in the classroom or in a separate location, to either small groups or individual students, and by an in-person facilitator or technological platform. Research shows that interventions that increase motivation and engagement in adolescents are more effective (Alvermann, 2005; Guthrie & Wigfield, 2000; Kamil et al., 2008).

Fidelity of Implementation

Fidelity of implementation has been found to be a significant predictor of student reading outcomes (O'Donnell, 2008). Implementing a program or an intervention with fidelity means delivering it the way it was designed and in the way it was implemented during the research studies that validated its effectiveness (National Center on Intensive Intervention, 2013). Fidelity of implementation increases the likelihood of consistently obtaining the intended student outcomes (Harn et al., 2013, 2017). Effective interventions implemented with higher degrees of fidelity tend to be more effective, and schools should use a systematic process to regularly evaluate implementation (Durlak & DuPre, 2008; Fixsen et al., 2013; Quinn & Kim, 2017). While the concept of fidelity seems simple, it is complex and not yet well understood; however, it is necessary to conclude if the results can be explained by the interventions and not because of other variables (Harn et al., 2013, 2017). Important components of fidelity for interventions include group size, dosage/schedule or session frequency and duration, and all elements of content (National Center of Intensive Intervention, 2013), and there also should be integrity of progress monitoring as well. Capin et al. (2018) noted that less than half of the reading intervention studies they reviewed reported fidelity data to track student progress, and they

emphasized the need for clear and comprehensive standards for reporting to determine internal and external validity of reading intervention studies.

Implementation measures capture fidelity to structure and fidelity to process. Structural dimensions involve an objective look at whether important components of the intervention were delivered (or not), and they are usually measured through direct observation, but some can be measured through participants' self-reports. Process dimensions involve an examination of the qualitative aspects of implementing an intervention, and they are usually measured by observers who attempt to rate how well or to what degree important components were delivered. Reliably and consistently capturing data on process dimensions across multiple raters often poses challenges (Harn et al., 2017). O'Donnell (2008) reviewed public health literature to determine five criteria for measuring fidelity of implementation that also work in K-12 curriculum intervention research:

1. Adherence: Are the components of the intervention being delivered as designed?
2. Duration: What is the number, length, or frequency of sessions implemented?
3. Quality of Delivery: Is the way the implementer delivers the program using the prescribed techniques, processes, or methods?
4. Participant Responsiveness: To what extent are participants engaged by and involved in the activities and content of the program?
5. Program Differentiation: Are the critical features that distinguish the program from the comparison condition present or absent during implementation?

Adherence and duration relate to fidelity to structure (O'Donnell, 2008). Quality of delivery and program differentiation relate to fidelity to process. Participant responsiveness relates to both

fidelity to structure and fidelity to process (O'Donnell, 2008). Dumas et al. (2001) noted that failing to establish fidelity can severely limit conclusions drawn from outcome evaluations.

Harn et al. (2017) examined the different approaches commonly used in schools to assess fidelity of implementation and presented them in relation to student outcomes within a 7-month intervention study with at-risk kindergarten students using descriptive, correlational, and qualitative approaches. Data for the study were taken from a randomized controlled trial of three research-based interventions in seven elementary schools across two districts in the Pacific Northwest. They found arguable the assumption that if an intervention is implemented with fidelity that it is the cause of the obtained outcomes, and they also found significant correlations between process measures and outcomes. It was interesting to note that Durlak and DuPre (2008) found *implementation thresholds*, in which increasing fidelity does not lead to better outcomes in programs with well-articulated implementation features such as teaching specific content through systematic and explicit instructional delivery, possibly because interventionists focused more on their delivery than on responsively engaging with their students. Harn et al. (2017) wondered if there may be a point of *diminishing returns* for increasing fidelity, and they also point out that it is important to match interventions to contextual features because each school is unique. Matching the local context is key to successful implementation and sustainability in future years (Harn et al., 2017; K. Swain et al., 2010). While their findings were not conclusive on frequency or methods of measuring fidelity of implementation, they demonstrated the complexity of collecting and examining fidelity data in schools and noted the goal is not obtaining perfect fidelity but rather to examine fidelity in relation to student outcomes as improving student outcomes is the most important goal. Limitations in the study included sample size, retrospective analysis, restricted range and nonnormal distribution of the data, and confounding observers and

schools because one observer/coach was responsible for collecting all observations across the study and consulting with schools to improve implementation.

In summary, fidelity of implementation for structure and process is a key component of intervention programs to improve reading proficiency. In implementing programs as they were designed, schools should analyze and report on adherence, duration, quality of delivery, participant responsiveness, and program differentiation (O'Donnell, 2008). Throughout the process, it is important to remember that the overall goal of any program is improving student outcomes.

i-Ready

One program that purports to combine effective reading instruction, coherent and adaptive assessment systems, and effective instructional interventions to improve adolescent literacy if implemented with fidelity is Curriculum Associates' i-Ready supplemental reading program. It uses data from the i-Ready Diagnostic to determine each student's strengths and instructional needs in specific reading domains and creates a personalized pathway of explicit and systematic lessons in i-Ready Instruction as next steps for skill building (Curriculum Associates, 2021a). It utilizes computer-adaptive assessments and aims to motivate and engage students.

i-Ready Instruction

The theoretical framework behind i-Ready Instruction in reading aligns with the science of reading (Cheatham & Allor, 2012; Liberman et al., 1989; Metsala & David, 2016; Molenaar & Roda, 2008), the Simple View of Reading showing reading comprehension requires strong decoding and language comprehension skills (Gough & Tunmer, 1986; Hoover & Tunmer, 2020), and Scarborough's (2001) Reading Rope blending word recognition (i.e., phonological

awareness, decoding, and sight recognition) and language comprehension (i.e., background knowledge, vocabulary, language structures, verbal reasoning, and literacy knowledge) into skilled reading (i.e., fluent execution and coordination of word recognition and text comprehension; Curriculum Associates, 2021a). The approach i-Ready Instruction uses to teach reading is structured literacy, which unifies multiple evidence-based programs; applies the science of reading to classroom instruction; and describes explicit and systematic teaching that focuses on phonological awareness, word recognition, phonics and decoding, spelling, and syntax at the sentence and paragraph levels (Lorimor-Easley & Reed, 2019).

Reading lessons in i-Ready Instruction are carefully sequenced and designed to improve skills of striving, on-grade level, and advanced learners. Word recognition lessons include phonological and phonemic awareness (for Grades K-1 and some in Grade 2), phonics (for Grades K-3 and some through Grade 8), and high-frequency words (for Grades K-2 and some through Grade 8). Language comprehension lessons include vocabulary (for Grades K-8), reading comprehension (for Grades K-8), and syntax (for Grades 6-8; Curriculum Associates, 2021a).

Additionally, i-Ready Instruction is designed to support all students (below, on, or above grade level) and is intended to be used for an average of 45 minutes or more per week for at least 18 weeks each school year. Curriculum Associates (2023a) created personalized lessons attempting to motivate students through culturally relevant topics, corrective feedback promoting productive struggle, active student participation, real-world connections, and student-owned progress monitoring.

i-Ready Diagnostic

The theoretical foundation of i-Ready Diagnostic is based on effective assessments, computer adaptive testing, and item response theory, specifically using the Rasch Model (Curriculum Associates, n.d.; Rasch, 1980). While many assessments are fixed form (i.e., assess proficiency on grade-level skills based on a prior design), the i-Ready Diagnostic leverages technology to assess proficiency on both on-grade and off-grade-level skills through students answering items presented to them based on ongoing calculations of their ability (i.e., when students answer correctly, they get a harder question and vice versa). Thus, i-Ready Diagnostic should facilitate more meaningful classroom assessments by providing a thorough, personalized evaluation of every student and track growth (as well as create both typical growth and stretch growth goals) across years on important reading skills.

The i-Ready Diagnostic in reading provides both a criterion-referenced score and a normative score. Criterion-referenced scores are students' grade-level placements for reading, and they are driven by students' scaled scores, which can be tracked and compared across grades and provide information about how students are performing relative to state standards. Normative scores, which are percentiles based on a nationally representative sample of students in the same grade level who took the i-Ready Diagnostic the same time of year and provide information about how students are performing relative to other students. Scores are provided for each of the five specific domains of academic reading ability three times per year with reports on each student's strengths and instructional needs, recommendations for next steps for skill-building in teacher-led instruction, and creation of a personalized pathway in i-Ready Instruction (Curriculum Associates, 2021a). These scores help teachers gain a more complete picture of student performance; assist in giving both students and parents feedback on students' current

performance level, strengths, and weaknesses; and provide targeted areas for interventions and personalized pathways for instruction. Word recognition strands are measured in phonological awareness (for Grades K-1 and some in Grade 2), phonics (for Grades K-3 and some through Grade 12), and high frequency words (for Grades K-2 and some through Grade 8). Language comprehension strands are measured in vocabulary (for Grades K-12) and reading comprehension in literature and informational text (for Grades K-12; Curriculum Associates, 2021a).

Philosophies of Teaching Reading in i-Ready

The theoretical models that provide the foundation for i-Ready Diagnostic and i-Ready Instruction are based on the science of reading (Curriculum Associates, 2021a) and align with the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Tunmer, 2020), and Scarborough's (2001) Reading Rope. Additionally, i-Ready emphasizes components of a science of reading curriculum including word recognition, decoding, phonological awareness, sight word recognition, language comprehension, vocabulary, background knowledge, language structures, literacy knowledge, and verbal reasoning (Curriculum Associates, 2021a). Curriculum Associates (2021a) stated they believe all students can become skilled readers through explicit, systematic, and evidence-based literacy instruction grounded in the science of reading, and they continue to improve their instruction to motivate all students, but especially adolescents in middle school, through age-appropriate visual graphics, more engagement with learning concepts, and responsive instruction to meet each student's personalized needs. Research shows that supporting reading motivation, reading engagement, and cognitive strategies for reading informational text is beneficial for adolescent literacy (Guthrie et al., 2013), and i-Ready allows educators to share students' data with them, so they can take ownership of their reading progress

and set their own meaningful learning goals, implement personalized learning paths based on their own unique strengths and growth opportunities, and engage in online lessons specifically designed to motivate adolescents. With Teacher Toolbox, the digital resource collection teachers can use for interventions aligned to students' i-Ready Diagnostic results (Curriculum Associates, 2023a), secondary teachers also receive explicit and systematic lessons to assist them in teaching reading skills to nonproficient adolescent students.

Coherent and Adaptive Assessment Systems in i-Ready

The i-Ready Diagnostic is a computer-adaptive assessment that uses a Rasch Model to adjust until it determines the level at which students need to receive instruction and pinpoints which skills students have mastered and which need improvement. When students answer questions correctly, i-Ready Diagnostic gives them more challenging questions, and when students answer questions incorrectly, it gives them fewer challenging questions. It is designed for students to get about 50% of the questions correct and 50% incorrect to determine students' placement level as quickly and efficiently as possible. Curriculum Associates (n.d.) had extensive field testing of items with over 2 million students to determine the difficulty level of each individual test item and its indicator group as well as to make probabilistic inferences about what students know and are likely able to do. Information gleaned from i-Ready Diagnostic coherently informs i-Ready instruction, which personalizes lessons based on individual student's strengths and weaknesses. This coherence and cohesion also link assessment results to instructional guidance in Teacher Toolbox, which assists educators in student placement decisions.

Interventions in i-Ready

The i-Ready suite is a blended learning intervention. Based on the i-Ready Diagnostic adaptive assessment and progress monitoring, students benefit from online personalized learning and intervention in i-Ready Instruction in addition to classroom instruction and interventions led by teachers using the Teacher Toolbox (Curriculum Associates, 2021b). Since all students participate in the i-Ready Diagnostic and i-Ready Instruction, early intervention is a reality. Also, i-Ready instruction uses the mastery learning cycle of feedback and corrective actions and action video game principles to motivate and engage students.

Fidelity of Implementation of i-Ready

In implementing the i-Ready program, fidelity of structure and fidelity of process are key components (Harn et al., 2017; O'Donnell, 2008). Adherence to i-Ready Diagnostic and i-Ready Instruction ensures that the program components are delivered as designed, so students should participate in both the i-Ready Diagnostic and i-Ready Instruction. Duration refers to the number of i-Ready Diagnostics students take each year, which should be three; length of time students work on i-Ready Instruction each week, which should be 30-49 minutes; and percentage of lessons passed, which should be at least 70%. Both adherence and duration can be measured through quantitative reports generated by the i-Ready program. Quality of delivery refers to the way that the school delivers the prescribed techniques, processes, and methods; educators should attend professional development, meaningfully use achievement and usage reports generated from i-Ready Diagnostic and i-Ready Instruction, as well as incorporate lessons from Teacher Toolbox. Participant responsiveness is the extent to which participants are engaged by and involved in i-Ready's activities and content, and ways to measure this engagement including looking at students' time on task during an i-Ready Instruction session or seeing if students

rushed or appeared to rush on their i-Ready Diagnostics or i-Ready Instruction lesson assessments. Lastly, program differentiation includes looking to see if the critical features that distinguish the program from the comparison condition are present or absent during implementation, so it is necessary to see if teachers or grade levels are using i-Ready Diagnostic and i-Ready Instruction with full or partial fidelity. Throughout the process, it is important to remember that the overall goal of any program is not perfect fidelity to the program, but rather improving student outcomes (Harn et al., 2017).

Effectiveness of i-Ready

Research Conducted by Curriculum Associates. Curriculum Associates has asserted that i-Ready Instruction has a positive effect on students' reading achievement. Their research team evaluated 2017-18 i-Ready Diagnostic data from more than one million K-8 students. Calculating and comparing descriptive statistics, including mean score gains, for the treatment group (students using i-Ready Instruction) and control (students who did not use i-Ready Instruction), they found that students using i-Ready Instruction as intended (i.e., for an average of 45 minutes per week for at least 18 weeks) showed greater learning gains on i-Ready Diagnostic (i.e., 46% more in reading) than those who did not use i-Ready Instruction as recommended, when controlling for prior achievement. This process was also used for each special population. Percent gains were calculated by taking the average percent gain within each grade level and population. They also found that, on average, K-8 students in special populations, including students with disabilities, English Learners, students who are economically disadvantaged, and students of color, using i-Ready Instruction as intended show greater learning than those from the same demographic groups not using i-Ready. Their rigorous study design (an analysis of covariance [ANCOVA] of more than 440,000 students using i-

Ready Instruction in reading) that controlled for prior spring test scores among students in Grades 1-8 and significant findings provide support for i-Ready Instruction as a program that meets criteria for ESSA Level 3 evidence standards. Limitations include overall sample sizes for middle school students (which are approximately half of elementary sizes), the much smaller treatment than control group, and inferences about the impact of fidelity on student learning gains are correlational and not causal (Curriculum Associates, 2020).

Another study conducted by Curriculum Associates looked at fall to spring score gains on the i-Ready Diagnostic for groups of students with different levels of exposure to i-Ready Instruction and different pass rates on Lesson Quizzes. The overall sample included 1,326,607 students who took the i-Ready Diagnostic in reading during 2017-18. Students using i-Ready Instruction as recommended (i.e., for an average of 30-49 minutes per week who passed at least 70% of their Lesson Quizzes) made greater gains than students who did not use i-Ready as recommended (i.e., for 10-29 minutes on average per week and passed fewer than 70% of their Lesson Quizzes). Interestingly, students who used i-Ready Instruction for more than 50 minutes per week and passed 70% of their lessons experienced greater growth than students who met the recommended guidance, but Curriculum Associates notes that this may be appropriate in some but not all implementations. This study also included an ANCOVA, which accounted for the prior fall scores and was a more rigorous test of change compared to simple raw gain scores, provides evidence that students in Grades 1-8 who used i-Ready Instruction in reading as recommended made positive, statistically significant gains and greater improvement than those not using i-Ready Instruction as recommended.

A major limitation is that this study is observational in nature, and inferences about the impact of fidelity of implementation on student learning gains are correlational and not causal

(Curriculum Associates, 2021c). Curriculum Associates conducted a follow-up to this study using the same procedures with data collected from 741,859 students during the 2018-19 school year. Based on students' fall-to-spring score gains on the i-Ready Diagnostic, findings show that K-8 students who use i-Ready Instruction as recommended make greater improvements in reading (including achieving a higher grade-level placement by spring, exceeding their typical growth targets, and achieving a higher percentage of their stretch growth targets) than students who do not. Again, a major limitation is that this study is observational in nature, and inferences about the impact of fidelity of implementation on student learning gains are correlational and not causal (Curriculum Associates, 2022b).

Curriculum Associates also conducted a large-scale study on the relationship between i-Ready Diagnostic and the 2021 North Carolina end-of-year summative assessment (NC EOG). Data were collected from approximately 28,000 students from 11 North Carolina school districts during the 2020-21 school year. Correlations between the Spring i-Ready Diagnostic and the 2021 NC EOG for reading were .76 (Grade 6), .78 (Grade 7), and .76 (Grade 8; Curriculum Associates, 2022a). One note of caution on this research that it is funded and published by Curriculum Associates, the publisher of i-Ready. All of these findings should be read with caution due to the lack of independent verification and validity of the findings.

Independent Research on i-Ready. While more independent research is needed to determine the effectiveness of i-Ready on nonproficient adolescent readers, two available research studies found i-Ready Instruction to be an effective online program in reading. Also, the National Center on Intensive Intervention (2021) gave i-Ready high ratings in academic screening on the i-Ready Diagnostic, academic progress monitoring on the i-Ready Diagnostic and Growth Monitoring, and academic intervention on the i-Ready Instruction.

The Center for Research and Reform in Education at Johns Hopkins University conducted a quasi-experimental design study on i-Ready Instruction's effectiveness on reading achievement of students from five Massachusetts school districts in the 2020-21 school year. The Center found that students who used i-Ready with fidelity outperformed students who did not use i-Ready on the Massachusetts end-of-year summative assessment (Massachusetts Comprehensive Assessment System for English Language Arts). This independent study provided ESSA Level 2 evidence of impact. Limitations of this evaluation include an inability to control for or conduct relevant subgroup analyses on some influential variables (i.e., economic disadvantage and English Learner status), as they were not available from all school districts involved in the evaluation. Also, researchers only had access to spring test scores and i-Ready Diagnostic score data from 2020-21, which limited their analyses to only one year and to strictly quantitative measures. Lastly, researchers were unable to draw any conclusions about fidelity of implementation within classrooms (Cook & Ross, 2022).

The Human Resources Research Organization (HumRRO) and Century Analytics partnered to examine the impact of i-Ready Instruction using a quasi-experimental design with propensity score matching using data from the 2018-19 school year. Their independent impact analysis using hierarchical linear modeling concluded that middle school students who used i-Ready Instruction with fidelity had higher reading achievement that students who did not use it and those students using i-Ready Instruction had a statistically significant higher reading i-Ready Diagnostic score than students in a matched comparison group. The researchers also found that these positive results were generally seen in specific subgroups of students including black students, Latinx students, English learners, students with disabilities, and students with economic disadvantages. This independent study provided ESSA Level 2 evidence of impact, which means

that there is at least one finding that shows moderate evidence of effectiveness (i.e., the study meets What Works Clearinghouse Standards with or without reservation under version 2.1 or later, has statistically significant positive effect, has at least 350 students, and involves at least two educational sites; U.S. Department of Education, n.d.).

The results of this study must be considered with caution as the intra-class correlations were above 0.20 for all grades, which suggests school differences may be important for matching and estimating treatment effects. Data shows large variations in how many students at a given school or grade use i-Ready Instruction with fidelity. Also, some schools only used i-Ready Diagnostic and i-Ready Instruction with low performing students, so baseline equivalence could not be achieved for students within a school. Also, the study was a quasi-experimental design and there was a lack of information of implementation decisions made at each school and within each classroom. Lastly, the treatment group was compared to the matched comparison group using the i-Ready Diagnostic, and it is possible that use of the i-Ready Diagnostic itself increases student achievement (M. Swain et al., 2020).

Summary

For over a century, as the name "reading wars" (Castles et al., 2018; Chall, 1967; Kim, 2008; Yaden et al., 2021) indicates, people have debated about the best way to teach reading (as if there is only one best way) from the various philosophies of teaching reading, which can be measured on a continuum of instructional approaches from skills-based to meaning-based. Phonics and the science of reading emphasize skills-based approaches, and whole language and balanced literacy emphasize meaning-based approaches. Research shows that beginning readers need explicit and systematic phonics instruction to help them understand how letters are linked to sounds to form letter-sound correspondences and spelling pattens as well as how to apply their

knowledge (Chall, 1967; Ehri et al., 2001; Flesch, 1955; NICHD, 2000), and phonics is one part (along with phonemic awareness, fluency, and comprehension strategies) of a complete reading program aligned to the science of reading (NICHD, 2000). While research does not support the effectiveness of the whole language and balanced literacy in teaching beginning reading because of the reliance on compensatory strategies, weak foundations in decoding, and vagueness in teacher understanding (Fisher et al., 2023; Gough & Tunmer, 1986; Hattie, 2009; Hollingsworth, 2023; Lorimor-Easley & Reed, 2019), the emphasis on motivating and engaging readers is important for adolescent literacy and trying to eliminate adolescent aliteracy (Alvermann, 2005; Decker, 1986; George, 2021; Guthrie et al., 2013; Verhoeven & Snow, 2001).

It is important for educators to know what students are expected to know, where each student currently is on that path, and how to get each one from where s/he is right now to where s/he needs to be. That responsibility is a great one, as educators work with many students and personalizing instruction is difficult without solid resources. Consequently, creating and using a coherent and adaptive assessment system, which has multiple approaches to collect formative and summative information and creates an in-depth, personalized assessment of each student that informs rich curriculum and effective instruction in equity-focused research on learning (Curriculum Associates, n.d.; Shepard et al., 2017), is important. Tracking student growth continuously and consistently over the entirety of a student's school career provides meaningful data for teachers and administrators to use to identify what students know and can do prior to instruction, how much they are learning throughout instruction, and when and how to differentiate instruction to meet students' specific needs. Aligning curriculum, instruction, assessments, and interventions is possible because of coherent and adaptive assessments. Closing learning gaps for students requires effective instruction interventions, which take different forms

and might be provided in the classroom or in a separate location, to either small groups or individual students, and by an in-person facilitator or technological platform.

Research shows that interventions that increase motivation and engagement in adolescents are more effective (Alvermann, 2005; Guthrie & Wigfield, 2000; Kamil et al., 2008). Fidelity of implementation is a key component of intervention programs to improve reading proficiency. In implementing programs as they were designed, schools should analyze and report on adherence, duration, quality of delivery, participant responsiveness, and program differentiation (O'Donnell, 2008) but remember that the overall goal of any program is improving student outcomes.

Ensuring that all students are proficient readers to help them succeed in school, work, and citizenship is one of the most important responsibilities of the educational system. Teachers must know what happens in the brain while people learn to read, how to implement instructional strategies that explicitly and systematically build literacy skills, and when and how to intervene with students are not demonstrating reading proficiency. Effective instructional approaches aligned with the science of reading with a cohesive and adaptive assessment system that drives effective instructional interventions to assist teachers in personalizing instruction to meet individual students' needs are necessary for coherence in students' educational experience. Using i-Ready with fidelity could provide alignment of reading instruction, assessment, and interventions in developing proficient readers. Implementing i-Ready as a systematic reading curriculum, instruction, and assessment program could create more equitable opportunities for students in learning foundational reading skills and closing learning gaps exacerbated by the COVID-19 pandemic. A summary of the review of related literature appears in Appendix B. In

the chapters that follow, the research methods, tentative findings, conclusion, and recommendations will be discussed.

CHAPTER 3

METHODS

The purpose of this convergent mixed methods design (Creswell & Creswell, 2018) program evaluation was to examine the fidelity of implementation (quantitative and qualitative design), attainment of intermediate-term outcomes (changes in reading achievement through quantitative design), and perceptions of users on the i-Ready program for middle school students in a small rural school district (quantitative and qualitative design). Its findings provided district leaders with information and recommendations on continuing, modifying, or discontinuing the program. The following evaluation questions guided this study:

Evaluation Questions

1. To what extent did teachers in the selected school implement with fidelity the i-Ready program (i.e., i-Ready Diagnostic and i-Ready Instruction) in the 2022-23 academic year?
2. To what extent, if any, did student reading proficiency improve during the 2022-23 academic year as measured by the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOGs?
3. What were the perceptions of teachers, instructional staff, and administrators regarding i-Ready in terms of guiding program elements such as the following: (a) the value of student data generated by the i-Ready Diagnostic, (b) the value of the i-Ready Instruction as an instructional resource, and (c) the extent to which i-Ready is having an impact on student achievement?

Program Evaluation Approach or Model

The guiding paradigm for this evaluation was pragmatic, which was responsive to the needs of stakeholders but mainly focused on usefulness in improving the program implementation in the stakeholders' specific context (Mertens & Wilson, 2019). The evaluation results needed to be meaningful and useful to the stakeholders who commissioned the evaluation. Consequently, the evaluation was designed to investigate the fidelity of i-Ready implementation, to determine if i-Ready was meeting the district's goal of increasing student reading proficiency, and to establish if the program seemed beneficial to specific stakeholders using it.

This program evaluation followed Stufflebeam's (2003) Context, Input, Process, and Product (CIPP) Evaluation Model. It was situated as a process and product evaluation, which assisted staff in revising activities, judging performance, and interpreting outcomes (Mertens & Wilson, 2019), and used mixed methods with quantitative review of Year 2 student usage and achievement outcomes (products) and qualitative analysis for stakeholder perception data (processes).

I sought to determine fidelity of implementation, effectiveness in producing desired outcomes of increasing student achievement, and perceptions of stakeholders. The i-Ready program would be valuable and possibly worth its cost if it provided staff and students with timely, understandable, and helpful instruction and data that promoted students' growth in reading, recovered unfinished learning, and improved their overall achievement. Results were intended to provide support for recommendations regarding the program's continuation, modifications, or termination.

Description of the Program Evaluation

This program evaluation relied on quantitative and qualitative data from the 2022-23 school year to examine the fidelity of activities by participants, student outcomes after participation in the program, and perceived benefit of the program for stakeholders using it as an instructional resource. During Summer 2023, i-Ready usage data (i.e., Diagnostic Status Report, Instructional Usage Report, and Personalized Instruction Report) from 2022-23 were reviewed and disaggregated by grade level and classroom teacher with descriptive statistics to determine the fidelity of implementation as prescribed by Curriculum Associates. Reading scores from the i-Ready Diagnostic (i.e., Diagnostic Results Report and Diagnostic Growth Report), NC Check-Ins/NC Check-Ins 2.0, which are benchmarks created and provided by the North Carolina Department of Public Instruction (NCDPI), and North Carolina End of Grade (NC EOG) assessments were reviewed and disaggregated by grade level and classroom teacher with descriptive statistics. Changes in the i-Ready Diagnostic and NC Check-Ins/NC Check-Ins 2.0 from the beginning, middle, and end of the year were analyzed using the Repeated Measures ANOVA. All end-of-the-year data points (i.e., i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOGs) were analyzed to determine how these three outcome measures correlated with each other and to determine if these outcome measures provided similar indicators of proficiency. Intended outcomes that were the focus of this evaluation were measured by data from formative and summative assessments as identified in the logic model (see Figure 1) and described in sections below.

Role of the Evaluator

As a senior administrator, I have positional authority, evaluate some participants, and set the direction for continuous improvement. For this study, I was an internal evaluator and

maintained an objective stance within the framework of the study. This duality required that I made it safe for stakeholders to express their perceptions transparently. I wanted honest responses in surveys and interviews, so staff needed to believe that their responses would not be used against them. My role may have posed an unknown degree of influence on the study's results. Using a reflexive journal helped clarify my thoughts as the researcher instead of a senior administrator and helped to maintain objectivity in data collection and analysis especially while coding interview transcripts. Additionally, I employed member checking by asking participants to review my transcriptions of their interviews.

Participants

Data from two primary stakeholder groups were used in this program evaluation. For quantitative analysis, there were 641 different students enrolled in grades six, seven, and eight who participated in at least one of the following reading assessments: beginning-of-year (BOY) i-Ready Diagnostic, middle-of-year (MOY) i-Ready Diagnostic, end-of-year (EOY) i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0 #1, NC Check-Ins/NC Check-Ins 2.0 #2, NC Check-Ins/NC Check-Ins 2.0 #3, or NC End of Grade (EOG) test. On the last day of the 2022-23 school year, there were 621 students enrolled in Cole Middle School in Grade 6 (n = 214), Grade 7 (n = 188), and Grade 8 (n = 219). All testing data from individual assessments taken by the 641 middle school students' reading data were included to address Evaluation Questions 1 and 2. Students had experience with i-Ready from the previous school year. All students, including students with disabilities and English learners, participated in i-Ready. Demographic data for all students appear in Table 3.

Table 3

Demographics of Cole Middle School Students

Grade	Male	Female	Asian	Black	Pacific Islander	White	Hispanic/ Latinx	2+	EL	SWD	SES	Total
6	110	104	6	70	0	107	19	12	19	32	152	214
	51%	49%	2%	33%	0%	50%	9%	6%	9%	15%	71%	100%
7	82	106	3	68	0	82	22	13	18	20	133	188
	44%	56%	2%	36%	0%	44%	11%	7%	10%	11%	71%	100%
8	124	95	6	70	0	113	22	8	16	34	140	219
	57%	43%	2%	32%	0%	52%	10%	4%	7%	16%	64%	100%

Note. n = 621. EL = English Learners; SWD = Students with Disabilities; SES = Socio-Economic Status and contains the number of students receiving free or reduced-price lunch.

For qualitative analysis, all middle school English language arts (ELA) teachers (n = 8), instructional staff (i.e., instructional coach), and administrators (i.e., assistant principal, principal, acting principal, and Executive Director of Instructional Services) in Cole Middle School addressed evaluation question three and provided a qualitative context for the quantitative data in the other two evaluation questions. In 2022-23, there were three Grade 6, three Grade 7, and two Grade 8 ELA teachers. The principal led during semester one, and the acting principal led during semester two while the principal was on leave.

The principal, acting principal, and assistant principal analyzed i-Ready Diagnostic achievement and i-Ready Instruction usage data in data dives with the executive director, who guided curriculum, instruction, and assessment for the district. The instructional coach used i-Ready Diagnostic and i-Ready Instruction data to guide professional learning communities, inform professional development, and coach teachers. Collectively, they looked for trends in

instruction and student needs to determine if i-Ready was improving student achievement in reading.

All teachers and the instructional coach participated in training provided by Curriculum Associates' consultants on using i-Ready designed for teachers facilitating i-Ready. The principal, acting principal, assistant principal, and executive director participated in training provided by Curriculum Associates' consultants on using i-Ready as an administrator to guide school-and district-level decision making. Demographic data for staff appear in Table 4.

Table 4

Demographics of Cole Middle School Staff

Position	Gender	Ethnicity	Years Experience	Highest Degree
Grade 6 Teacher	Female	White	19	Bachelor's
Grade 6 Teacher	Female	White	23	Bachelor's
Grade 6 Teacher	Female	White	21	Bachelor's
Grade 7 Teacher	Female	White	12	Bachelor's
Grade 7 Teacher	Female	White	27	Bachelor's
Grade 7 Teacher	Female	White	18	Bachelor's
Grade 8 Teacher	Female	White	27	Bachelor's
Grade 8 Teacher	Male	White	18	Bachelor's
Instructional Coach	Female	White	26	Master's
Assistant Principal	Female	White	26	Master's
Principal	Female	Black	27	Master's
Acting Principal	Male	White	28	Master's
Executive Director of Instructional Services	Female	White	13	Master's

Note. $n = 13$.

Data Sources

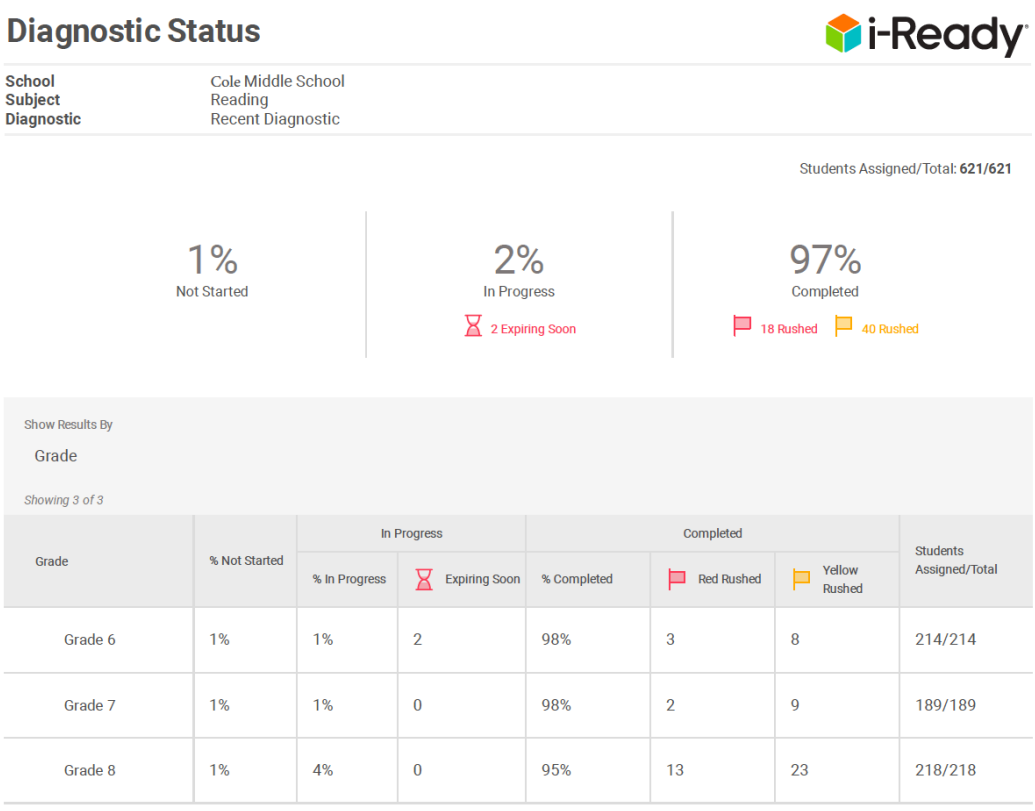
Both quantitative and qualitative data were collected for analyses. Quantitative data sources included i-Ready Reports from BOY, MOY, and EOY; NC Check-Ins/NC Check-Ins 2.0 Results Reports from benchmarks #1, #2, and #3; NC EOG Results Reports; and a teacher survey. Qualitative data sources included a semi-structured interview.

Measure 1: Quantitative Data: i-Ready Reports

Reports existed in real-time within the i-Ready Reports online portal. To determine if teachers were implementing the i-Ready Diagnostic as intended, I used Diagnostic Status, which tracked progress to ensure all students successfully completed the i-Ready Diagnostic within each of the three testing windows. Figure 2 shows a sample i-Ready Diagnostic Status report.

Figure 2

Sample i-Ready Diagnostic Status Report



Diagnostic Status

i-Ready

School: Cole Middle School
Subject: Reading
Diagnostic: Recent Diagnostic

Students Assigned/Total: **621/621**

1%	2%	97%
Not Started	In Progress	Completed
	2 Expiring Soon	18 Rushed, 40 Rushed

Show Results By: Grade

Showing 3 of 3

Grade	% Not Started	In Progress		Completed			Students Assigned/Total
		% In Progress	Expiring Soon	% Completed	Red Rushed	Yellow Rushed	
Grade 6	1%	1%	2	98%	3	8	214/214
Grade 7	1%	1%	0	98%	2	9	189/189
Grade 8	1%	4%	0	95%	13	23	218/218

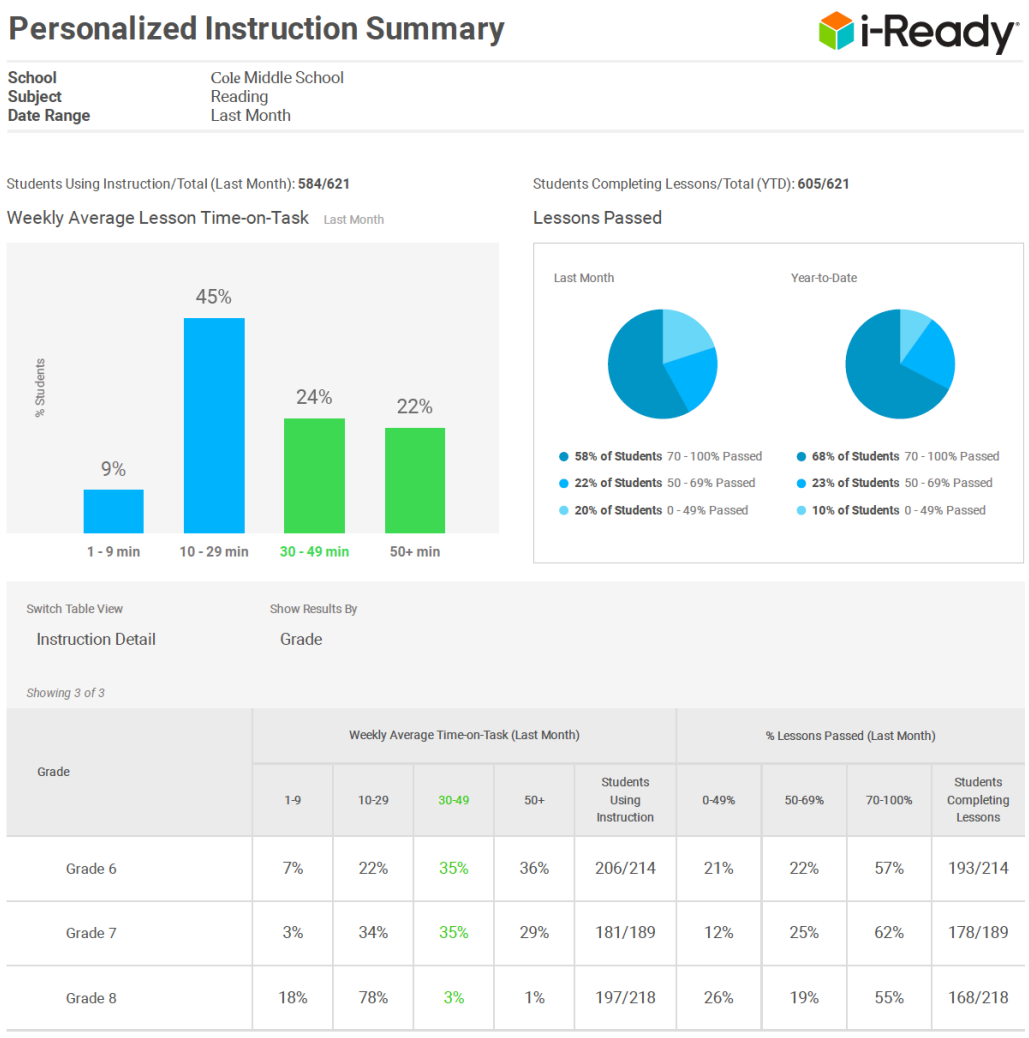
Note. This report is an example of an actual school's i-Ready report using a pseudonym in place of the actual school's name (Curriculum Associates, n.d.).

To determine if students were using the i-Ready Instruction as intended (i.e., 30-49 minutes per week for an average of 45 minutes per week), I used the Instructional Usage spreadsheet, which showed how much time individual students have spent in the personalized reading instruction by month as well as the percentage of lessons passed by month, and Personalized Instruction Summary, which tracked the amount of time on task and lessons passed and summarized if it was within the intended range (i.e., passing 70-100% of the completed lessons). The Personalized Instruction Summary showed the weekly average lesson time-on-task

in a bar graph and lessons passed in a pie chart. Figure 3 shows a sample Personalized Instruction Summary report.

Figure 3

Sample i-Ready Personalized Instruction Summary Report



Personalized Instruction Summary

i-Ready

School Cole Middle School
Subject Reading
Date Range Last Month

Students Using Instruction/Total (Last Month): **584/621**

Weekly Average Lesson Time-on-Task Last Month

Students Completing Lessons/Total (YTD): **605/621**

Lessons Passed

Switch Table View: Instruction Detail
Show Results By: Grade

Showing 3 of 3

Grade	Weekly Average Time-on-Task (Last Month)					% Lessons Passed (Last Month)			
	1-9	10-29	30-49	50+	Students Using Instruction	0-49%	50-69%	70-100%	Students Completing Lessons
Grade 6	7%	22%	35%	36%	206/214	21%	22%	57%	193/214
Grade 7	3%	34%	35%	29%	181/189	12%	25%	62%	178/189
Grade 8	18%	78%	3%	1%	197/218	26%	19%	55%	168/218

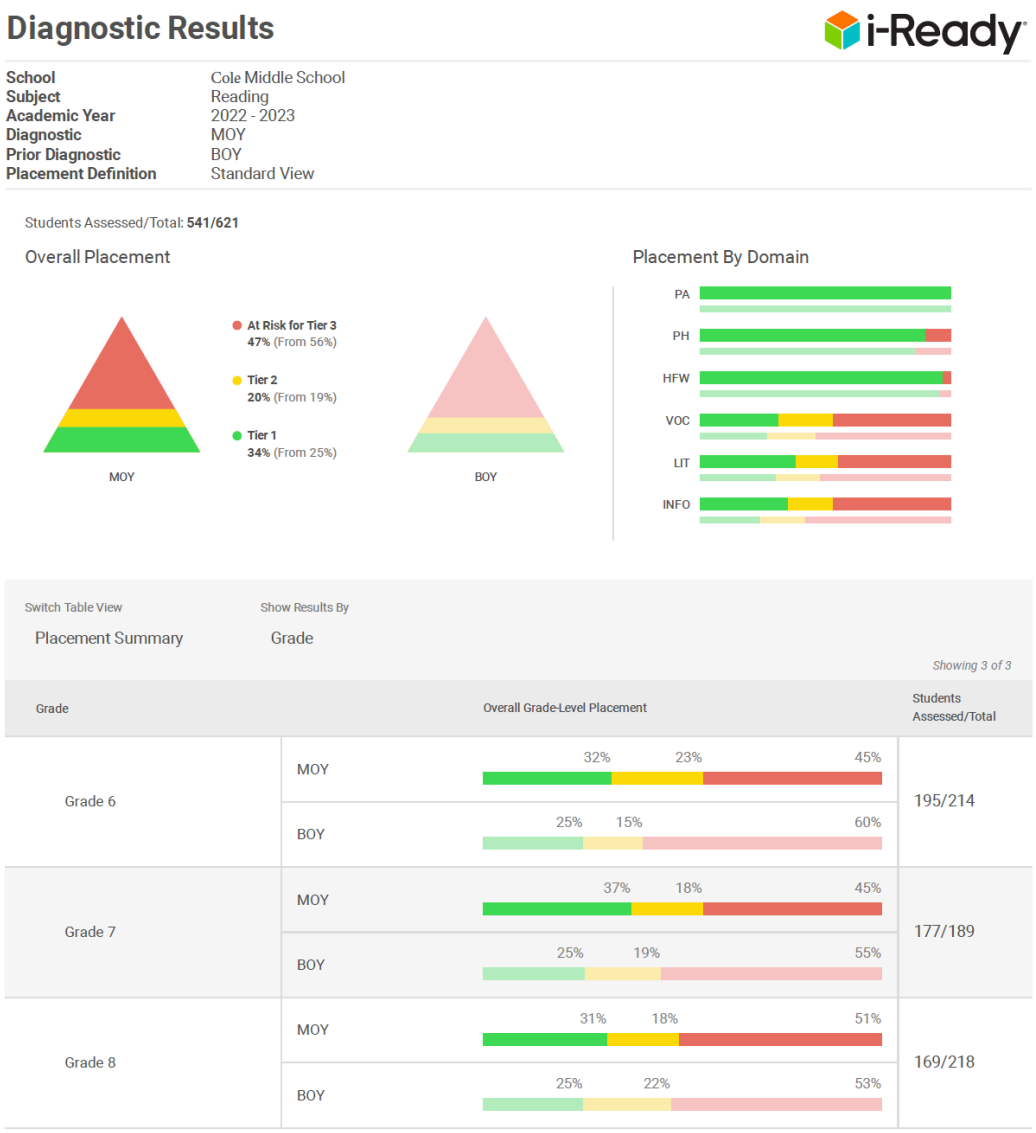
Note. This report is an example of an actual school's i-Ready report using a pseudonym in place of the actual school's name (Curriculum Associates, n.d.).

To determine if students' reading achievement on the i-Ready Diagnostic was increasing during the 2022-23 academic year, I used Diagnostic Results, which included students' overall

Diagnostic placement level on each specific domain and compared their performance across assessments. This chart included a prior diagnostic as well to illustrate growth between different assessment windows. Figure 4 shows a sample Diagnostic Results report from i-Ready.

Figure 4

Sample i-Ready Diagnostic Results Report

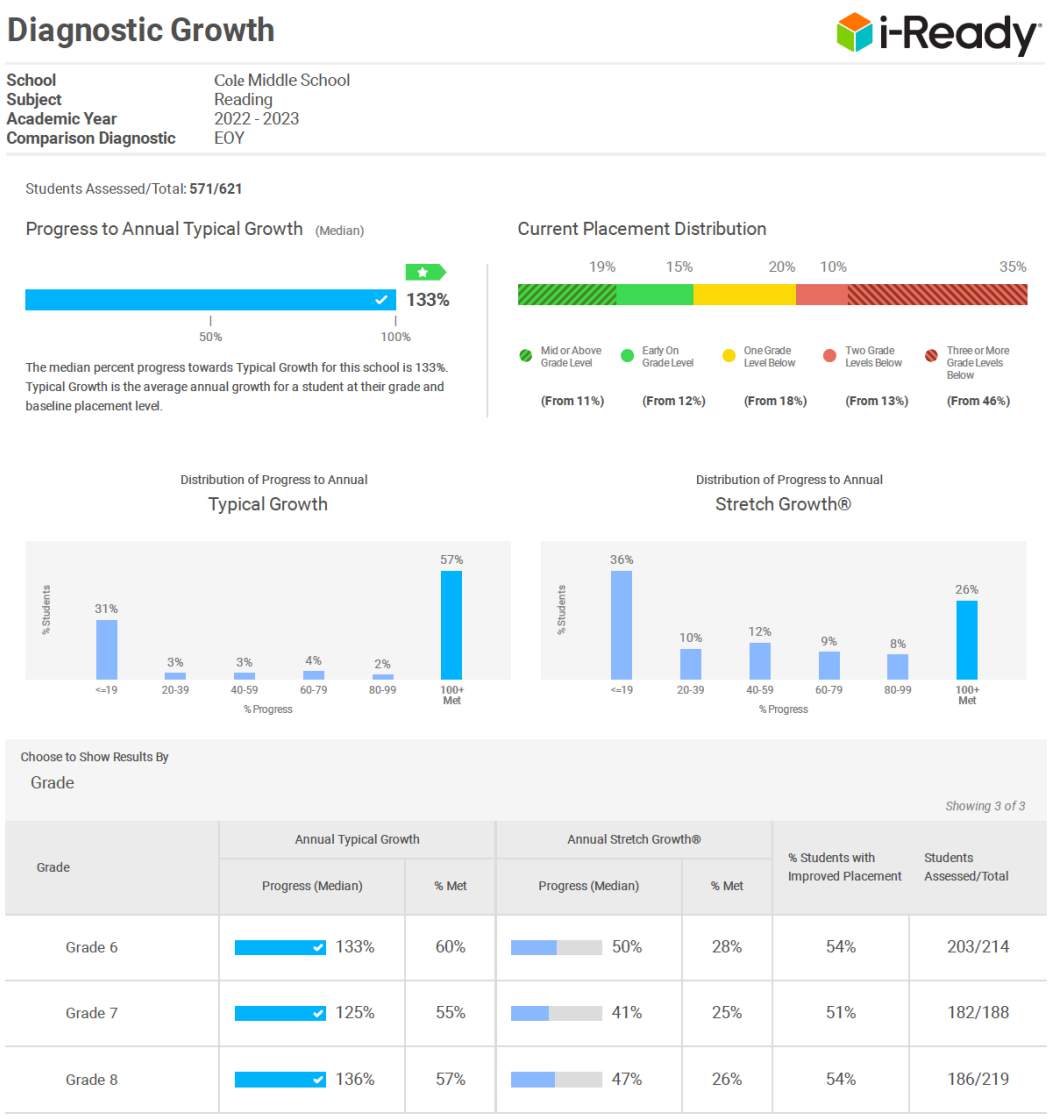


Note. This report is an example of an actual school's i-Ready report using a pseudonym in place of the actual school's name (Curriculum Associates, n.d.).

I also used Diagnostic Growth, which evaluated student growth and shared progress toward placement-based growth measures. These reports described usage of and growth on i-Ready Diagnostic and i-Ready Instruction for schools, grades, classrooms, and individual students. Figure 5 shows a sample Diagnostic Growth report from i-Ready.

Figure 5

Sample i-Ready Diagnostic Growth Report



Note. This report is an example of an actual school's i-Ready report using a pseudonym in place of the actual school's name (Curriculum Associates, n.d.).

According to the National Center on Intensive Intervention (n.d.), the i-Ready Diagnostic provided two types of reliability estimates: Item Response Theory-based measures (e.g., marginal reliability estimates and standard error of measurement [SEM]) and test-retest reliability coefficients. Marginal reliability operated on the variance of the theta scores and the

average of the expected error variance; its definition of reliability was the proportion of variance in the total observed score due to true score under the Item Response Theory model. The median coefficient was 0.97 for Grade 6, Grade 7, and Grade 8. SEM were affected by various factors including how well the data fit the underlying model, student response consistency, student location on the ability continuum, match of items to student ability, and test length. The median coefficient was 0.86 for Grade 6, 0.855 for Grade 7, and 0.853 for Grade 8. Test-retest reliability was provided because the i-Ready Diagnostic was often used as an interim assessment, and students could take the assessment multiple times per year. It provided stability estimates for the same students who took two i-Ready Diagnostic assessments. For validity, the internal structure of the i-Ready Diagnostic assessments was supported by the construct maps and ordering of skills addressed at different stages on the map. Its concurrent validity with the NC EOG was 0.82 for Grade 6, 0.81 for Grade 7, and 0.79 for Grade 8. Both the reliability and validity information were considered convincing evidence. As indicated on the Diagnostic Status Report, i-Ready flagged students for rushing to help educators identify students who may have demonstrated low effort on the i-Ready Diagnostic. Red flags indicated that a student spent less than eleven seconds on average per item, and yellow flags indicated that a student spent between twelve and fifteen seconds on average per item. Both the red and yellow flags pointed out that student scores may be lower than their true ability levels, so teachers were encouraged to have students with red flags re-take the i-Ready Diagnostic and consider if results were consistent for students with yellow flags.

Measure 2: Quantitative Data: NC Check-Ins/NC Check-Ins 2.0

NC Check-Ins and NC Check-Ins 2.0 were three interim assessments per grade level developed by NCDPI. Their main purpose was to provide stakeholders with actionable data and

a snapshot of students' current performance on the same reading standards but on passages of increasing text complexity throughout the year. NC Check-Ins have been replaced by NC Check-Ins 2.0, which were developed as a pilot study for a flexible summative assessment (i.e., the North Carolina Personalized Assessment Tool). NC Check-Ins 2.0 were used for Grades 7-8 students in 2022-23. NC Check-Ins were used for Grade 6 students in 2022-23 (NCDPI moved Grade 6 to NC Check-Ins 2.0 in 2023-24). Each NC Check-Ins (Grade 6) assessment contained three reading selections and a total of 22-24 multiple choice (four-response-option) questions (NCDPI, 2022a). Each NC Check-Ins 2.0 (Grades 7 and 8) assessment was online, fixed form, and contained 24 multiple-choice or technology enhanced items (NCDPI, 2022c). Assessments were not timed and could be administered in one school day or over multiple days; the estimated administration time was ninety minutes. While the NC Check-Ins/NC Check-Ins 2.0 shared a common item bank with NC EOGs, NCDPI did not have validity evidence to support using NC Check-Ins/NC Check-Ins 2.0 as predictor of student performance on NC EOGs (NCDPI, 2022a). According to personal communication with Dr. Kinge Mbella, NCDPI's lead psychometrician, who was in the process of compiling a full technical report for the NC Check-Ins, the average NC Check-Ins reliability across all grade levels was about .82, based on Cronbach's alpha internal consistency measure. An example of the reliability data is provided in Table 5.

Table 5

Reliability of Grade 7 NC Check-Ins 2.0 Scores

Assessment	NC Check-In 2.0 A	NC Check-In 2.0 B	NC Check-In 2.0 C
NC Check-In 2.0 A	(.827)	-	-
NC Check-In 2.0 B	.769	(.804)	-
NC Check-In 2.0 C	.783	.792	(.819)

Note. Values on the diagonal are Cronbach's alphas; values on the off-diagonal are correlation coefficients. (Source: Personal Communication with Dr. Kinge Mbella)

Measure 3: Quantitative Data: NC EOG Reading Tests

NC EOGs measured students' reading proficiency on the NC Standard Course of Study (updated in June 2017), and results were used for school accountability. In Grades 6-8, 36-41% of items were reading for literature, 43-47% were reading for informational text, and 11-16% were language. There are no Depth of Knowledge 1 questions; 60-82% were Depth of Knowledge 2 questions, and 18-40% were Depth of Knowledge 3 questions. There were six operational reading selections with five to eight operational items per selection as well as one field test selection with eight field test items, for a test total of seven reading selections with 52 items. Tests contained four-response-option multiple-choice and technology-enhanced item types. NCDPI estimated students needed 2 hours to complete the NC EOG, and the maximum time allowed was 3 hours (unless students have documented special needs requiring accommodations). Online administration was required, and the NC EOG was administered only during the last 10 days of the school year. NCDPI states that using NC educators to write and review items strengthens content validity of NC EOGs (NCDPI, 2022b). According to NCDPI (2021), the content of the test forms reflected the goals and objectives of the North Carolina Standard Course of Study for the subject (content validity) and reflected the goals and objectives as taught in North Carolina schools (instructional validity). Cronbach alpha reliability estimates

for the forms of the reading NC EOG were 0.89-0.90 (Grade 6 – three forms), 0.89-0.90 (Grade 7 – two forms), and 0.88-0.89 (Grade 8 – two forms). Additional validity evidence including standard error estimates, classification consistency summary of reported achievement levels, and an exploratory principal component analysis supported the unidimensional interpretation of the test scores as well as an EOG Lexile framework linking study and a summary of procedures used to ensure assessments are accessible and fair for all students.

Measure 4: Quantitative Data: Survey

Surveys are easy and quick to administer, can reach all teachers at the same time, provide a high volume of data quickly, and allow for anonymous responses (Mertens & Wilson, 2019). The survey protocol was created, and questions were purposefully sequenced as introductory (begin discussion of topic), transition (move from introduction to key questions), key (insights on central concern of the study) and ending (bring closure to the discussion; Krueger, 1998). The survey was designed to provide information about teachers' fidelity of implementation and perceptions of i-Ready, to encourage reflection on i-Ready program implementation and effectiveness in meeting intermediate-term outcomes of increased student reading achievement, and to compare the stakeholders' perceptions of their usage with their actual usage as indicated in i-Ready reports.

William & Mary professors with expertise in program evaluation and/or reading curricula reviewed the survey questions as part of the plan for expert review. Necessary revisions were made prior to pilot testing. Testing prior to implementation involved seven non-participant elementary school teachers, instructional staff, and administrators (four teachers, two instructional coaches, and one director) familiar with i-Ready to ensure participants in the study understood the questions in the same way and were comfortable during the process. Necessary

revisions were made following the pilot test. The pilot test and revisions built utility through relevant information and meaningful processes and products, feasibility through contextual viability, propriety through clarity and fairness, and accuracy through valid information and reliable information (Yarbrough et al., 2011). To ensure propriety and feasibility, all middle school reading teachers were emailed the survey with a rating scale via Google Forms as well as the survey participant informed consent form (see Appendix C). Survey questions are included with their question type and corresponding evaluation question in Table 6 and Appendix D.

Table 6

Alignment of Survey Questions and Evaluation Questions

Survey Question	Question Type	Evaluation Question
What grade do you teach?	I	
In which academic years have you used i-Ready?	T	3
Respond to the following questions using this scale: (a) daily, (b) weekly, (c) monthly, (d) quarterly, (e) yearly, or (f) never. 1. I use i-Ready's Historical Reports to see my class's testing history. 2. I use i-Ready's Historical Reports to see an individual student's testing history. 3. I lead data chats with my students using i-Ready Diagnostic Results Report and Diagnostic Growth Report. 4. I use the Diagnostic Growth Report to create and discuss typical and stretch growth goals with students. 5. I use i-Ready's Diagnostic Results Report to create student groups to intervene on specific skill deficits. 6. I use the Personalized Instruction Summary Report to see students' weekly usage of i-Ready Instruction. 7. I use the Personalized Instruction Summary Report to see students' lessons passed in i-Ready Instruction. 8. I use the instructional usage spreadsheet to guide conversations with students about their i-Ready Instruction usage and progress. 9. I use Teacher Toolbox to provide specific interventions that are personalized to their needs to small groups or individual students. 10. I use Think Up!-ELA resources with my students to integrate critical thinking instruction on ELA standards. 11. I use Ready-NC resources with my students to build conceptual understanding through reasoning, practice, and productive discussion around real-world scenarios. 12. I use Ready Reading Discourse Cards with my students to promote richer, deeper dialogue to engage all students in meaningful ELA conversations. 13. I use PHONICS for Reading to help struggling students become fluent and independent readers. 14. I use the Grade Level Scaffolding Report to plan small groups and/or use reading buddies.	K	1
Respond to the following questions using this scale: (a) always, (b) often, (c) sometimes, or (d) never. 15. I prepare students to take the i-Ready Diagnostic by using the preparation materials found in i-Ready Central. 16. I administer and proctor the i-Ready Diagnostic for all students in my class three times per year. 17. I chunk the i-Ready Diagnostic over multiple days, so students do not test for more than 30 minutes per day. 18. I use i-Ready's Diagnostic Status Report to check student progress on the i-Ready Diagnostic. 19. I reset the i-Ready Diagnostic for individual students whose data on the Diagnostic Status Report indicates a red flag or yellow flag for rushing.	K	1
Respond to the following questions using this scale: (a) yes, (b) no, or (d) I am not sure. 20. Over 80% of students completed all three i-Ready Diagnostic assessments in 2022-23. 21. Over 80% of my students used i-Ready Instruction for 30-49 minutes per week for at least 18 weeks in 2022-23. 22. Over 80% of my students passed 70-100% of lessons in i-Ready Instruction in 2022-23.	K	1

Survey Question	Question Type	Evaluation Question
Respond to the following questions using this scale: (a) strongly agree, (b) agree, (c) disagree, or (d) strongly disagree. 23. I have a scheduled weekly time for my students to use i-Ready Reading that ensures they access it 30-49 minutes per week.	K	1
Rate the value of each of the following i-Ready program elements using this scale: (a) very valuable, (b) somewhat valuable, or (c) not valuable. 24. Student data generated from the i-Ready Diagnostic assessment 25. Student instruction and practice in i-Ready Instruction 26. Instructional resources provided in Teacher Toolbox 27. Overall impact that i-Ready is having on student achievement in reading	K	3
Should the middle school continue, modify, or discontinue i-Ready use?	E	3
If you selected "modify," please share what should be modified.	E	3
Is there anything else you would like to share about the i-Ready program?	E	1, 2, 3

Note. I = Introductory question; T = Transitional question; K = Key question; E = Ending question

Measure 5: Qualitative Data: Semi-Structured Individual Interviews

Semi-structured interviews are well-suited for asking probing, open-ended questions and learning individual thoughts from members of a group of key stakeholders (Adams, 2015). While all participants responded to the same interview guide to support consistency of the data sets, the nature of semi-structured interviews allowed participants to elaborate on specific questions.

To plan for expert review, William & Mary professors with expertise in program evaluation and/or reading curricula reviewed the interview protocol and questions. The protocol was pilot tested prior to implementation using elementary school instructional staff and administrators familiar with i-Ready to ensure participants understood the questions in the same way and were comfortable during the process, and revisions were made. The pilot test and revisions built utility through relevant information and meaningful processes and products, feasibility through contextual viability, propriety through clarity and fairness, and accuracy through valid information and reliable information (Yarbrough et al., 2011).

Interviews were conducted with all consenting middle school administrators, instructional staff, and teachers to learn more about their perceptions of i-Ready in terms of guiding program elements. It was expected that perceptions may differ based on participant roles. These

differences were considered during analysis. Each interview was designed to last for 45-60 minutes but went no longer than 45 minutes. The interview began with introductions, purpose of the interviews, pledging confidentiality, and explaining and securing informed consent (Adams, 2015). Once participants signed the informed consent, I used the interview guide (see Appendix E) and engaged in rapport building through customary pleasantries. Interview questions were purposefully sequenced as opening (participants become comfortable), introductory (begin discussion of topic), transition (move from introduction to key questions), key (insights on central concern of the study) and ending (bring closure to the discussion; Krueger, 1998). Probes were used to ask for more information and for an explanation of ideas (Creswell & Creswell, 2018). Following interviews, I used member checking by transcribing interviews and sending to participants to ensure accuracy. Questions for the semi-structured interviews are included with their question type and aligned with their corresponding evaluation question in Table 7 and Appendix F.

Table 7

Alignment of Semi-Structured Interview Questions and Evaluation Questions

Semi-Structured Interview Question	Question Type	Evaluation Question
What do you find most rewarding about being an educator?	O	
What are some of the online curriculum, instruction, and assessment programs that you have observed teachers or students using in the middle school?	I	
When you were introduced to using i-Ready, what were your initial impressions?	T	3
What are the most valuable aspects of the i-Ready Diagnostic in reading? *(What aspects are most beneficial to teachers as they provide reading instruction? What aspects are most beneficial to students as they participate in reading instruction?)*	K	3
What are the most valuable aspects of the i-Ready Instruction in reading? *(What aspects are most beneficial to teachers as they provide reading instruction? What aspects are most beneficial to students as they participate in reading instruction?)*	K	3
What are the most valuable aspects of the i-Ready Teacher Toolbox in reading?	K	3
Describe how you use data generated from i-Ready Diagnostic to inform your lesson planning.	K	1
Describe how you have seen students' reading proficiency improve after participation in i-Ready Instruction. (*Can you tell me a story about a student whose reading proficiency or enthusiasm for reading increased?)*	K	2
How do you ensure that students are meeting the required amount of time using i-Ready Instruction each week?	K	1
What barriers exist to students meeting the required amount of time using i-Ready Instruction each week?	K	1
What other instructional strategies (or programs) are you using to supplement reading instruction? – *Listen for acceleration, personalization, gamification*	K	1, 3
What are the challenges of implementing the i-Ready program in reading as it is intended? What ideas do you have for overcoming these challenges?	K	1, 3
Describe the types of reading instruction you provide to students who test out of i-Ready Instruction.	K	1
In what ways has implementing the i-Ready program impacted student reading proficiency?	K	1, 2, 3
What ideas do you have to modify the i-Ready program, so it will be more effective in reading?	E	3
Is there anything else we should have talked about but did not?	E	1, 2, 3

Note. O = Opening question; I = Introductory question; T = Transitional question; K = Key question; E = Ending question

Data Collection

Qualitative data from the semi-structured interviews were collected to support the larger quantitative data set; however, dialogue occurred between the quantitative data sets and findings of the qualitative analysis and converged to determine overall results (Creswell & Creswell, 2018; Mertens & Wilson, 2019). Quantitative data were collected from i-Ready Reports, NC Check-Ins/NC Check Ins 2.0, and NC EOGs for all enrolled middle school students in 2022-23 as well as from the teacher survey. Reports from i-Ready and assessment data were collected at the end of the 2022-23 academic year. Qualitative data were collected from semi-structured interviews as well as open-ended questions in the survey. All participants were invited through written and personal contact with the evaluator, who gave (orally and in the written informed consent form) assurance of participant rights, confidentiality, and ability to withdraw from the evaluation without harm (Appendices C and D). Surveys were disseminated during Summer 2023. Semi-structured interviews were scheduled at a convenient time during Fall 2023 for participants, digitally recorded, and transcribed. Following survey and interview completion, I sent participants a thank-you note to show my gratitude for them (Adams, 2015).

Data Analysis

Both quantitative and qualitative data were analyzed to evaluate implementation of the i-Ready program. Descriptive statistics and repeated measures ANOVA were used for quantitative measures. Open coding and axial coding to determine themes across participant groups were used for qualitative measures.

Evaluation Question 1: To what extent did teachers in the selected school implement with fidelity the i-Ready program (i.e., i-Ready Diagnostic and i-Ready Instruction) in the 2022-23 academic year?

Quantitative i-Ready Diagnostic completion results (from the three administrations at the beginning, middle, and end of the 2022-23 school year) as appear on the Diagnostic Status Report and i-Ready Instruction usage (weekly minutes and lessons passed) as appear on the Instructional Usage spreadsheet and Personalized Instruction Summary Report were disaggregated by grade level and classroom teacher with descriptive statistics (frequencies) using SPSS software programming (Creswell & Creswell, 2018). I identified trends including in fidelity of administration of the i-Ready Diagnostic (i.e., three administrations) and fidelity of usage of the i-Ready Instruction (i.e., 0-9 minutes, 10-29 minutes, 30-49 minutes, and 50 or more minutes per week for at least 18 weeks) across grade levels and classrooms. Also, there was quantitative analysis of survey rating scale responses (using frequencies) to compare staff's reported fidelity with i-Ready usage reports.

Evaluation Question 2: To what extent, if any, did student reading proficiency improve during the 2022-23 academic year as measured by the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOGs?

Quantitative i-Ready Diagnostic and NC Check-Ins/NC Check-Ins 2.0 achievement results (from the three administrations at the beginning, middle, and end of the 2022-23 school year) as well as NC EOG scores (from the last 3 years) were disaggregated by grade level and/or classroom teacher with descriptive statistics (mean, standard deviation, and range) and repeated measures ANOVA using SPSS software programming (Creswell & Creswell, 2018). I identified trends including reading achievement changes on individual measures to see if students were

advancing as well as to see if the outcomes across the three final assessments were in alignment. I used a repeated measures ANOVA to analyze the quantitative i-Ready Diagnostic and NC Check-Ins/NC Check-Ins 2.0 achievement results since there were three assessments during the year. I looked for trends by examining the progress at each of the three test administrations. Because the NC EOG is only assessed once a year, I conducted a correlational analysis on the NC EOG, the final i-Ready Diagnostic, and the final NC Check-Ins/NC Check-Ins 2.0.

Evaluation Question 3: What were the perceptions of teachers, instructional staff, and administrators regarding i-Ready in terms of guiding program elements such as the following: (a) the value of student data generated by the i-Ready Diagnostic, (b) the value of the i-Ready Instruction as an instructional resource, and (c) the extent to which i-Ready is having an impact on student achievement?

Qualitative responses from semi-structured interviews were read in entirety, examined individually for their gist or major topics, clustered by topics that become codes, annotated, and organized by codes, reanalyzed for new categories or codes, and recombined to create summary statements to create the key themes of the program evaluation (Creswell & Creswell, 2018; Saldaña, 2016; Tesch, 1990). Open coding (also known as initial coding) broke down qualitative data into distinct parts, closely studied them and compared them; it employed in vivo coding by incorporating the interviewee's actual words (Saldaña, 2016). After open coding, axial coding related codes to each other and constructed linkages among data. Participants' perceptions could not be generalized to others who have implemented the i-Ready program. Table 8 represents the table of specification that includes the evaluation questions, data sources and methods of analysis.

Table 8

Aurelius School District's Evaluation Plan of the i-Ready Program

Evaluation Question	Data Sources	Data Analysis
Fidelity of Implementation Output (Process) 1. To what extent did teachers in the selected school implement the i-Ready program (i.e., i-Ready Diagnostic and i-Ready Instruction) as intended in the 2022-23 academic year?	i-Ready Reports (for all Grade 6-8 students)	Descriptive statistics (frequencies)
Intermediate-Term Outcome (Product) 2. To what extent, if any, did student reading proficiency improve during the 2022-23 academic year as measured by the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOGs?	i-Ready Reports (for all Grade 6-8 students) NC Check-Ins (for all Grade 6 students) NC Check-Ins 2.0 (for all Grade 7-8 students) NC EOGs (for all Grade 6-8 students)	Quantitative comparison of 2022-23 BOY, MOY, and EOY reading scores disaggregated by grade level and classroom with descriptive statistics (mean, standard deviation, and range) Repeated Measures ANOVA (changes in i-Ready and NC Check-Ins/NC Check-Ins 2.0 over all three testing administrations) Correlation Analysis (all three EOY data points)
Output (Process) 3. What were the perceptions of teachers, instructional staff, and administrators regarding the i-Ready program in terms of guiding program elements such as the following: a. The value of student data generated by the i-Ready Diagnostic b. The value of the i-Ready Instruction as an instructional resource c. The extent to which i-Ready is having an impact on student achievement	Semi-structured interviews Survey	Quantitative analysis of rating scale responses from survey questions using descriptive statistics (frequencies) Qualitative analysis of themes emerging from interviews using open and axial coding

Note. NC = North Carolina; EOG = End of Grade; BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year; ANOVA = analysis of variance.

Delimitations, Limitations, and Assumptions

Delimitations

This evaluation was delimited by the evaluator's decisions affecting the scope and focus of the evaluation. This evaluation of the i-Ready program was delimited by the choice of evaluating the intermediate outcomes of the program in the middle school context only. Some of

the activities and inputs, including professional development and the role of professional learning communities, were excluded. Additionally, the evaluation did not examine the role and perceptions of some key stakeholders, such as students, parents, and other school staff. These delimitations were based on feasibility and because the program is in its second year of implementation.

Limitations

Limitations are weaknesses within a study that may influence outcomes and conclusions of the research (Ross & Bibler Zaidi, 2019). Growth on the i-Ready Diagnostic or NC Check-Ins/NC Check-Ins 2.0 could be attributed to instructional resources or initiatives other than i-Ready. Qualitative data were collected with surveys and interviews from staff who reported on their own work, so the possibility of self-reporting bias must be considered. Findings were not generalizable to other schools in the district. Because North Carolina mandated LETRS professional development and reading curriculum aligned with the science of reading, these state mandates could have affected the district's ability to implement i-Ready as intended. Lastly, while the district was only in its second year of implementing i-Ready, staff turnover occurred, which could have affected its effectiveness in using i-Ready.

Assumptions

Assumptions and external factors were included in the logic model (Figure 1). The overarching assumption was the i-Ready program was a program that when implemented with fidelity could produce increased student achievement. Likewise, there was an assumption that the i-Ready program was internally aligned, so weekly usage of i-Ready Instruction as intended would increase i-Ready Diagnostic, which correlated with increased NC Check-Ins/NC Check-Ins 2.0 and NC EOG scores (Curriculum Associates, 2022a). There were other assumptions

inherent to the i-Ready program as well. Consultants with i-Ready facilitated robust, effective professional development, so staff had a thorough understanding of the i-Ready platform, increased self-efficacy in using i-Ready, and ability to onboard new users to the i-Ready suite. The district facilitated effective professional learning communities (PLCs) with i-Ready data dives and i-Ready pacing guides, so staff could implement resources from Teacher Toolbox; prepare for NC EOGs; use i-Ready assessment data and pacing to improve differentiated core (Tier I) instruction, targeted small group (Tier II) interventions, and intensive personalized (Tier III) interventions; and have districtwide alignment, coherence, and consistency in reading assessment, instruction, and intervention. Teachers administered the i-Ready Diagnostic with fidelity, so students' scores are valid and reliable. Finally, district culture promoted continuous improvement, so stakeholders were eager to understand i-Ready's impact and supportive of the evaluation, professional development, professional learning community meetings, i-Ready Diagnostic, and i-Ready Instruction.

Ethical Considerations

Ethical considerations can be categorized using the Joint Commission on Standards for Program Evaluation (Yarbrough et al., 2011). This evaluation considered utility, feasibility, propriety, and accuracy. All of these categories were important for a robust program evaluation.

Utility standards center on stakeholders finding evaluation processes and products useful. Since the beginning of i-Ready implementation, staff had been curious about whether the time students spent on i-Ready Instruction was translating into increased reading achievement as measured by the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and the NC EOG. Surveys and interviews allowed staff to express, rediscover, reinterpret, or revise their views to

the evaluator, who had pre-established credibility as a district administrator and had control over continuing, modifying, or discontinuing use of the i-Ready program.

Feasibility standards are intended to increase the effectiveness and efficiency of the evaluation. All delimitations occurred to enhance feasibility. Teacher surveys were used, so teachers could respond at a convenient time instead of having to schedule with the evaluator. Individual interviews allowed individuals to share their views without worrying their views conflict with another group member, and open-ended questions about what they find rewarding as an educator helped to clarify individual values underpinning their judgments. Evaluation results were shared with participants in a timely, ongoing way throughout the process.

Propriety standards support that which is proper, right, fair, legal, and just. This evaluation was responsive to the stakeholders and the community, as there was a high monetary cost in implementing the i-Ready program, so it should be used with fidelity and modified or discontinued if it was not meeting its intended outcomes. Conflict of interest was openly acknowledged by the evaluator, who initiated the i-Ready contract and believed it would increase student achievement, by assuring participants there were no right or wrong answers and no retaliation for honestly sharing. I maintained a reflexive journal and conducted member checks to monitor my actions to ensure I did not overlay my biases on data interpretations, to triangulate data, and to increase clarity and fairness as well as transparency and disclosure. Additionally, one of the district's Board members works for Curriculum Associates, but she did not benefit monetarily from the district's i-Ready contract, and she disclosed her possibly perceived conflict of interest prior to any Board votes. Interview protocols included sharing that findings were shared with district leaders regarding the future of the i-Ready program, that individual responses were confidential, and that participants could leave without penalty at any time.

Accuracy standards should increase dependability and truthfulness of evaluation findings through valid and reliable information, sound program and context description, and strong procedures for collecting and analyzing data and justifying and communication findings. Although findings were not generalizable, they could be helpful starting points in other contexts and schools.

This study was submitted to the William & Mary Education Institutional Review Committee (EDIRC) for approval prior to conducting any program evaluation. ASD's process to gather data is informal and required a discussion with the superintendent. I had permission to gather data and evaluate ASD's implementation of i-Ready. Participants in surveys and semi-structured interviews were invited through electronic and in-person contact with the evaluator. Assurance of participants' rights, confidentiality, and their ability to withdraw at any time from the evaluation were given orally and in writing through the participant informed consent forms for the survey and for semi-structured interviews (see Appendices B and C). Survey responses, interview notes, and other data collected and analyzed in the program evaluation have been kept confidential and have not been used in any employee's personnel evaluation.

CHAPTER 4

FINDINGS

This study aimed to evaluate the fidelity of implementation, impact on student reading proficiency, and perceptions of users regarding i-Ready Reading in Aurelius School District to determine if the district should continue, modify, or discontinue the program in Cole Middle School. Using a convergent mixed methods program evaluation design (Creswell & Creswell, 2018) following Stufflebeam's (2003) Context, Input, Process, and Product (CIPP) Evaluation Model, this study provided district leaders with information and recommendations based on quantitative and qualitative data, including i-Ready Diagnostic and Instruction reports, NC Check-Ins/NC Check-Ins 2.0 results, NC EOG achievement scores, survey results, and semi-structured interviews, to understand i-Ready's effect on middle school reading achievement. The previous chapter outlined the methods used to collect and analyze data in the study. This chapter is structured by evaluation question and includes data analysis and findings for each question.

During the 2022-23 school year, Cole Middle School had 641 different students enrolled in grades six, seven, and eight who participated in at least one of the following reading assessments: beginning-of-year (BOY) i-Ready Diagnostic, middle-of-year (MOY) i-Ready Diagnostic, end-of-year (EOY) i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0 #1, NC Check-Ins/NC Check-Ins 2.0 #2, NC Check-Ins/NC Check-Ins 2.0 #3, or NC EOG. On the last day of the 2022-23 school year, there were 621 students enrolled in Cole Middle School in Grade 6 six (n = 214), Grade 7 (n = 188), and Grade 8 (n = 219). Of these 621 students, 51% were male

and 49% were female; 49% were white, 34% were black, 10% were Hispanic/Latino; 5% were two or more race categories, and 2% were Asian. While all student scores were included in individual descriptive statistics for each assessment, only students with matched scores were included in the repeated measures ANOVA and correlation analysis. Table 9 shows the number of student participants for each assessment.

Table 9

Cole Middle School Student Participants in 2022-23 Assessments

Grade	Students	NC Check-Ins			i-Ready Diagnostic			NC EOG
		1	2	3	BOY	MOY	EOY	
6	219	202	205	207	207	211	210	211
7	194	184	185	185	189	189	184	185
8	228	212	213	208	205	212	195	215

Note. *N* = 641. NC = North Carolina. BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year. NC EOG = North Carolina End of Grade.

Evaluation Question 1: To what extent did teachers in the selected school implement with fidelity the i-Ready program (i.e., i-Ready Diagnostic and i-Ready Instruction) in the 2022-23 academic year?

Evaluation question one investigated the fidelity of i-Ready implementation during the 2022-23 academic year in Cole Middle School. The i-Ready Diagnostic Status Report, i-Ready Diagnostic Results Report, and i-Ready Personalized Instruction Summary provided data for analysis. Fidelity was measured by examining the percentage of students who participated in the three i-Ready Diagnostic assessments, the percentage of students passing 70-100% of lessons in i-Ready Instruction, and the amount of weekly student usage minutes in i-Ready Instruction.

i-Ready Diagnostic

Implementing the i-Ready Diagnostic with fidelity means that all enrolled students complete three assessments (i.e., one BOY, one MOY, and one EOY i-Ready Diagnostic). These assessment data create a personalized pathway for each student in i-Ready Instruction and are intended to be used by teachers to create intervention groups and provide small-group instruction using resources found in Teacher Toolbox. The i-Ready Diagnostic Status Reports and i-Ready Diagnostic Results Reports provided data to determine how many students participated in each i-Ready Diagnostic. Of the 641 different students who took at least one reading assessment during the 2022-23 school year, 619 students were enrolled during the i-Ready BOY assessment window (i.e., September 6-27, 2022), 625 students were enrolled during the i-Ready MOY assessment window (i.e., January 3-31, 2023), and 616 students were enrolled during the i-Ready EOY assessment window (i.e., May 1-19, 2023).

All students enrolled during each assessment window were expected to take each assessment available during that testing window, but for various reasons, not all of them participated in all three assessments. During the 2022-23 school year, 601 students completed the i-Ready BOY assessment, 612 students completed the i-Ready MOY assessment, and 589 students completed the i-Ready EOY assessment. Table 10 contains the total number of individual students who completed at least one assessment in 2022-23, the number of students enrolled during each i-Ready Diagnostic assessment window, the total number who completed each i-Ready Diagnostic assessment, and the percentage of completed versus enrolled students for each assessment. Results indicated that 98% or more of enrolled Grade 6 and Grade 7 students and 90% or more of enrolled Grade 8 students completed i-Ready Diagnostic assessments in 2022-23.

Table 10

i-Ready Diagnostic Completion by Grade Level

Grade	Students	BOY			MOY			EOY		
		Completed	Enrolled	%	Completed	Enrolled	%	Completed	Enrolled	%
6	219	207	208	99%	211	212	99%	210	213	98%
7	194	189	189	100%	189	192	98%	184	187	98%
8	228	205	222	92%	212	221	96%	195	216	90%
Total	641	601	619	97%	612	625	98%	589	616	96%

Note. $N = 641$. BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year.

Teacher perceptions of student assessment completion aligned with data collected in i-Ready Reports. When surveyed about their perceptions of student completion of the i-Ready Diagnostic assessments in 2022-23, seven of the eight teachers affirmed that over 80% of their students completed all three assessments, and one teacher said they were not sure. As shown in Table 10, the Diagnostic Status Report indicated that sixth grade and seventh grade had nearly perfect completion rates on the i-Ready Diagnostic assessments (between 98% and 100% completion). The eighth-grade completion rate was lower (between 90% and 96%) than sixth and seventh grades. All eight teachers reported on the survey that they *consistently* administer and proctor the i-Ready Diagnostic for all students in their class three times per year; however, only two of the eight teachers administered all three i-Ready Diagnostic assessments to all their enrolled students. For example, one teacher did not administer 26 of the 281 i-Ready Diagnostic assessments, which is almost 10%, expected to be administered to the 111 different students on their roster. Six teachers reported on the survey that they *consistently* use the i-Ready Diagnostic Status Report to check student progress on the i-Ready Diagnostic, and two teachers reported they *frequently* use it. These data are summarized in Table 11.

Table 11

Missing i-Ready Diagnostic by Classroom Teacher

Grade	Teacher	Missing Scores			
		BOY	MOY	EOY	Total
6	A	0	1	2	3
6	B	0	0	0	0
6	C	1	0	1	2
7	A	0	2	3	5
7	B	0	0	0	0
7	C	0	1	0	1
8	A	8	3	10	21
8	B	9	6	11	26
All School		18	13	27	58

Note. n = 58. BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year.

i-Ready Instruction

Implementing i-Ready Instruction with fidelity meant that all enrolled students engaged with the personalized instruction pathway for 45 minutes per week and passed at least 70% of the i-Ready lessons. Students' performance on the i-Ready Diagnostic provided an individual entry point for the personalized pathway, which was designed to close specific skill gaps through a series of online, scaffolded lessons. The i-Ready Personalized Instruction Summary provided data to determine the lesson time-on-task and percentage of lessons passed. During the 2022-23 school year, 621 individual students used i-Ready Instruction. Of these students, 214 students were in Grade 6, 188 students were in Grade 7, and 219 students were in Grade 8. To determine the fidelity of usage, expected usage was determined by multiplying the number of students enrolled in each teacher's class on the last day of the 2022-23 school year by 45, the number of minutes i-Ready should have been used by students each week. One limitation of examining expected usage was the different variables related to student absenteeism such as student tardiness, student discipline, student mobility, teacher absences, or other testing windows. As shown in Table 12, the i-Ready Personalized Instruction Summary provided the average weekly

usage minutes by month disaggregated by teacher; expected usage was only met or exceeded in one month (either February 2023 or March 2023) by only four teachers (two in Grade 6 and two in Grade 7).

Table 12

i-Ready Instruction Average Weekly Usage Minutes by Classroom Teacher

Grade	Expected Usage	Sept. 2022	Oct. 2022	Nov. 2022	Dec. 2022	Jan. 2023	Feb. 2023	Mar. 2023	Apr. 2023	May 2023
6-A	2070	521	695	966	716	804	1502	1077	71	433
6-B	3375	1202	2159	2090	1977	2336	3159	3811	1985	1859
6-C	3285	1378	2282	3129	1528	2393	4465	3129	1923	2898
7-A	2970	1409	1771	1960	1747	2250	2587	3002	1366	896
7-B	2205	1821	2008	1802	1136	1986	2073	2511	780	1017
7-C	2970	2224	2625	1893	1748	2110	2297	1851	644	992
8-A	4950	693	1131	1457	1032	333	1519	1624	789	32
8-B	4770	674	178	1184	513	911	669	1419	832	225
Total	26595	9922	12849	14481	10397	13123	18271	18424	8390	8352

Note. Expected usage is the number of students who used i-Ready Instruction in each regular education classroom multiplied by 45 minutes per week. The i-Ready Personalized Instruction Summary Report provides average weekly usage minutes for each classroom teacher for each month as well as an Excel spreadsheet with each classroom for every month in the year. Grade 6, $n = 194$. Grade 7, $n = 181$. Grade 8, $n = 216$.

To examine trends across grade levels, the expected usage for the total number of students in a grade level was determined by multiplying all students by 45 minutes. Next, each classroom teacher's average usage was combined for each month to determine the grade level monthly average usage. These data are shown in Table 13, and as a grade level, expected usage was not met during the 2022-23 school year.

Table 13

i-Ready Instruction Average Weekly Usage Minutes by Grade Level

Grade	Expected Usage	Sept. 2022	Oct. 2022	Nov. 2022	Dec. 2022	Jan. 2023	Feb. 2023	Mar. 2023	Apr. 2023	May 2023
6	9630	3476	5591	6795	4505	5752	9444	8316	4380	5781
7	8640	5630	6889	6229	4990	6731	7364	7693	3028	3250
8	9855	1379	1347	2801	1545	1357	2218	3063	1621	338
Total	28125	10485	13827	15825	11040	13840	19026	19072	9029	9369

Note. Expected usage represents the expected usage hours when all students participate in i-Ready instruction for 45 minutes per week as required. The i-Ready Personalized Instruction Summary Report provides average weekly usage minutes for each classroom teacher for each month as well as an Excel spreadsheet with each classroom for every month in the year. Grade 6, $n = 214$. Grade 7, $n = 188$. Grade 8, $n = 219$.

Each month's actual average weekly usage minutes' total for the grade level was divided by the expected usage minutes to determine the percentage of usage for each grade level. The actual percentage varied between 1% (i.e., Grade 8 in May 2023) and 98% (i.e., Grade 6 in February 2023) of expected usage. Grade 6 ranged from 36% (September 2022) to 98% (February 2023). Grade 7 ranged from 36% (April 2023) to 91% (March 2023). Grade 8 ranged from 1% (May 2023) to 31% (March 2023). Average usage for the 9 months between September 2022 and May 2023 was 62% for Grade 6, 68% for Grade 7, and 17% for Grade 8. These data are shown in Table 14.

Table 14

i-Ready Instruction Percentage of Actual Versus Expected Usage by Grade Level

Grade	Expected Usage	Sept. 2022	Oct. 2022	Nov. 2022	Dec. 2022	Jan. 2023	Feb. 2023	Mar. 2023	Apr. 2023	May 2023
6	9630	36%	58%	71%	47%	60%	98%	86%	45%	60%
7	8640	67%	81%	74%	59%	80%	87%	91%	36%	38%
8	9855	14%	14%	28%	16%	14%	23%	31%	16%	1%

Note. Expected usage is the number of students who used i-Ready Instruction in each grade level multiplied by 45 minutes per week, and 100% for the month represents the program expectation of averaging 45 minutes per week. The i-Ready Personalized Instruction Summary Report provides average weekly usage minutes for each classroom teacher for each month as well as an Excel spreadsheet with each classroom for every month in the year. Grade 6, n = 214. Grade 7, n = 188. Grade 8, n = 219.

Data suggested that classroom teachers used i-Ready Instruction inconsistently, and usage did not match the program's expectations for all students for 45 minutes per week. In 2 months, 4 teachers exceeded the number of minutes recommended for i-Ready Instruction usage (i.e., February 2023 for Grade 6-C and March 2023 for Grade 6-B, Grade 7-A, and Grade 7-B). Data indicated that individual teachers' usage varied within a grade level. For example, Grade 6-A ranged from 3% to 73%, Grade 6-B ranged from 36% to 113%, and Grade 6-C ranged from 42% to 136% of expected usage. The lowest usage month for Grade 6-B and Grade 6-C was September 2022, and Grade 6-A's lowest usage was April 2023. Grade 7-A ranged from 30% to 101%, Grade 7-B ranged from 35% to 114%, and Grade 7-C ranged from 22% to 88% of expected usage. The lowest usage month for Grade 7-B and Grade 7-C was April 2023, and the lowest usage month for Grade 7-A was May 2023. By far, the grade level with the lowest usage of i-Ready Instruction was Grade 8. Data showed that Grade 8-A ranged from 1% to 33%, and Grade 8-B ranged from 4% to 30% of expected usage. The lowest usage month for Grade 8-B was October 2022 and for Grade 8-A was May 2023. Table 15 shows i-Ready Instruction percentage of actual usage versus expected usage disaggregated by classroom teacher.

Table 15

i-Ready Instruction Percentage of Actual Versus Expected Usage by Classroom Teacher

Grade	Expected Usage	Sept. 2022	Oct. 2022	Nov. 2022	Dec. 2022	Jan. 2023	Feb. 2023	Mar. 2023	Apr. 2023	May 2023
6-A	2070	25%	34%	47%	35%	39%	73%	52%	3%	21%
6-B	3375	36%	64%	62%	59%	69%	94%	113%	59%	55%
6-C	3285	42%	69%	95%	47%	73%	136%	95%	59%	88%
7-A	2970	47%	60%	66%	59%	76%	87%	101%	46%	30%
7-B	2205	83%	91%	82%	52%	90%	94%	114%	35%	46%
7-C	2970	75%	88%	64%	59%	71%	77%	62%	22%	33%
8-A	4950	14%	23%	29%	21%	7%	31%	33%	16%	1%
8-B	4770	14%	4%	25%	11%	19%	14%	30%	17%	5%

Note. Expected usage is the number of students who used i-Ready Instruction in each regular education classroom multiplied by 45 minutes per week, and 100% for the month represents the program expectation of averaging 45 minutes per week. The i-Ready Personalized Instruction Summary Report provides average weekly usage minutes for each classroom teacher for each month as well as an Excel spreadsheet with each classroom for every month in the year. Grade 6, n = 194. Grade 7, n = 181. Grade 8, n = 216.

Of the 641 individual students who took at least one reading assessment during 2022-23, 621 students used i-Ready Instruction. Of these 621 students, 606 students completed at least one lesson. For the 2022-23 school year, 15 students in Grades 6, 7, and 8 accessed i-Ready Instruction but never completed a lesson. One limitation of examining lessons completed included the different variables related to student absenteeism such as student tardiness, student discipline, student mobility, teacher absences, or other testing windows.

Of the 606 students who completed at least one lesson in i-Ready Instruction, 65%, which was 406 students, passed 70-100% of lessons. The highest percentage of lessons passed was in seventh grade with 72%; eighth grade and sixth grade had 63% and 62%, respectively. It was concerning that 33% of students engaged with i-Ready Instruction but passed less than 70% of their lessons when the platform provided opportunities for students to retake lessons and receive

additional interventions facilitated in-person through Teacher Toolbox. One limitation of reporting lessons passed included the different variables related to student absenteeism such as student tardiness, student discipline, student mobility, teacher absences, or other testing windows. These data are shown in Table 16.

Table 16

i-Ready Instruction Lessons Passed

		1+ lessons completed	0 lessons completed		1-49% lessons passed		50-69% lessons passed		70-100% lessons passed	
Grade	Total students	No.	No.	%	No.	%	No.	%	No.	%
6	214	211	3	1%	20	9%	59	28%	132	62%
7	188	186	2	1%	10	5%	41	22%	135	72%
8	219	209	10	5%	29	13%	41	19%	139	63%
Total	621	606	15	2%	59	10%	141	23%	406	65%

Note. Expected lesson passage rate is 70-100%. The i-Ready Personalized Instruction Summary Report as well as an Excel spreadsheet provide percentage of lessons passed for each classroom teacher.

Teacher perceptions of student average weekly usage of i-Ready Instruction were not closely aligned with data collected in i-Ready Reports. Lesson participation data were collected in each grade level over a 9-month period, totally 27 months of instruction. As indicated in Table 15, out of the total 27 months for the three grade levels (9 months each), only 6 months had 80% or higher of usage; of those six, Grade 6 had 2 months and Grade 7 had 4 months. When surveyed about i-Ready Instruction weekly student usage, 7 of the 8 teachers affirmed that over 80% of their students used i-Ready Instruction for 30-49 minutes per week for at least 18 weeks in 2022-23, and one teacher said they were not sure. Six teachers reported they *strongly agree* they have a scheduled weekly time for their students to use i-Ready Instruction that ensures students access it 30-49 minutes per week, and two teachers reported they *agree*. Six teachers said they use the Personalized Instruction Summary Report *weekly* to see their students' weekly

usage of i-Ready Instruction; one teacher reported using it *monthly*; and one teacher reported *never* using it. Teachers reported much lower usage of the Personalized Instruction Summary Report to monitor student performance related to lessons passed: four teachers reported using it *weekly*, three teachers reported using it *monthly*, and one teacher reported *never* using it. Likewise, four teachers reported they were *not sure* if over 80% of their students passed 70-100% of lessons in i-Ready Instruction in 2022-23, two teachers reported *no*, and two teachers reported *yes*. Lastly, two teachers reported using the instructional usage spreadsheet *monthly* to guide conversations with their students about their i-Ready Instruction usage and progress, three teachers reported using it *quarterly*, and three teachers reported *never* using it.

Summary of Data Analysis of Evaluation Question 1

In summary, there was fidelity of implementation in administration of the i-Ready Diagnostic. Over 95% of all enrolled students participated in the i-Ready Diagnostic, which established a personalized instructional pathway for each student in i-Ready Instruction. Overall, there was a lack of fidelity in implementation of i-Ready Instruction as measured by program expectations of student usage of 45 minutes per week and passing 70% or more lessons. Meeting or exceeding program expectations for usage only occurred in four classrooms for one of the nine months in the 2022-23 school year and was not met during any month for an entire grade level. Similarly, almost 40% of students did not pass 70-100% of lessons.

Evaluation Question 2: To what extent, if any, did student reading proficiency improve during the 2022-23 academic year as measured by the i-Ready Diagnostic, NC Check-Ins/ NC Check-Ins 2.0, and NC EOGs?

Evaluation Question 2 investigated the extent to which, if any, student reading proficiency improved during the 2022-23 academic year in Cole Middle School as measured by

the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOG. The i-Ready Diagnostic Results Report, NC Check-Ins/NC Check-Ins 2.0 Results Reports, and NC EOGs Results Report provided data for analysis. Improvement was measured by examining the changes in mean scores for each administration of each assessment as well as the correlation of scores on all final assessments.

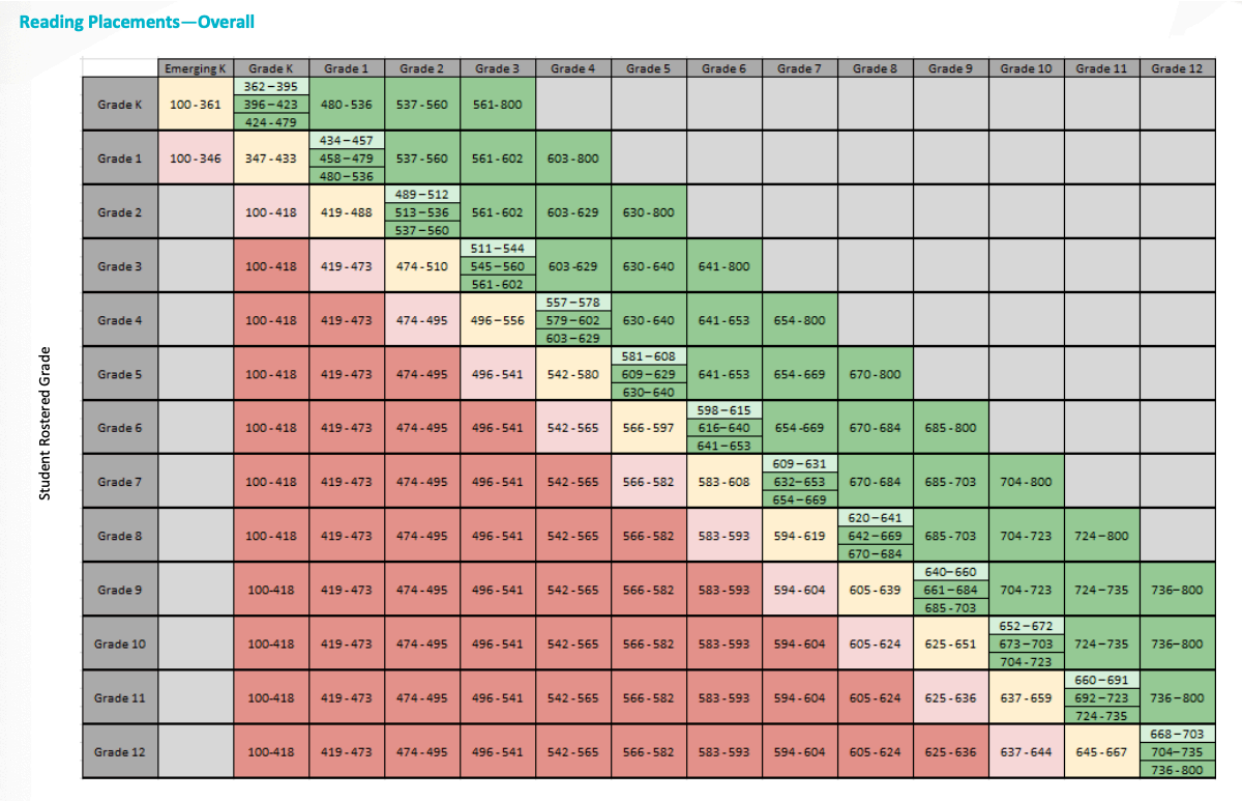
i-Ready Diagnostic

The i-Ready Diagnostic assessments helped teachers make appropriate instructional recommendations and placements for students who were performing at different levels along the K-12 spectrum (Curriculum Associates, 2023b). Upon completion of an assessment, i-Ready calculated an overall (for all items) and domain-level (for items within a given area) scale score based on the difficulty levels of items each student received. Scale scores were vertically aligned across years, so a student's scores could be compared from year to year. These scores were compared to placement tables to provide each student with a grade-level placement, which was an indication of a student's performance based on grade level. Grade-level placements were based on specific scale score ranges for each chronological grade, which was the grade in which a student was currently enrolled, and were criterion referenced, so they reflected what students were expected to know at each grade level. Students received an absolute placement, which was the grade level associated with their performance, and for students who placed on grade level, i-Ready indicated a relative placement of early on grade level, mid-grade level, and late grade level, which were relative to their grade level. While students were considered on grade level if their scale score was between 598 and 653 (Grade 6), 609 and 669 (Grade 7), and 620 and 684 (Grade 8), Curriculum Associates stated that the decision to deem a student as proficient should be left up to individual school districts. They noted that students who achieved the mid-grade

level placement were generally considered to have achieved grade-level proficiency; however, proficiency on the i-Ready Diagnostic did not necessarily have the same meaning as proficiency on an EOY state assessment. Figure 6 shows i-Ready Diagnostic's reading placement table with students' rostered grade in the horizontal rows and reading grade-level placement in the vertical columns.

Figure 6

i-Ready Diagnostic Reading Placement Table



Reading Placements—Overall

Student Rostered Grade

	Emerging K	Grade K	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12
Grade K	100-361	362–395 396–423 424-479	480-536	537-560	561-800									
Grade 1	100-346	347-433	434–457 458–479 480–536	537-560	561-602	603-800								
Grade 2		100-418	419-488	489–512 513–536 537–560	561-602	603-629	630-800							
Grade 3		100-418	419-473	474-510	511–544 545–560 561-602	603-629	630-640	641-800						
Grade 4		100-418	419-473	474-495	496–556	557–578 579–602 603–629	630-640	641-653	654-800					
Grade 5		100-418	419-473	474-495	496-541	542-580	581–608 609–629 630–640	641-653	654-669	670-800				
Grade 6		100-418	419-473	474-495	496-541	542-565	566-597	598–615 616–640 641–653	654-669	670-684	685-800			
Grade 7		100-418	419-473	474-495	496-541	542-565	566-582	583-608	609–631 632–653 654–669	670-684	685-703	704-800		
Grade 8		100-418	419-473	474-495	496-541	542-565	566-582	583-593	594-619	620–641 642–669 670–684	685-703	704-723	724-800	
Grade 9		100-418	419-473	474-495	496-541	542-565	566-582	583-593	594-604	605-639	640–660 661–684 685-703	704-723	724–735	736–800
Grade 10		100-418	419-473	474-495	496-541	542-565	566-582	583-593	594-604	605-624	625-651	652–672 673–703 704-723	724–735	736–800
Grade 11		100-418	419-473	474-495	496-541	542-565	566-582	583-593	594-604	605-624	625-636	637-659	660–691 692–723 724-735	736–800
Grade 12		100-418	419-473	474-495	496-541	542-565	566-582	583-593	594-604	605-624	625-636	637-644	645-667	668–703 704–735 736-800

Note. Green indicates mid, late, or above grade level. Light green indicates early on grade level. Yellow indicates one grade level below. Pink indicates two grade levels below. Red indicates three or more grade levels below. Gray indicates a score at this grade level is not possible (N/A). From "i-Ready Diagnostic Grades K-12 Scale Score Placement Tables," by Curriculum Associates, 2023, (https://i-readycentral.com/pdfs/i-ready-scale-score-placement-tables). Copyright 2023 by Curriculum Associates.

Sixth-grade students showed improvement on the i-Ready Diagnostic during the 2022-23 school year. Students scoring early, mid, late, or above grade level increased from 23% at the beginning of the year to 36% at the end of the year. Students scoring three or more grade levels below decreased from 42% at the beginning of the year to 29% at the end of the year. Table 17 shows scale scores for Grade 6 during all 2022-23 testing administrations of the i-Ready Diagnostic.

Table 17

Grade 6 Students i-Ready Diagnostic Scale Scores

Test	3+ grade levels below 100-541	2 grade levels below 542-565	1 grade level below 566-597	Early on grade level 598-615	Mid, late, or above grade level 616-800
BOY	87 (42%)	41 (20%)	31 (15%)	19 (9%)	29 (14%)
MOY	66 (31%)	33 (16%)	47 (22%)	17 (8%)	48 (23%)
EOY	61 (29%)	34 (16%)	39 (19%)	28 (13%)	48 (23%)

Note. BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year. BOY, $n = 207$. MOY, $n = 211$. EOY, $n = 210$.

Seventh-grade students showed improvement on the i-Ready Diagnostic during the 2022-23 school year. Students scoring early, mid, late, or above grade level increased from 26% at the beginning of the year to 37% at the end of the year. Students scoring three or more grade levels below decreased from 46% at the beginning of the year to 34% at the end of the year. Table 18 shows scale scores for grade seven during all 2022-23 testing administrations of the i-Ready Diagnostic.

Table 18

Grade 7 Students i-Ready Diagnostic Scale Scores

Test	3+ grade levels below	2 grade levels below	1 grade level below	Early on grade level	Mid, late, or above grade level
	100-565	566-582	583-608	609-631	632-800
BOY	87 (46%)	20 (11%)	34 (18%)	30 (16%)	18 (10%)
MOY	71 (38%)	18 (10%)	32 (17%)	34 (18%)	34 (18%)
EOY	63 (34%)	13 (7%)	41 (22%)	31 (17%)	36 (20%)

Note. BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year. BOY, n = 189. MOY, n = 189. EOY, n = 184.

Eighth-grade students showed improvement on the i-Ready Diagnostic during the 2022-23 school year. Students scoring early, mid, late, or above grade level increased from 24% at the beginning of the year to 32% at the end of the year. Students scoring three or more grade levels below decreased from 48% at the beginning of the year to 42% at the end of the year. Changes from the beginning of the year to the middle of the year, however, were not as positive as the number of students scoring at three or more grade levels below and two grade levels below both increased, which was a pattern not occurring at grades six and seven. Table 19 shows scale scores for grade eight during all 2022-23 testing administrations of the i-Ready Diagnostic.

Table 19

Grade 8 Students i-Ready Diagnostic Scale Scores

Test	3+ grade levels below 100-582	2 grade levels below 583-593	1 grade level below 594-619	Early on grade level 620-641	Mid, late, or above grade level 642-800
BOY	98 (48%)	14 (7%)	44 (21%)	29 (14%)	20 (10%)
MOY	103 (49%)	16 (8%)	38 (18%)	24 (11%)	31 (15%)
EOY	82 (42%)	13 (7%)	38 (19%)	33 (17%)	29 (15%)

Note. BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year. BOY, $n = 205$. MOY, $n = 212$. EOY, $n = 195$.

The i-Ready Diagnostic scale score means increased between each administration in all three grade levels; however, the increases were sometimes less than one point (e.g., the Grade 6 change from middle to end of the year was 0.4, and the Grade 8 change from beginning to middle of year was 0.9). According to the reading placement table (Figure 6), the means of three Grade 6 administrations were designated in the one grade level below range. The mean of the first Grade 7 administration was designated in the two grade levels below range, and the other two administrations were designated in the one grade level below range. The means of the first two Grade 8 administrations were in the two grade levels below range, and the final mean was in the one grade level below range. Overall, from beginning to end of year, Grade 6 increased 17.1 points, Grade 7 increased 14.8 points, and Grade 8 increased 12.5 points. Table 20 indicates the

ranges and means of scale scores as well as the changes in means between administrations for all 2022-23 i-Ready Diagnostic assessments disaggregated by grade level.

Table 20

i-Ready Diagnostic Assessment Results by Grade Level

	BOY Scale Score			MOY Scale Score				EOY Scale Score				
Grade	Lowest	Highest	*M*	Lowest	Highest	*M*	Change in BOY to MOY *M*	Lowest	Highest	*M*	Change in MOY to EOY *M*	Change in BOY to EOY *M*
6	340	678	547.5	321	681	564.2	16.7	277	682	564.6	0.4	17.1
7	360	682	561.0	317	698	574.1	13.1	330	722	575.8	1.7	14.8
8	337	712	574.1	303	718	575.0	0.9	268	719	586.6	11.6	12.5

Note. BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year. Grade 6 BOY, *n*=207; MOY, *n*=211; EOY, *n*=210. Grade 7 BOY, *n*=189; MOY, *n*=189; EOY *n*=184. Grade 8 BOY, *n*=205; MOY, *n*=212; EOY, *n*=195.

Unlike the grade levels combined, the i-Ready Diagnostic scale score means disaggregated by classroom did not increase between each administration in all classrooms. Although most classrooms increased (from 8.4 to 23.2 points) from beginning to middle of year, one classroom (i.e., Grade 8-A) decreased by 4.3 points. From middle to end of year, three classrooms (i.e., Grade 6-A, Grade 6-C, and Grade 7-A) decreased (from -8 to -2.9 points) and five classrooms increased (from 0.9 to 11.9 points). All classrooms increased from beginning to end of year between 3.7 (i.e., Grade 6-C) to 27.3 points (i.e., Grade 6-B). According to the reading placement table (Figure 6), the means of the beginning-of-year administration were three or more grade levels behind in three classrooms and two grade levels behind in five classrooms; the means of the middle-of-year administration were two grade levels behind in two classrooms and one grade level behind in five classrooms; and the means of the end-of-year administration were three or more grade levels behind in two classrooms; two grade levels behind in one classroom; and one grade level behind in five classrooms. Overall, classroom means increased

between 3.7 and 27.3 points in Grade 6, 5.8 and 22.7 points in Grade 7, and 7.6 and 20.4 points in Grade 8. Table 21 indicates the ranges and means of scale scores as well as the changes in means between administrations for all 2022-23 i-Ready Diagnostic assessments disaggregated by classrooms.

Table 21

i-Ready Diagnostic Assessment Results by Classroom Teacher

	BOY Scale Score			MOY Scale Score				EOY Scale Score				
Grade	Lowest	Highest	*M*	Lowest	Highest	*M*	Change in BOY to MOY *M*	Lowest	Highest	*M*	Change in MOY to EOY *M*	Change in BOY to EOY *M*
6-A	498	623	562.7	513	642	585.9	23.2	480	649	583.0	-2.9	20.3
6-B	382	678	555.3	421	681	576.1	20.8	428	682	582.6	6.5	27.3
6-C	440	666	560.2	473	666	568.6	8.4	277	666	563.9	-4.7	3.7
7-A	360	655	554.7	317	664	568.5	13.8	330	659	560.5	-8	5.8
7-B	404	659	576.7	440	666	589.6	12.9	401	652	599.4	9.8	22.7
7-C	409	682	570	414	698	585.6	15.6	407	722	586.5	0.9	16.5
8-A	359	712	573.3	393	718	569.0	-4.3	378	703	580.9	11.9	7.6
8-B	407	693	578.0	409	704	587.0	9	412	719	598.4	11.4	20.4

Note. BOY = beginning-of-year; MOY = middle-of-year; EOY = end-of-year. **Grade 6-A** BOY, *n*=43; MOY, *n*=44; and EOY, *n*=44. **Grade 6-B** BOY, *n*=72; MOY, *n*=74; and EOY, *n*=75. **Grade 6-C** BOY, *n*=69; MOY, *n*=72; and EOY, *n*=72. **Grade 7-A** BOY, *n*=64; MOY, *n*=64; and EOY, *n*=63. **Grade 7-B** BOY, *n*=49; MOY, *n*=49; and EOY, *n*=49. **Grade 7-C** BOY, *n*=65; MOY, *n*=65; and EOY, *n*=65. **Grade 8-A** BOY, *n*=102; MOY, *n*=109; and EOY, *n*=99. **Grade 8-B** BOY, *n*=99; MOY, *n*=101; and EOY, *n*=94.

Since students took the i-Ready Diagnostic three times in the 2022-23 year, these data were analyzed using a one-way within-subjects, or repeated measures, ANOVA. For Grade 6, there were 198 students who participated in all three i-Ready Diagnostic assessments (M = 548.3 [BOY], 565.7 [MOY], 564.6 [EOY]; SD = 63.3 [BOY], 60.0 [MOY], 65.7 [EOY]). For Grade 7, there were 179 students who participated in all three i-Ready Diagnostic assessments (M = 562.3 [BOY], 576.5 [MOY], 577.6 [EOY]; SD = 63.7 [BOY], 65.5 [MOY], 66.2 [EOY]). For Grade 8,

there were 174 students who participated in all three i-Ready Diagnostic assessments (M = 573.3 [BOY], 578.0 [MOY], 588.9 [EOY]; SD = 67.2 [BOY], 71.1 [MOY], 66.2 [EOY]). One of the statistical analyses, the Tests of Within-Subjects Effects, which includes Sphericity Assumed, indicated the means of the i-Ready Diagnostic were significantly different during at least one of the time points, for Grade 6, $F(2, 394) = 39.09$, $p < .001$, partial $\eta^2 = .166$; Grade 7: $F(2, 356) = 28.50$, $p < 001$, partial $\eta^2 = .138$; and Grade 8, $F(2, 346) = 17.16$, $p < .001$, partial $\eta^2 = .090$. Table 22 shows repeated measures ANOVA results.

Table 22

Summary of Repeated Measures ANOVA Results for Comparison of BOY, MOY, and EOY i-Ready Diagnostic Assessments

Grade	Test	*M*	*SD*	*F*	*df*1	*df*2	partial η^2
6				39.09	2	394	.166
	BOY	548.3	63.3				
n = 198	MOY	565.7	60.0				
	EOY	564.6	65.7				
7				28.50	2	356	.138
	BOY	562.3	63.7				
n = 179	MOY	576.5	65.5				
	EOY	577.6	66.2				
8				17.16	2	346	.090
	BOY	573.3	67.2				
n = 174	MOY	578.0	71.1				
	EOY	588.9	66.2				

Note. ANOVA = analysis of variance. BOY = beginning-of-year. MOY = middle-of-year. EOY = end-of-year. $p < .001$.

NC Check-Ins/NC Check-Ins 2.0

NC Check-Ins and NC Check-Ins 2.0 were NCDPI-created interim assessments that provided a snapshot of students' current performance on the same reading standards but on passages of increasing text complexity throughout the year. NC Check-Ins 2.0 were used for

Grades 7 and 8 in 2022-23, and NC Check-Ins were used for Grade 6 students in 2022-23. Each NC Check-Ins (Grade 6) assessment contained three reading selections and a total of 22-24 multiple choice (four-response-option) questions (NCDPI, 2022a). Each NC Check-Ins 2.0 (Grades 7 and 8) assessment was online, fixed form, and contained 24 multiple-choice or technology enhanced items (NCDPI, 2022c). Assessments were not timed and could be administered in 1 school day or over multiple days; the estimated administration time was ninety minutes. While the number of students participating in the NC Check-Ins/NC Check-Ins 2.0 varied slightly, it did not appear to affect the overall means.

Means of the number of correct items on the NC Check-Ins/NC Check-Ins 2.0 increased between the first and second assessments in all three grade levels (between 0.6 and 2.1 points); however, all three means decreased from the second to third assessments in all three grade levels (between -2.0 and -0.5 points). Overall, from the first to the third assessment, Grade 6 increased 0.1, Grade 7 increased 0.4, and Grade 8 increased 0. Table 23 indicates the ranges and means of the number correct as well as the changes in means between administrations for all 2022-23 NC Check-Ins/NC Check-Ins 2.0 disaggregated by grade level.

Table 23

NC Check-Ins/NC Check-Ins 2.0 Assessment Results by Grade Level

	C1			C2				C3				
Grade	Lowest Number Correct	Highest Number Correct	*M*	Lowest Number Correct	Highest Number Correct	*M*	Change in C1 to C2 *M*	Lowest Number Correct	Highest Number Correct	*M*	Change in C2 to C3 *M*	Change in C1 to C3 *M*
6	3	23	13.2	1	24	15.3	2.1	2	22	13.3	-2.0	0.1
7	2	24	11.4	2	24	12.3	0.9	2	23	11.8	-0.5	0.4
8	0	23	11.3	0	23	11.9	0.6	2	24	11.3	-0.6	0

Note. NC = North Carolina. C1 = NC Check-Ins/NC Check-Ins 2.0 #1. C2 = NC Check-Ins/NC Check-Ins 2.0 #2. C3 = NC Check-Ins/NC Check-Ins 2.0 #3. Grade 6 C1, *n*=202; C2, *n* =205; and C3, *n* =207. Grade 7 C1, *n* =184; C2, *n* =185; and C3, *n*=185. Grade 8 C1, *n* =212; C2, *n* =213; and C3, *n* =208.

Like the grade levels combined, the means of the NC Check-Ins/NC Check-Ins 2.0 number of items correct disaggregated by classroom increased between the first and second administrations in all classrooms (between 0.1 and 3.2 points). While most classrooms' means decreased (between -3.5 and -0.3 points) from the second to the third assessments, one classroom's mean (i.e., Grade 7-B) increased by 0.2 point. From the first to the third assessment, four of the eight classrooms' means increased (from 0.1 to 1.2 points), two classrooms' means remained consistent, and two classrooms' means decreased (between -0.3 and -0.1 point). Table 24 indicates the ranges and means of the number correct as well as the changes in means between administrations for all 2022-23 NC Check-Ins/NC Check-Ins 2.0 disaggregated by classrooms.

Table 24

NC Check-Ins/NC Check-Ins 2.0 Assessment Results by Classroom Teacher

	C1			C2				C3				
Grade	Lowest Number Correct	Highest Number Correct	*M*	Lowest Number Correct	Highest Number Correct	*M*	Change in C1 to C2 *M*	Lowest Number Correct	Highest Number Correct	*M*	Change in C2 to C3 *M*	Change in C1 to C3 *M*
6-A	5	23	14.3	7	24	16.4	2.1	7	20	14.3	-2.1	0
6-B	3	23	14.6	3	24	16.2	1.6	6	22	15.8	-0.4	1.2
6-C	3	22	12.4	4	24	15.6	3.2	4	21	12.1	-3.5	-0.3
7-A	2	23	9.7	3	21	11.5	1.8	2	23	10.7	-0.8	1.0
7-B	3	22	13.3	2	20	13.4	0.1	6	22	13.6	0.2	0.3
7-C	3	24	12.1	2	24	12.8	0.7	2	23	12.1	-0.7	0
8-A	2	23	11.2	0	22	11.4	0.2	2	23	11.1	-0.3	-0.1
8-B	0	21	11.4	0	23	12.3	0.9	2	24	11.5	-0.8	0.1

Note. NC = North Carolina. C1 = NC Check-Ins/NC Check-Ins 2.0 #1. C2 = NC Check-Ins/NC Check-Ins 2.0 #2. C3 = NC Check-Ins/NC Check-Ins 2.0 #3. **Grade 6-A** C1, *n*=43; C2, *n*=44; and C3, *n* =44. **Grade 6-B** C1, *n*=72; C2, *n*=73; and C3, *n*=73. **Grade 6-C** C1, *n*=69; C2, *n*=70; and C3, *n*=73. **Grade 7-A** C1, *n*=64; C2, *n*=64; and C3, *n*=65. **Grade 7-B** C1, *n*=48; C2, *n*=48; and C3, *n*=49. **Grade 7-C** C1, *n*=65; C2, *n*=66; and C3, *n*=66. **Grade 8-A** C1, *n*=104; C2, *n*=109; and C3, *n*=109. **Grade 8-B** C1, *n*=108; C2, *n*=104; and C3, *n*=99.

Since students took the NC Check-Ins/NC Check-Ins 2.0 three times in the 2022-23 year, these data were analyzed using a one-way within-subjects, or repeated measures, ANOVA. For Grade 6, there were 195 students who participated in all three NC Check-Ins assessments (M = 13.2 [BOY], 15.5 [MOY], 13.5 [EOY]; SD = 5.0 [BOY], 5.5 [MOY], 5.5 [EOY]). For Grade 7, there were 177 students who participated in all three NC Check-Ins 2.0 (M = 11.5 [BOY], 12.5 [MOY], 12.0 [EOY]; SD = 5.5 [BOY], 4.9 [MOY], 4.9 [EOY]). For Grade 8, there were 195 students who participated in all three NC Check-Ins 2.0 (M = 11.6 [BOY], 12.0 [MOY], 11.4 [EOY]; SD = 4.7 [BOY], 5.2 [MOY], 4.7 [EOY]). One of the statistical analyses, the Tests of Within-Subjects Effects, which includes Sphericity Assumed, indicated the means of the NC Check-Ins/NC Check-Ins 2.0 were significantly different during at least one of the time points, for Grade 6, $F(2, 388) = 42.57$, $p < .001$, partial $\eta^2 = .180$; Grade 7: $F(2, 352) = 6.50$, $p = .002$, partial $\eta^2 = .036$; and Grade 8, $F(2, 388) = 3.17$, $p = .043$, partial $\eta^2 = .016$. Table 25 shows the summary of repeated measures ANOVA results.

Table 25

Summary of Repeated Measures ANOVA Results for Comparison of NC Check-Ins/NC Check-Ins 2.0 #1, #2, and #3 Assessments

Grade	NC Check-Ins/ NC Check-Ins 2.0	*M*	*SD*	*F*	*df*1	*df*2	partial η^2	*p*
6				42.57	2	388	.180	< .001
	C1	13.2	5.0					
n = 195	C2	15.5	5.5					
	C3	13.5	5.5					
7				6.50	3	352	.036	.002
	C1	11.5	5.5					
n = 177	C2	12.5	4.9					
	C3	12.0	4.9					
8				3.17	2	388	.016	.043
	C1	11.6	4.7					
n = 195	C2	12.0	5.2					
	C3	11.4	4.7					

Note. ANOVA = analysis of variance. NC = North Carolina. C1 = NC Check-In/NC Check-In 2.0 #1. C2 = NC Check-In/NC Check-In 2.0 #2. C3 = NC Check-In/NC Check-In 2.0 #3.

End-of-Year Assessments

North Carolina End of Grade (NC EOG) assessments measured students' reading proficiency on the North Carolina Standard Course of Study (updated in June 2017), and results were used for school accountability. There were six operational reading selections with five to eight operational items per selection as well as one field test selection with eight field test items, for a test total of seven reading selections with 52 items. Tests contained four-response-option multiple-choice and technology-enhanced item types. NCDPI estimated students needed 2 hours to complete the NC EOG, and the maximum time allowed was 3 hours (unless students had documented special needs requiring accommodations). Online administration was required, and

the NC EOG was administered only during the last ten days of the school year (NCDPI, 2022b). For the last 3 years, Cole Middle School's growth designations in each grade level remained consistent (i.e., Grade 6 exceeded expected growth; Grade 7 met expected growth; Grade 8 did not meet expected growth). In 2022-23, 43.9% of sixth graders, 40.6% of seventh graders, and 38.7% of eighth graders were proficient on the NC EOG. Table 26 indicates NC EOG scores for 2020-21, 2021-22, and 2022-23 as well as the number of students assessed.

Table 26

North Carolina End of Grade (NC EOG) Assessment Results

	2020-21			2021-22			2022-23		
Grade	No. Students	% Proficient	Growth Status	No. Students	% Proficient	Growth Status	No. Students	% Proficient	Growth Status
6	201	37.6%	Exceeded	186	33.9%	Exceeded	214	43.9%	Exceeded
7	216	42.3%	Met	214	40.2%	Met	187	40.6%	Met
8	251	38.7%	Not Met	229	40.6%	Not Met	217	38.7%	Not Met

Note. These data include students on the extended content standards, who took an alternative assessment, and have not been reflected in earlier tables. For 2022-23, there were 211 sixth graders, 185 seventh graders, and 215 eighth graders who took the actual NC EOG test.

NC EOG scores were reported in four levels: Not Proficient, Level 3, Level 4, and Level 5. Cut scores were slightly different in each grade level. Table 27 shows the 2022-23 NC EOG students' scores disaggregated by grade level and score level.

Table 27

2022-23 NC End of Grade Assessment Results Disaggregated by Score Level

Grade	Not Proficient	Level 3	Level 4	Level 5
6	119 (56%)	48 (23%)	34 (16%)	10 (5%)
7	111 (60%)	36 (19%)	29 (16%)	9 (5%)
8	133 (62%)	42 (19%)	28 (13%)	12 (6%)

Note. Scores vary by grade level as follows: Grade 6: Not Proficient = < 551; Level 3 = 552–557; Level 4 = 558 – 566; Level 5 = > 567. Grade 7: Not Proficient = < 553; Level 3 = 554 – 558; Level 4 = 559 – 565; Level 5 = > 566. Grade 8: Not Proficient = < 556; Level 3 = 557 – 562; Level 4 = 563 – 571; Level 5 = > 572.

A Pearson product-moment correlation coefficient (Pearson's *r*) was computed to assess the relationships among the EOY i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0 #3, and NC EOG. There was a large positive association between each set of assessments for all grade levels ranging between .740 and .804. Because results on the EOY i-Ready Diagnostic appeared to be correlated with results on the NC Check-Ins/NC Check-Ins 2.0 #3 and NC EOG, administrators, teachers, and other instructional staff should feel confident that making data-based instructional decisions for students using i-Ready Diagnostic results positively impacts student achievement on NC Check-Ins/NC Check-Ins 2.0 and NC EOG.

For Grade 6, there was a significant positive correlation between the i-Ready Diagnostic and NC Check-In #3, $r = .740$, $n = 204$, $p = < .001$; between the i-Ready Diagnostic and the NC EOG, $r = .784$, $n = 208$, $p = < .001$; and between the NC Check-In #3 and the NC EOG $r = .804$, $n = 207$, $p = < .001$. All correlations were strong with the strongest between the NC Check-In #3 and NC EOG. Table 28 presents the Pearson product-moment correlations among the three EOY assessments: EOY i-Ready Diagnostic, NC Check-Ins #3, and NC EOG for Grade 6.

Table 28

Grade 6 Correlations on End-of-Year Assessments

EOY Assessment	i-Ready Diagnostic	NC Check-Ins	NC EOG
i-Ready Diagnostic	-----	.740 *n* = 204	.784 *n* = 208
NC Check-Ins	.740 *n* = 204	-----	.804 *n* = 207
NC EOG	.784 *n* = 208	.804 *n* = 207	----

Note. EOY = end-of-year. NC = North Carolina. NC EOG = North Carolina End of Grade. $p <$ .001.

For Grade 7, there was a significant positive correlation between the i-Ready Diagnostic and NC Check-In 2.0 #3, $r = .752$, $n = 181$, $p = < .001$; between the i-Ready Diagnostic and the NC EOG, $r = .797$, $n = 181$, $p = < .001$; and between the NC Check-In 2.0 #3 and the NC EOG $r = .785$, $n = 184$, $p = < .001$. All correlations were strong with the strongest between the i-Ready Diagnostic and NC EOG. Table 29 presents the Pearson product-moment correlations among the three EOY assessments: EOY i-Ready Diagnostic, NC Check-Ins 2.0 #3, and NC EOG for Grade 7.

Table 29

Grade 7 Correlations on End-of-Year Assessments

EOY Assessment	i-Ready Diagnostic	NC Check-Ins 2.0	NC EOG
i-Ready Diagnostic	-----	.752 *n* = 181	.797 *n* = 181
NC Check-Ins 2.0	.752 *n* = 181	-----	.785 *n* = 184
NC EOG	.797 *n* = 181	.785 *n* = 184	----

Note. EOY = end-of-year. NC = North Carolina. NC EOG = North Carolina End of Grade. $p <$ .001.

For Grade 8, there was a significant positive correlation between the i-Ready Diagnostic and NC Check-In 2.0 #3, $r = .743$, $n = 189$, $p = < .001$; between the i-Ready Diagnostic and the NC EOG, $r = .768$, $n = 193$, $p = < .001$; and between the NC Check-In 2.0 #3 and the NC EOG $r = .770$, $n = 207$, $p = < .001$. All correlations were strong with the strongest between the NC Check-Ins 2.0 and NC EOG. Table 30 presents the Pearson product-moment correlations among the three EOY assessments: EOY i-Ready Diagnostic, NC Check-Ins 2.0 #3, and NC EOG for Grade 8.

Table 30

Grade 8 Correlations on End-of-Year Assessments

EOY Assessment	i-Ready Diagnostic	NC Check-Ins 2.0	NC EOG
i-Ready Diagnostic	-----	.743 *n* = 189	.768 *n* = 193
NC Check-Ins 2.0	.743 *n* = 189	-----	.770 *n* = 207
NC EOG	.768 *n* = 193	.770 *n* = 207	----

Note. EOY = end-of-year. NC = North Carolina. NC EOG = North Carolina End of Grade. $p < .001$.

Summary of Data Analysis of Evaluation Question 2

In summary, student reading proficiency increased during the 2022-23 school year as measured by the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOG assessments. Appendix G contains descriptive statistics for all 2022-23 assessments. The means of the i-Ready Diagnostic increased 17 points for Grade 6, 15 points for Grade 7, and 13 points for Grade 8 from the beginning to the end of the school year. Growth of means on the NC Check-Ins/NC Check-Ins 2.0 was smaller, but still present, for Grade 6 at 0.1 and Grade 7 at 0.4; but the mean of Grade 8 remained the same. Achievement on the NC EOG followed a similar pattern to the i-Ready Diagnostics with Grade 6 increasing 10 points and Grade 7 increasing 0.4, but Grade 8 decreasing by -1.9 points. Based on calculating a Pearson product-moment correlation coefficient, there was a strong positive association among results on the EOY i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0 #3, and NC EOG.

Evaluation Question 3: What were the perceptions of teachers, instructional staff, and administrators regarding i-Ready in terms of guiding program elements such as the

following: (a) the value of student data generated by the i-Ready Diagnostic, (b) the value of the i-Ready Instruction as an instructional resource, and (c) the extent to which i-Ready is having an impact on student achievement?

Following the 2022-23 school year, all middle school English language arts teachers (n = 8), instructional staff (n = 2), and administrators (n = 3) participated in semi-structured interviews to share their perceptions of the i-Ready program and implementation. All teachers also participated in an online survey, in which they reported their classroom usage and perceptions of key aspects of the i-Ready program. Data from the semi-structured interviews and online survey are organized by evaluation question and are reported in two groups: teachers and administrators, which includes instructional staff.

i-Ready Diagnostic

One theme that emerged from teachers and administrators was that i-Ready Diagnostic clearly identified students' strengths and gaps in reading. When interviewed about the most valuable aspects of the i-Ready Diagnostic in reading, all staff (n = 13) shared that the i-Ready Diagnostic identified students' current level of performance on specific reading domains. Specifically, three staff said that i-Ready Diagnostic "pinpoints" those strengths and weaknesses. Three staff commented on the benefits of knowing students' level when students entered their classrooms, but one teacher qualified this benefit for students who tried their best on the assessment. One administrator described the adaptive nature of the i-Ready Diagnostic as valuable, and another administrator described it as a "consistent" screener for kindergarten through grade eight that was directly correlated to the NC EOG.

Another theme that emerged from six teachers and one administrator was that i-Ready Diagnostic set and monitored specific goals for students. Two staff specifically mentioned

focusing on students' attaining their stretch goals. When surveyed about using the Diagnostic Growth Report to create and discuss typical and stretch growth goals with students, one teacher reported this occurring *daily*, one teacher reported it *weekly*, one teacher reported it *monthly*, and six teachers reported it *quarterly*.

Throughout the interviews, another theme emerged that i-Ready Diagnostic data were more meaningful when stakeholders were motivated. Ten staff referred to students needing to try their best and take their time. Curriculum Associates (2023c) suggests on i-Ready Central (i.e., an online resource with checklists and action plans to help staff using i-Ready) that to get good data, teachers should get organized, prepare and motivate students, actively proctor, and track completion. When surveyed about preparing students to take the i-Ready Diagnostic by using the preparation materials found in i-Ready Central, one teacher reported *always*, two teachers reported *often*, four teachers reported *sometimes*, and one teacher reported *never*. Twelve staff provided at least one specific strategy to motivate students on i-Ready Diagnostic. Table 31 lists the strategies teachers and administrators suggested for motivating students on i-Ready Diagnostic.

Table 31

Strategies for Motivating Students on i-Ready Diagnostic

Strategy	Additional Notes	No. Staff Reporting
Rewards	Additional recess time, candy, eating outside passes, fast food gift cards, "Free Fridays," homework passes, parties with ice cream or cake, and stickers	8
Individual recognition	For most lessons passed, most usage minutes, highest growth, and/or highest score	7
Encouragement and celebration	Teachers encouraging students, and administrators encouraging teachers (includes teacher and administrator buy-in)	6
Individual student data conferencing (or data chats)	Includes using i-Ready Diagnostic Results Report and Diagnostic Growth Report and creating goals with students	5
Class or grade level competition	For most lessons passed, most usage minutes, highest growth, and/or highest score	3
Chucking the Diagnostic	Divide test administration over multiple sessions, so students do not test more than 30 minutes per day.	1
Resetting the Diagnostic	Restart test for students who have a red flag or yellow flag for rushing or who indicate in individual data conferences that they did not do their best.	1

Note. Staff (n = 12) could suggest more than one strategy.

Two administrators discussed the importance of staff buy-in to increase student motivation on the i-Ready Diagnostic. Individual student data conferencing was one strategy suggested by five staff members. When surveyed about leading data chats using i-Ready Diagnostic Results Reports and Diagnostic Growth Report, one teacher reported conducting them *weekly*, two teachers reported conducting them *monthly*, and six teachers reported conducting them *quarterly*. Two other strategies suggested by staff members were chunking and resetting the i-Ready Diagnostic. When surveyed about chunking the i-Ready Diagnostic over

multiple days, so students did not test for more than 30 minutes per day, two teachers reported *always*, one teacher reported *often*, and six teachers reported *sometimes* chunking. When surveyed about resetting the i-Ready Diagnostic for individual students whose data on the Diagnostic Status Report indicated a red flag or yellow flag for rushing, two teachers reported *always*, four teachers reported *often*, and two teachers reported *sometimes* resetting. Eleven staff members discussed i-Ready results informing lesson planning and pacing; however, two teachers shared they had i-Ready results, but they did not use them to adjust their pacing or lesson planning. When surveyed about how often they used i-Ready's Historical Reports to see the class's testing history, four teachers reported *weekly*, three teachers reported *monthly*, and one teacher reported *yearly*. When surveyed about how often they used i-Ready's Historical Reports to see an individual student's testing history, three teachers reported *weekly*, four teachers reported *monthly*, and one teacher reported *yearly*.

A final theme emerging from nine staff about the i-Ready Diagnostic was that the i-Ready Diagnostic created student groups targeted on specific needs. One teacher reported having awareness of the student groups and using that feature for intervention and enrichment time, but intentionally choosing not to use that feature in the regular classroom setting because of her belief that middle school students needed choices about with whom they work in small groups. When surveyed about how often they used i-Ready's Diagnostic Results Report to create student groups to intervene on specific skill deficits, three teachers reported *never*, one teacher reported *weekly*, three teachers reported *monthly*, and one teacher reported *quarterly*.

Based on interviews and survey responses, teachers and administrators appeared to have a good understanding of key features of the i-Ready Diagnostic, but there was the potential for continuous growth in implementing those key features in their classroom practices. Two areas of

high potential growth are in preparing and motivating students for the i-Ready Diagnostic using resources on i-Ready Central and in targeted instruction for student groups based on specific needs. In general, when surveyed about the value of the student data generated from the i-Ready Diagnostic assessment, five teachers (62.5%) responded that it was *very valuable*, two teachers (25%) responded that it was *somewhat valuable*, and one teacher (12.5%) responded that it was *not valuable*.

i-Ready Instruction

One theme that emerged from teachers and administrators was that i-Ready Instruction delivered personalized lessons to remediate specific standards and skills for individual students. Eleven staff specifically described i-Ready Instruction as "personalized," "tailored to each student," "differentiated," or creating for students their own "path" for "remediation" of "standards" and "skills." While one teacher qualified this benefit for students who "actually take it seriously like they should," two teachers elaborated that i-Ready Instruction personalized for every student in a way that one teacher would not be able to do. Three teachers referred to i-Ready Instruction as "interactive," providing "breaks along the way," and "engaging" for students. Four staff discussed how i-Ready lesson passage rates informed teachers' next steps when students were unsuccessful online and needed additional remediation directly facilitated by a teacher. When surveyed about the value of the student instruction and practice in the i-Ready Instruction, six teachers (75%) responded that it was *very valuable*, one teacher (12.5%) responded that it was *somewhat valuable*, and one teacher (12.5%) responded that it was *not valuable*.

In addition to the personalized online lessons, i-Ready Instruction also included access to Teacher Toolbox, which is a digital collection of instructional resources and assessments aligned

to standards that teachers can use with individuals, small groups, or whole classrooms of students. Another theme that emerged from teachers and administrators was teachers' knowledge and usage of the Teacher Toolbox and its tools were inconsistent. Two teachers shared they frequently used Teacher Toolbox and found it very beneficial, but five teachers reported either they were unsure what it was, or they found it hard to navigate. When surveyed about how often they used Teacher Toolbox to provide specific interventions that were personalized to the needs of small groups or individual students, one teacher responded *weekly*, one teacher responded *monthly*, three teachers responded *quarterly*, one teacher responded *yearly*, and two teachers responded *never*. When surveyed about how often they used specific resources in Teacher Toolbox, four teachers (50%) responded *never* using the Grade Level Scaffolding Report to plan small groups and/or use reading buddies; five teachers (62.5%) responded *never* using Ready-NC resources with students to build conceptual understanding through reasoning, practice, and productive discussion around real-world scenarios; six teachers (75%) responded *never* using PHONICS for Reading to help struggling students become fluent and independent readers; seven teachers (87.5%) responded *never* using Think Up! – ELA resources with students to integrate critical thinking instruction on reading standards; and eight teachers (100%) responded *never* using Ready Reading Discourse Cards with students to promote richer, deeper dialogue to engage all students in meaningful reading conversations. When interviewed, one teacher said, "I love the Teacher Toolbox in reading" and explained its lessons matched standards, its lessons were easy to deliver to individual students or small groups, its assessments were useful, and it was helpful for reteaching. Another teacher described using Teacher Toolbox for specific lessons in intervention and enrichment block as well as when the whole class needed reteaching on a specific topic. One administrator noted that it was a "wonderful tool" but was unsure whether

staff would use it. Another administrator noted the Teacher Toolbox had lessons and interventions with "everything in one place." When surveyed about the value of the instructional resources generated from the i-Ready Teacher Toolbox, two teachers (25%) responded that it was *very valuable*, and six teachers (75%) responded that it was *somewhat valuable*.

Another theme that emerged from teachers and administrators was the importance of setting and enforcing a culture of high expectations regarding i-Ready usage to overcome challenges of implementing i-Ready as intended. Research showed that students who used i-Ready as intended demonstrated improved reading achievement (Cook & Ross, 2022; Curriculum Associates, 2020, 2021c, 2022a, 2022b; M. Swain et al., 2020). Teachers need to use these effective programs even though there are implementation challenges. Nine staff indicated that meeting the program expectations of 45 minutes of student usage per week was a challenge. Staff elaborated about the insufficient length of classes needed for the large amount of reading content other than i-Ready, frequent student absences and tardies, students who were bored or "burned out" on i-Ready, and students who tested out of i-Ready. Because leaders build the culture in the buildings and the district, administrators must establish using i-Ready is both important and expected for teachers and students. To try to achieve fidelity of i-Ready Instruction implementation, administrators shared that they "asked teachers to monitor it by allotting specific time and not just 'work on i-Ready if you finish.'" Interview responses also indicated the importance of administrative monitoring of student usage rates, downloading and discussing the usage reports, and discussing student usage rates with teachers in professional learning community meetings to let teachers know that "administration is aware if the program is being utilized or not." One spoke about the importance of administrators communicating their belief about i-Ready's effectiveness and expectation of usage fidelity to staff and said,

"Communicating that with the staff and letting them know that as an administrator, we feel it's important that they use the tool, and whether we believe it or not, we must be positive about it." Another administrator felt initially that the teacher time commitment was "a lot" but realized later there was "plenty" of time to fit in the 45 minutes per week; furthermore, in Year 1, "we didn't try very hard," but in Year 2, "we committed." An administrator also noted that fidelity of usage came from actual teacher monitoring in various ways (e.g., usage of GoGuardian [software that helps teacher manage student devices to make sure students were not opening other tabs instead of i-Ready on their Internet browsers] and walking around the room).

i-Ready's Impact on Student Achievement

All staff (n = 13) indicated that i-Ready positively affected student growth. One administrator said, "i-Ready has been kind of an anchor...a place for [staff] to go back to, which has been helpful." Ten staff specifically indicated that i-Ready helped close gaps, and 11 staff noted that it helped individual students. Three staff emphasized improved reading ability. Eight staff credited i-Ready for positively affecting reading proficiency, but three staff shared that they could not attribute improved reading proficiency only to i-Ready. One administrator said that they could not attribute causation completely to i-Ready but said, "All our scores went up and our growth was up." Additionally, four staff noted increases in students' confidence that they attributed to i-Ready, and one staff member shared that i-Ready puts more accountability on students.

All administrators mentioned that students showed growth when the i-Ready program was implemented with fidelity, and two teachers indicated that the i-Ready professional development assisted their implementation fidelity. Various staff expressed concerns or raised questions about student achievement. Three staff expressed concerns about advanced readers and

gifted students, especially those who tested out of i-Ready Instruction. One teacher expressed that i-Ready was a "better tool for struggling readers" than advanced ones, and another hypothesized that i-Ready was more effective for elementary school students than middle school students. When surveyed about the overall impact that i-Ready had on student achievement in reading, five teachers (62.5%) responded that it was *very valuable*, two teachers (25%) responded that it was *somewhat valuable*, and one teacher (12.5%) responded that it was *not valuable*.

Summary of Data Analysis of Evaluation Question 3

In summary, during the interviews and surveys about i-Ready Diagnostic and i-Ready Instruction, themes emerged and included that i-Ready Diagnostic pinpointed students' strengths and gaps in learning, which i-Ready Instruction addressed with personalized lessons for remediation of those standards and skills; set and monitored specific goals for students; created student groups targeted on specific needs; and had more meaningful data when stakeholders were motivated. Other themes about i-Ready Instruction included teachers' knowledge and usage of the Teacher Toolbox and its tools were inconsistent and the importance of setting and enforcing a culture of high expectations regarding i-Ready usage to overcome challenges of implementing i-Ready as intended. A final theme that emerged was that all staff agreed that i-Ready had positively affected student growth. These themes illustrated key concepts found in the literature review including coherent and adaptive assessment systems, effective instructional interventions, fidelity of implementation, and adolescent literacy. While no staff used the word "coherence," all interviews included descriptions of the relationship between i-Ready Diagnostic and i-Ready Instruction showing the curriculum, instruction, and assessment effectively working together toward the same valuable, compelling, and reachable goals and students moving closer

to shared definitions of what they ought to know and be able to do (Shepard et al., 2017). One staff member described the adaptive nature of the i-Ready Diagnostic, and others alluded to benefits of adaptive assessments including maximizing information on student performance (including performance descriptions on multiple grade levels), promoting accurate measurement of a growth across a student's school career, and helping administrators make long-term decisions and measure impact (Curriculum Associates, n.d.). No staff used the phrase "adolescent aliteracy" (Verhoeven & Snow, 2001), but some staff questioned some student performance levels because they noted characteristics of students being bored, pretending to do i-Ready Instruction but doing something else, and feeling "burned out" or "sick of" i-Ready. Administrators noted the need to motivate and engage teachers, and teachers noted the need to motivate and encourage students to ensure students experienced positive effects from i-Ready. When surveyed about whether the middle school should continue, modify, or discontinue i-Ready use, five teachers (62.5%) responded that it should *continue*, two teachers (25%) responded that it should *be modified* (i.e., having another resource for students who test out i-Ready and having a separate time during the school day set aside for i-Ready without taking English class time for usage), and one teacher (12.5%) responded that it should *be discontinued*. When all staff were asked about their ideas for modifying i-Ready, responses were widespread. Some responses (e.g., increased student ownership/choice, shorter lessons/reading passages, including grammar, including performance streaks, and broadening grade-level instruction of reading domains) involved changes to the program itself and were outside of the realm of school district's control, and some responses (e.g., creative scheduling of i-Ready time or having all

professional development prior to beginning the program) were about the school's implementation.

Summary of Findings

In summary, there was fidelity of implementation in administration of the i-Ready Diagnostic as over 95% of all enrolled students participated in the i-Ready Diagnostic, and there was a lack of fidelity in implementation of i-Ready Instruction as expected usage was not met during any month for an entire grade level and almost 40% of students did not pass 70-100% of lessons. Student reading proficiency increased during the 2022-23 school year as measured by the i-Ready Diagnostic (M = +13-17 points), NC Check-Ins/NC Check-Ins 2.0 (M = +0-0.4 points), and NC End of Grade assessments (M = +0.4 – 10 points [except for Grade 8, which decreased 1.9 points]). Based on calculating a Pearson product-moment correlation coefficient, there was a strong positive association among results on the EOY i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0 #3, and NC EOG. Stakeholder perceptions included that the i-Ready Diagnostic and i-Ready Instruction were valuable because i-Ready Diagnostic pinpointed students' strengths and gaps in learning, which i-Ready Instruction addressed with personalized lessons for remediation of those standards and skills, and all teachers and administrators agreed that i-Ready has positively affected student achievement by increasing student growth. When surveyed about whether the middle school should continue, modify, or discontinue i-Ready use, five teachers (62.5%) responded that it should *continue*, two teachers (25%) responded that it should *be modified*, and one teacher (12.5%) responded that it should *be discontinued*.

CHAPTER 5

RECOMMENDATIONS

Aurelius School District used Curriculum Associates' i-Ready program, which included i-Ready Diagnostic, i-Ready Instruction, and Teacher Toolbox, to improve Cole Middle School students' reading achievement in 2022-23. This program evaluation sought to determine the fidelity of usage of diagnostic assessments, personalized instruction, and teacher resources in the implementation of the i-Ready program; assess the effectiveness of the i-Ready program in increasing student achievement as measured by i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOG; and determine stakeholders' perceptions of the value of the program to inform recommendations to continue, modify, or discontinue usage of the i-Ready program. Because it provided staff and students with timely, understandable, and helpful instruction and data that helped students grow, recover unfinished learning, and improve their achievement, the i-Ready program was valuable and worth its cost, though there are some recommendations that, if implemented, could increase the benefits for students. Over 95% of all enrolled students participated in the i-Ready Diagnostic, which showed fidelity of implementation in administration of the i-Ready Diagnostic. Expected usage of i-Ready Instruction was not met during any month for an entire grade level and almost 40% of students did not pass 70-100% of lessons, which together demonstrated a lack of fidelity in implementation of i-Ready Instruction. Even with usage time that did not meet expectations and insufficient lesson passage rates, as measured by the i-Ready Diagnostic (M = +13-17 points), NC Check-Ins/NC Check-Ins 2.0 (M = +0-0.4 points), and NC End of Grade assessments (M = +0.4 – 10 points [except for Grade 8,

which decreased 1.9 points]), student reading proficiency increased during the 2022-23 school year. Based on calculating a Pearson product-moment correlation coefficient, there was a strong positive association among results on the EOY i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0 #3, and NC EOG. Perceptions of both teachers and administrators included that the i-Ready Diagnostic and i-Ready Instruction were valuable because i-Ready Diagnostic pinpointed students' strengths and gaps in learning, which i-Ready Instruction addressed with personalized lessons for remediation of those standards and skills, and all stakeholders agreed that i-Ready positively impacted student achievement by increasing student growth. When surveyed about whether the middle school should continue, modify, or discontinue i-Ready use, five teachers (62.5%) responded that it should *continue*, two teachers (25%) responded that it should *be modified*, and one teacher (12.5%) responded that it should *be discontinued*. The previous chapter was structured by evaluation question and included data analysis and findings for each question. This chapter is also structured by evaluation question and synthesizes the findings and relevant literature to provide recommendations.

Summary of Major Findings

The purpose of this program evaluation was to analyze fidelity of implementation, impacts on student reading achievement, and perceptions of stakeholders to inform recommendations to continue, modify, or discontinue usage of the i-Ready program. Major findings are structured by evaluation question. Overall, data supported continuing the i-Ready program in Cole Middle School.

Evaluation Question 1: To what extent did teachers in the selected school implement with fidelity the i-Ready program (i.e., i-Ready Diagnostic and i-Ready Instruction) in the 2022-23 academic year?

Fidelity of Implementation. A significant predictor of student reading outcomes is fidelity of implementation (O'Donnell, 2008), which means a program is delivered the way it was designed and in the way it was implemented during the research studies that validated its effectiveness (National Center on Intensive Intervention, 2013). Effective interventions implemented with higher degrees of fidelity tend to be more effective, and schools need a systematic process to regularly evaluate implementation (Durlak & DuPre, 2008; Fixsen et al., 2013; Quinn & Kim, 2017). In considering fidelity of implementation, five criteria to measure are adherence, duration, quality of delivery, participant responsiveness, and program differentiation (O'Donnell, 2008).

Fidelity of i-Ready Diagnostic: A Coherent and Adaptive Assessment System. During the 2022-23 academic year, teachers implemented the i-Ready program with varying degrees of fidelity. Overall, there was fidelity of implementation in administration of the i-Ready Diagnostic, which means that all enrolled students completed all three assessments. Over 95% of all enrolled middle school students participated in the i-Ready Diagnostic assessments, which established a personalized instructional pathway for each student in i-Ready Instruction. Results indicated that 98% or more of enrolled Grade 6 and Grade 7 students and 90% or more of enrolled Grade 8 students completed i-Ready Diagnostic assessments in 2022-23. While only two of the eight teachers administered all three i-Ready Diagnostic assessments to all their enrolled students, most of the missing assessments were in two classrooms. For example, one of

those teachers did not administer 26 of the 281 i-Ready Diagnostic assessments, which is almost 10%, expected to be administered to the 111 different students on their roster.

The i-Ready Diagnostic demonstrates characteristics found in the literature on coherent and adaptive assessment systems. The assessment sought to determine how well students were learning, and it could be used to assist learning, measure individual achievement, and evaluate programs (Pellegrino et al., 2001). As a formative assessment, it provided feedback to inform on skill progress for small groups, targeted instruction, and scaffolding (McGlynn & Kelly, 2017) and provided accurate reflections of achievement (Stiggins & Chappuis, 2005). By its results informing i-Ready Instruction, it drove systematic instruction (Reutzel, 2022), tracked student growth continuously and consistently (Curriculum Associates, n.d.), and provided students with continuous access to descriptive feedback on improvement in their work (Black & Wiliam, 1998; Bloom, 1984; Pellegrino et al., 2001; Rodriguez, 2004; Stiggins & Chappuis, 2005). It charted progress over time (Pellegrino et al., 2001) and created an in-depth, personalized assessment of each student. Lastly, it was a computer-adaptive assessment, with different questions personalized for each student (CSAI, 2019) using the Rasch Model (1980).

Fidelity of i-Ready Instruction: An Effective Instructional Intervention. Overall, fidelity in implementation of i-Ready Instruction was not achieved as measured by program expectations, which were all enrolled students should have used i-Ready Instruction for 45 minutes per week and passed 70% or more of their lessons. Meeting or exceeding program expectations for usage only occurred in four classrooms (two in Grade 6 and two in Grade 7) for one of the 9 months (either February or March 2023) and was not met during any month for an entire grade level. The lowest usage months for teachers were either one of the first usage months (September or October 2022) or one of the last months (April or May 2023). Similarly,

33% of students engaged with i-Ready Instruction but passed less than 70% of their lessons. The highest percentage of lessons passed was in Grade 7 with 72%, and Grades 8 and 6 had 63% and 62% respectively.

The i-Ready Instruction demonstrates characteristics found in the literature on effective instructional interventions. As a blended learning intervention, it combined student completion of online activities using digital technology (i-Ready Instruction) with teacher-led instruction (Teacher Toolbox) to help non-proficient students master skills (Horn & Staker, 2011; Macaruso et al., 2020; Pytash & O'Byrne, 2018). It was research-based, systematically delivered, based on targeted areas of need (from i-Ready Diagnostic results), and used a proactive and catch-up stance to enable acceleration to become on-grade level or above (Neal & Kelly, 2002). It was scaffolded sequentially to promote mastery and build confidence (Hirsh-Pasek et al., 2015). Employing mastery learning, it used instruction, formative assessment, feedback, and either individual corrective activities and second chance formative assessments or enrichment and extension activities (Bloom, 1971). It provided real time performance data to teachers to provide individualized instruction (Macaruso et al., 2020). For students who demonstrated significantly low achievement, i-Ready Instruction provided individualized support (National Center on Intensive Intervention, 2013), and Teacher Toolbox provided lessons for teachers to deliver individually.

Evaluation Question 2: To what extent, if any, did student reading proficiency improve during the 2022-23 academic year as measured by the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOGs?

Student reading proficiency increased during the 2022-23 school year as measured by the i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOG. The means of the i-Ready

Diagnostic increased 17 points for Grade 6, 15 points for Grade 7, and 13 points for Grade 8 from the beginning to the end of the school year. Growth of means on the NC Check-Ins/NC Check-Ins 2.0 was smaller, but still present, for Grade 6 at 0.1 and Grade 7 at 0.4; but the mean of Grade 8 remained the same. Achievement on the NC EOG followed a similar pattern to the i-Ready Diagnostics as Grade 6 increased 10 points and Grade 7 increased 0.4, but Grade 8 decreased by -1.9 points. Based on calculating a Pearson product-moment correlation coefficient, there was a strong positive association among results on the EOY i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0 #3, and NC EOG.

Adolescent Literacy. Examining changes in student reading proficiency after participation in the i-Ready program relates to literature on adolescent literacy, which is a complex concept that includes comprehending different kinds of texts, mastering new vocabulary, and sharing ideas with others (Kamil et al., 2008). According to the 2022 NAEP, only 31% of eighth grade students are reading proficiently and able to comprehend text at their grade level (NCES, 2022). While sixth and seventh grades increased student reading proficiency on the summative NC EOG, eighth grade dropped back to only 38.7% proficient, which was the same level as the 2020-21 school year. As the number of non-proficient readers increases, more students miss out on more background knowledge, skills, and specialized vocabulary, which are important for understanding the schools' core curriculum (Alvermann, 2005).

Adolescent Aliteracy. Examining changes in student reading proficiency after participation in the i-Ready program also relates to literature on adolescent aliteracy, which is the condition of being able to read but choosing not to do so (Verhoeven & Snow, 2001). As students enter middle school, adolescent aliteracy becomes more prevalent. Teachers must focus on motivating and engaging students (Dredger, 2013) to do their best on all coursework and

assessments. Based on their students' performance, this concept seems even more important for the eighth-grade teachers.

Evaluation Question 3: What were the perceptions of teachers, instructional staff, and administrators regarding i-Ready in terms of guiding program elements such as the following: (a) the value of student data generated by the i-Ready Diagnostic, (b) the value of the i-Ready Instruction as an instructional resource, and (c) the extent to which i-Ready is having an impact on student achievement?

Themes emerged during the interviews and surveys about i-Ready Diagnostic, i-Ready Instruction, and Teacher Toolbox. Overall, staff perceptions were positive about i-Ready Diagnostic. While not as strongly positive, staff expressed contentment with i-Ready Instruction. All staff gave credit to i-Ready for increased student growth in reading. Perceptions of Teacher Toolbox revealed many teachers were unfamiliar with its resources and were not using it to its potential with students.

Perceptions of i-Ready Diagnostic. Themes about the i-Ready Diagnostic included characteristics of coherent and adaptive assessment systems. The i-Ready Diagnostic pinpointed students' strengths and gaps in learning. It set and monitored specific goals for students. It created student groups targeted on specific needs. It had more meaningful data when stakeholders are motivated. Pellegrino et al. (2001) note that adaptivity in education assessment results in charting progress over time and modeling performance at various levels, which the i-Ready Diagnostic did.

Perceptions of i-Ready Instruction. Themes about i-Ready Instruction included characteristics of effective instructional interventions. The i-Ready Instruction addressed gaps identified on i-Ready Diagnostic with personalized lessons for remediation of those standards

and skills. Staff noted it was important to set and enforce a culture of high expectations regarding i-Ready Instruction usage to overcome challenges of implementing i-Ready as intended. Also, a theme that emerged about Teacher Toolbox was that teachers' knowledge and usage of the Teacher Toolbox and its tools are inconsistent. Neal and Kelly (2002) point out that effective interventions are provided by trained professionals, and based on their responses, training about Teacher Toolbox and its resources would be beneficial.

Perceptions of i-Ready's Impact on Student Achievement. One theme that emerged about the i-Ready program was that all staff agree that i-Ready has positively impacted student growth. This finding resonates with other literature showing i-Ready's positive effect on student achievement including Curriculum Associates (2020, 2021c, 2022a, 2022b) and independent research (Cook & Ross, 2022; M. Swain et al., 2020). When surveyed about whether the middle school should continue, modify, or discontinue i-Ready use, five teachers (62.5%) responded that it should *continue*, two teachers (25%) responded that it should *be modified* (e.g., having another resource for students who test out i-Ready and having a separate time during the school day set aside for i-Ready without taking English class time for usage), and one teacher (12.5%) responded that it should *be discontinued*. When all staff were asked about their ideas for modifying i-Ready, responses were widespread. Some responses (e.g., increased student ownership/choice, shorter lessons/reading passages, including grammar, including performance streaks, and broadening grade-level instruction of reading domains) involved changes to the program itself and outside of the realm of school district's control, and some responses (e.g., creative scheduling of i-Ready time or having all professional development prior to beginning the program) were about the school's implementation.

Perceptions of Staff: Self-Efficacy, Motivation, and Engagement. Staff perceptions on i-Ready's impact on student achievement demonstrated characteristics of the literature on self-efficacy in teaching non-proficient adolescent readers, motivation, and engagement. At the secondary level, the responsibility for teaching reading skills often seems to belong to no one specifically, and some teachers reported feeling unprepared to help students improve their reading proficiency (Heller & Greenleaf, 2007), so i-Ready has provided the middle school with a way to reach students who were reading significantly below grade level. Even when teachers have a sense of self-efficacy, which Bandura (1997) describes as the "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p. 2), about teaching in their specific content area, they may have a low sense of efficacy regarding teaching reading to students they perceive as struggling or unmotivated (Fine et al., 2011). Motivation, or the desire, reason, or predisposition to become involved in a task or activity, and engagement, or the degree to which a student processes text deeply using active strategies, thought processes, and prior knowledge, play a key role in adolescents' academic literacy development (Alvermann, 2005; Guthrie & Wigfield, 2000; Kamil et al., 2008). Some staff interviews referred to the gamification built into i-Ready as a motivator for students. More staff interviews described ways that teachers actively motivated and encouraged students including individual data conferences, individual recognitions, class competitions, rewards, and celebration. These strategies aligned with the four recommendations to increase student motivation and engagement in literacy learning provided by Kamil et al. (2008):

1. Establish meaningful and engaging content learning goals around the essential ideas of a discipline as well as around the specific learning processes used to access those ideas.

2. Provide a positive learning environment that promotes student autonomy in learning.
3. Make literacy experiences more relevant to student interests, everyday life, or important current events.
4. Build classroom conditions to promote higher reading engagement and conceptual learning through such strategies as goal setting, self-directed learning, and collaborative learning.

Discussion of Findings

These findings illustrated key concepts found in the literature review including coherent and adaptive assessment systems, effective instructional interventions, fidelity of implementation, and adolescent literacy. While an important part of the literature review was the philosophies of teaching reading, this concept was less evident in the findings in the middle school context than it probably would have been in an elementary school context. Findings are organized by concept.

Coherent and Adaptive Assessment Systems

While no staff used the word "coherency," all interviews included descriptions of the relationship between i-Ready Diagnostic and i-Ready Instruction; of curriculum, instruction, and assessment effectively working together toward the same valuable, compelling, and reachable goals; and of students moving closer to shared definitions of what they ought to know and be able to do (Shepard et al., 2017). One staff member described the adaptive nature of the i-Ready Diagnostic, and others alluded to benefits of adaptive assessments including maximizing information on student performance (including performance descriptions on multiple grade levels), promoting accurate measurement of a growth across a student's school career, and helping administrators make long-term decisions and measure impact (Curriculum Associates,

n.d.). As a formative assessment, i-Ready Diagnostic met Wiliam and Thompson's (2007) framework that (a) establishes where learners are in their learning, (b) establishes where they are going, and (c) establishes how to get there. Also, it provided student grouping for targeted instruction and scaffolding at students' levels (McGlynn & Kelly, 2017). These characteristics of i-Ready Diagnostic illustrate coherent and adaptive assessment systems.

Effective Instructional Interventions

The i-Ready program was found to be an effective instructional intervention. It met Neal and Kelly's (2002) characteristics of successful late intervention programs: (a) considering individual student needs, (b) implementing an apprenticeship model of teaching and learning, (c) selecting appropriate materials, (d) establishing a focus on accelerative instruction, (e) considering role of fluent responding, and (f) providing for affirmation of success. In interviews, teachers noted they felt i-Ready Instruction benefitted low achievers more than other students, which aligns with the findings of Black and Wiliam (1998) about student-involved formative assessment. It was interesting to see correlations between fidelity of implementation and student achievement. Two grade levels (sixth and seventh) came closer to achieving full fidelity of implementation, and they both experienced increased means and higher student reading achievement on all three measures. One grade level (eighth) only partially met fidelity of implementation for i-Ready Diagnostic and i-Ready Instruction and had decreased or unchanged means and lower student reading achievement on NC Check-Ins 2.0 and NC EOG. Ensuring students actively engage for 45 minutes per week in i-Ready Instruction is important as Guthrie and Wigfield (2000) found the level of student engagement is the mediating factor through which classroom instruction influences reading outcomes.

Fidelity of Implementation

In this evaluation of the implementation of the i-Ready program, O'Donnell's (2008) criteria were examined: adherence, duration, quality of delivery, participant responsiveness, and program differentiation. *Adherence* means the components were delivered as designed, so students should have participated in both the i-Ready Diagnostic and i-Ready Instruction. *Duration* means the number, length, or frequency of sessions implemented, so each student should have taken three i-Ready Diagnostics, engaged in 45 minutes of i-Ready Instruction per week, and passed at least 70% of lessons. *Quality of delivery* means teachers used the prescribed techniques, processes, or methods to deliver the program, so educators should have attended professional development, meaningfully used achievement and usage reports generated from i-Ready Diagnostic and i-Ready Instruction, and incorporated lessons from Teacher Toolbox. *Participant responsiveness* means the extent to which the participants were engaged by and involved in the activities and content of the program, so teachers and administrators should have shared their perspectives on i-Ready and staff should have used i-Ready reports to make data informed decisions (e.g., monitoring students' time on task during i-Ready Instruction and checking if students rushed or appeared to rush on their i-Ready Diagnostics or i-Ready Instruction lesson assessments and reassigning). *Program differentiation* means that critical features that distinguish the program were present during implementation, so teachers or grade levels should have used i-Ready Diagnostic and i-Ready Instruction with full fidelity, which should have resulted in student growth. All five of these criteria were met for i-Ready Diagnostic. For i-Ready Instruction, adherence (relating to fidelity to structure) and program differentiation (relating to fidelity to process) were met; however, duration (relating to fidelity to

structure), quality of delivery (relating to fidelity to process), and participant responsiveness (relating to fidelity to structure and process) were only partially met.

Adolescent Literacy

No staff used the phrase "adolescent aliteracy" (Verhoeven & Snow, 2001), but they questioned some student performance levels because they noted characteristics of students being bored, pretending to do i-Ready Instruction but doing something else, and feeling "burned out" or "sick of" i-Ready. Administrators noted the need to motivate and engage teachers, and teachers noted the need to motivate and encourage students to ensure students experienced positive effects from i-Ready.

While the i-Ready program was robust and exhibited many positive characteristics of the literature, vocabulary for middle school was one area that administrators felt was lacking. One recommendation to support non-proficient adolescent readers from the Institute of Education Science's National Center for Education Evaluation and Regional Assistance is to provide explicit vocabulary instruction (Kamil et al., 2008). In reviewing resources on the i-Ready Central website, there are new middle school vocabulary lessons designed to build students' depth of knowledge of academic vocabulary and support their comprehension of passages in future lessons on reading comprehension. Teachers would benefit from professional development on these vocabulary resources.

Implications for Policy and Practice

The findings of this program evaluation suggest that implementing the i-Ready program (including i-Ready Diagnostic and i-Ready Instruction with Teacher Toolbox) with fidelity helped to facilitate the intermediate term outcomes (i.e., increased student reading achievement on i-Ready Diagnostic, NC Check-Ins/NC Check-Ins 2.0, and NC EOG) for students at the

middle school that served as the context for this evaluation. It is recommended that the district continue to provide the i-Ready program for students in middle school to increase their reading proficiency. Recommendations for practice include that the program should continue in its current form, but slight improvements could be made to achieve stronger fidelity of implementation and meet long-term intended outcomes. This section aligns recommendations for policy and practice with the program evaluation findings and supporting literature. Table 32 provides an overview of the recommendations.

Table 32

Recommendations for i-Ready Program Continuation

Findings	Related Recommendations	Supporting Literature
Teachers administered i-Ready Diagnostic consistently. Over 95% of all enrolled students participated in the i-Ready Diagnostic.	Continue use of i-Ready Diagnostic assessments and ensure robust preparation of teachers and students.	Pellegrino et al. (2001)
Students used i-Ready Instruction inconsistently. Expected usage of i-Ready Instruction was not met during any month for a grade level; the 2022-23 average for Grade 6 was 62%, Grade 7 was 68%, and Grade 8 was 17%. Expected lesson passing level was not met as 33% of Grades 6-8 students did not pass at least 70% of i-Ready Instruction lessons.	Continue use of i-Ready Instruction and monitor student usage and lesson passing rates using the Personalized Instruction Summary report.	Curriculum Associates (2022b); Neal & Kelly (2002); O'Donnell (2008)
Grade 8 had fewer students take i-Ready Diagnostic (only 90% for the EOY) and never exceeded 31% of expected usage in any month. Achievement in Grade 8 decreased from BOY to EOY on NC Check-Ins 2.0 and NC EOG.	Create a successful student and teacher incentive system to increase student and teacher motivation and engagement and fidelity of implementation of the i-Ready program that includes • Hosting more inclusive celebrations that connect individual goals with group goals; • Implementing a choice board of different rewards and activities that students/teachers can select when they achieve their goal; • Connecting with families/ community to recognize achievements; and • Asking for feedback from students/teachers about the incentive system.	Harn et al. (2017); Ivey & Fisher (2006); Kamil et al. (2008); O'Donnell (2008)
5 of 8 teachers reported they were unsure what Teacher Toolbox was or they found it hard to navigate, and 6 teachers reported using Teacher Toolbox quarterly, yearly, or never, but 33% of students did not pass 70% of i-Ready lessons.	Implement professional development on Teacher Toolbox to increase use of Teacher Toolbox to provide targeted instruction to small groups of students based on specific needs	Curriculum Associates (2023a); Neal & Kelly (2002)

Note. BOY = beginning of year; EOY = end of year. NC = North Carolina. NC EOG = North Carolina End of Grade.

A Culture of High Expectations

All of the recommendations about i-Ready implementation have a foundation in district and school administrators creating a culture of high expectations regarding student achievement. The culture should prioritize student needs over staff comfort, normalize staff sharing and using data to inform instruction, and monitor progress and achievement. District administrators must set and enforce expectations for school administrators, who must set and enforce expectations for teachers, who must set and enforce expectations for students. High expectations require buy-in at all levels from leaders to students, and all involved stakeholders should understand and work toward the unified goal of improving student achievement. Once the goal is established and actions are delineated, everyone must work together to complete the actions and meet the goal. Administrators must monitor action steps and goal progress because what is inspected gets respected, and they model how teachers should do this for students. In a culture of high expectations, all involved share the same values, live the same mission, and work toward making the same goal a reality. Research has shown that i-Ready improves student achievement when it is implemented with fidelity (Cook & Ross, 2022; Curriculum Associates, 2020, 2021c, 2022a, 2022b; M. Swain et al., 2020), so part of living the culture of high expectations is implementing i-Ready with fidelity. The four recommendations share ways that the district can improve the implementation of i-Ready.

Recommendation 1: Continue Use of i-Ready Diagnostic Assessments and Ensure Robust Preparation of Teachers and Students

Administering the i-Ready Diagnostic with fidelity requires actions from administrators and teachers. As a coherent and adaptive assessment, the i-Ready Diagnostic determines how well students are learning, and its results can be used to assist learning, measure individual

achievement, and evaluate programs (Pellegrino et al., 2001). Recommendation #1 is to continue using i-Ready Diagnostic assessments and ensure robust preparation of teachers and students. It is also important to communicate with families about the i-Ready Diagnostic and share the purpose of the assessment and when students will take it.

Before facilitating i-Ready Diagnostic assessments, administrators need to set a clear vision of why the school is using i-Ready, establish priorities (including a testing schedule and expectations) for i-Ready Diagnostic administration, set goals for student growth and proficiency, and demonstrate excitement and build momentum for teachers. They should communicate and enforce expectations to teachers about preparing and motivating students to take the i-Ready Diagnostic assessments; administering assessments to all enrolled students; and using data from assessments to inform student groupings, lesson planning, and course pacing. District administrators such as the Executive Director of Instructional Services and Superintendent should meet with school administrators such as principals and assistant principals to ensure these action steps are achieved.

Teachers need to be organized and plan for time to prepare and motivate students to take the i-Ready Diagnostic as well as to administer it to students. They should communicate a clear vision to students about why the i-Ready Diagnostic is important, share expectations and tips for test-taking, and create individual and class goals. Resources including PowerPoint presentations and short videos are available on the i-Ready website in the i-Ready Central for teachers to use with students. Individual data conferences, or data chats, allow teachers to discuss students' previous test results, testing experiences, and goals for upcoming assessments. School administrators should meet with all teachers to ensure these action steps are achieved.

While facilitating the i-Ready Diagnostic, teachers must actively monitor students to observe for off-task behaviors (such as additional open tabs on internet browsers), rushing, or being stuck. During testing windows, district and school administrators and the instructional coach should conduct classroom walkthrough observations to ensure optimal testing conditions and consistent monitoring occur. Also, both teachers and administrators need to use the i-Ready Diagnostic Status Report to check student progress on the i-Ready Diagnostic and ensure that all students complete an assessment before their test expires and within the window on the testing schedule.

Recommendation 2: Continue Use of i-Ready Instruction and Monitor Student Usage and Lesson Passing Rates using the Personalized Instruction Summary Report.

Duration is the number, length, or frequency of sessions implemented (O'Donnell, 2008). Quality of delivery means teachers used the prescribed techniques, processes, or methods to deliver the program (O'Donnell, 2008). Fidelity of implementation of i-Ready Instruction as measured by duration and quality of delivery were only partially met, as usage data showed most students were not engaged in 45 minutes of i-Ready Instruction per week and over 30% did not pass at least 70% of lessons. Surveys and interviews indicated that teachers could improve how they accessed and used achievement and usage reports generated from i-Ready Diagnostic and i-Ready Instruction and incorporated lessons from Teacher Toolbox.

Students receive personalized interventions on i-Ready Instruction based on their needs identified in i-Ready Diagnostic. Several teachers spoke about how i-Ready Instruction can provide individualized differentiation to students in a way that teachers are unable to because of so many students on different levels. Effective interventions are research-based and systematically delivered in small groups or individually by a trained professional based on a

targeted area of need with a proactive, catch-up stance to enable acceleration to become on-grade level (Neal & Kelly, 2002). Research conducted by Curriculum Associates (2022b) showed that K-8 students who use i-Ready Instruction as recommended make greater improvements in reading (including achieving a higher grade-level placement by spring, exceeding their typical growth targets, and achieving a higher percentage of their stretch growth targets) than students who do not.

Recommendation #2 is to continue use of i-Ready Instruction and monitor student usage and lesson passing rates using the Personalized Instruction Summary report, classroom walkthrough observations, and discussions with teachers. Monitoring i-Ready Instruction is an important action in responding to students' needs. During weekly professional learning community meetings and monthly staff meetings, school administrators should model how to monitor expectations using the Diagnostic Status Report and Personalized Instruction Summary Report for classrooms, grade levels, and the school. Discussing these data should be a required focus of professional learning community meetings for teachers even if administrators are not in attendance. Administrators should review Lesson Time-on-Task for each classroom to meet the goal of 30-49 minutes of i-Ready Instruction per week to meet the goal average of 45 minutes per week. If classes are not meeting this expectation, administrators should review teachers' lesson plans to monitor whether teachers are scheduling 60 minutes of i-Ready Instruction time per week. Scheduling one hour per week accounts for transitions and necessary classroom interruptions. Administrators should also review the Percentage of Lessons Passed to meet the goal of lessons passed at or above 70%. Resources to document the data review are available in i-Ready Central. District administrators such as the Executive Director of Instructional Services

and Superintendent should monitor these reports and meet with school administrators to ensure usage expectations are met.

Recommendation 3: Create a Successful Student and Teacher Incentive System to Increase Student and Teacher Motivation and Engagement and Fidelity of Implementation of the i-Ready Program.

Participant responsiveness is the extent to which the participants were engaged by and involved in the activities and content of the program (O'Donnell, 2008). While improving student outcomes (not perfect fidelity) is the most important goal of an instructional intervention (Harn et al., 2017), these program evaluation findings indicated participant responsiveness to the i-Ready program was much lower in eighth grade than in sixth or seventh grades. Eighth grade had fewer students take i-Ready Diagnostic (i.e., only 90% for the EOY i-Ready Diagnostic) and never exceeded 31% of expected usage in any month. One might wonder if Cole Middle School should end i-Ready usage at seventh grade due to this limited usage in eighth grade; however, only 38% of eighth graders were proficient on the NC EOG and achievement in Grade 8 remained stagnant or decreased from BOY to EOY on NC Check-Ins 2.0 and NC EOG. Ivey and Fisher (2006) found that a whole-school approach is needed to improve literacy skills in secondary school students. They recommend curriculum materials that draw students into literary experiences that they find relevant and meaningful. Likewise, it is important to increase student motivation and engagement in literacy learning (Kamil et al., 2008). Perceptions of teachers and administrators indicated that teacher motivation for using i-Ready also needed to increase.

Teachers with the highest usage of i-Ready Diagnostic and i-Ready Instruction discussed rewards for growth and/or proficiency as a motivator for students; however, there was not a consistent, whole-school incentive system present. Recommendation #3 is to create a successful

student and teacher incentive system to increase student and teacher motivation and engagement and fidelity of implementation of the i-Ready program. To develop the system, school administration should empower the School Improvement Team to lead the effort. The student system should celebrate learning with clear criteria for success that connect individual goals with either classroom or grade-level goals, incorporate student choice, connect with families, and are relevant and meaningful for students. It should exist for all students in all classrooms, and the specific kinds of goals and rewards should align with program expectations but also include teacher and administrator feedback.

This program evaluation found that celebrations for the i-Ready Diagnostic mostly occurred for a small number of students (i.e., the top five in growth and/or proficiency in a classroom), so a recommendation is to create more inclusive celebrations that connect individual goals with either classroom or grade-level goals (e.g., an ice cream or pizza party for all students in the class or grade level when their class or grade meets the goals or for all students who meet their individual typical or stretch growth goals). Progress toward these goals can be documented visually in the hallways as a scoreboard to see when the students earn the larger celebration. Some teachers were celebrating all students who met their weekly goals for lessons passed and lesson time-on-task with a reward provided by the teacher, and a recommendation is to create a choice board of different rewards and activities that students can select when they achieve their goal. Examples of rewards and activities that are already used by individual teachers include stickers, candy, additional recess time, eat outside passes, homework passes, or $5 fast food gift cards. Connecting with families to recognize students' achievement is another way to increase student engagement, and i-Ready Central provides resources including postcards that teachers can use to acknowledge student achievement and growth and send to families. Recognizing

individual student achievement and growth on i-Ready Diagnostic and usage expectations in i-Ready Instruction during school awards ceremonies is another way for students to be celebrated by their peers and families. Relevance and meaning are important for adults, but they are also important for middle school students. Teachers and administrators should use a survey to get feedback from students about the incentive system and what celebrations would motivate them to do their best on i-Ready Diagnostic, meet expectations for usage time on i-Ready Instruction, and increase their lesson passing rates.

The teacher system should be similar to the student system. There should be clear criteria for success to connect individual goals with grade-level or school goals, incorporate teacher choice, connect with the community, and be relevant and meaningful for teachers. Teachers should be celebrated when all enrolled students finish i-Ready Diagnostic assessments within the testing window and when all enrolled students meet weekly usage of i-Ready Instruction or pass 70% or more of their lessons. As teachers meet goals, there should be a choice board for them with rewards that are relevant and meaningful to them based on their feedback (e.g., off-campus lunch pass, no bus duty pass, or preferred parking spot). Incentives should be more enticing for teachers during the months of least frequent usage in 2022-23 (i.e., September, October, April, and May). Teachers should be celebrated and receive recognition during staff meetings and on school social media for meeting their goals. This kind of motivation system serves to celebrate what administrators want to see more of while increasing fidelity of implementation, which should increase student achievement. District administrators can provide similar kinds of incentives to school administrators as school goals are achieved.

Recommendation 4: Implement Professional Development on Teacher Toolbox to Increase Use of Teacher Toolbox to Provide Targeted Instruction to Small Groups of Students Based on Specific Needs

Effective instructional interventions are research-based, systematically delivered in small groups or individually by a trained professional based on a targeted area of need, and foster students becoming on-grade level (Neal & Kelly, 2002). Teacher Toolbox provides resources aligned to standards and research-based interventions that teachers can use to target areas of students' needs, especially when students are not passing at least 70% of lessons (Curriculum Associates, 2023a). Teachers should use the i-Ready Diagnostic Results Report to create student groups to intervene on specific skill deficits; however, four teachers reported *never* or only *quarterly* using that report to intervene with students. Five teachers reported they were unsure what Teacher Toolbox was, or they found it difficult to navigate. Recommendation #4 is to implement professional development on Teacher Toolbox to increase use of Teacher Toolbox to provide targeted instruction to small groups of students based on specific needs. If funds allow, Curriculum Associates trainers could provide this training to all teachers; however, if funds do not allow for an external consultant, then the school's instructional coach and the two teachers who reported *frequently* using Teacher Toolbox and find it beneficial could provide the training. During this training, the following resources should be addressed: Grade Level Scaffolding Report, Ready-NC Resources, PHONICS for Reading, Think Up!-ELA Resources; Ready Reading Discourse Cards; and Middle School vocabulary. Administrators should monitor the classrooms and see to what extent the Teacher Toolbox is being used with students.

Additional Recommendations

Although this program evaluation examined fidelity of implementation, effects on student achievement, and perspectives of stakeholders on the i-Ready program in one middle school, ideas for improving the i-Ready program surfaced. Both teachers and administrators expressed concerns that students were testing out of the i-Ready program but still needed personalized instruction to assist with their growth. One additional recommendation is that Curriculum Associates increase its reading domains, so vocabulary and comprehension instruction continue into high school grade levels as well as include a grammar component, which would make the i-Ready program more robust for high achieving middle school students. Another recommendation is to increase the gamification components of i-Ready Instruction to include performance streaks for students, which could make it more motivating and engaging for students.

Recommendations for Future Research

The design of this program evaluation prevents generalizations about the i-Ready program in other schools in the district using i-Ready. Because other schools may have had different experiences with professional development or varying degrees of fidelity of implementation, the value of the i-Ready program could be perceived differently by stakeholders at other schools. Also, in this context, neither student nor parent voices were heard. The following recommendations for future research are provided:

1. Conduct surveys and focus groups with students and parents in the middle school and focus on their perspectives of the value of the i-Ready program.
2. Evaluate fidelity of implementation of the i-Ready program and student reading achievement at elementary schools within the district that are using i-Ready. Compare results with those found in this program evaluation.

3. After 5 years of implementation of the i-Ready program in the middle school, examine student reading achievement to see if the long-term program outcomes of eliminating COVID-19 learning loss and ensuring all students are reading on grade level were met.

Summary

ASD used Curriculum Associates' i-Ready program, which was philosophically aligned to the science of reading and included i-Ready Diagnostic, i-Ready Instruction, and Teacher Toolbox, to improve Cole Middle School students' reading achievement. This program evaluation determined the fidelity of implementation of the i-Ready program; assessed the effectiveness of the i-Ready program in increasing student achievement; and determined stakeholders' perceptions of the value of the program to inform recommendations to continue, modify, or discontinue usage of the i-Ready program. The program evaluation found that i-Ready was a coherent and adaptive assessment system and an effective instructional intervention designed to increase adolescent literacy, which occurred in grade levels where the program was implemented with fidelity.

As a district administrator, I face the constant juggle of competing demands for limited resources, and as the September 2024 ESSER funding cliff approaches, the struggle becomes even more challenging. Knowing how to conduct a program evaluation and then actually evaluating the programs the district uses are essential for district administrators, so we know what to continue, what to modify, and what to discontinue. This evaluation showed that the i-Ready program positively impacted student achievement when it was implemented with fidelity, but I am left with some lingering questions.

One question is about the BOY i-Ready Diagnostic: Is it a harder assessment than either the MOY or the EOY? Mean scores increased 17 points (Grade 6) and 13 points (Grade 7) from BOY to MOY, but then they only increased 0.37 (Grade 6) and 1.72 points (Grade 7) from MOY to EOY. Could lower BOY scores indicate something else (e.g., summer slide), or could they really be accurate reflections of student growth from the beginning to the middle of the year? If the latter is true, then we need to look at what was happening between BOY and MOY and what was not happening from MOY to EOY. Data from NC Check-Ins/NC Check-Ins 2.0 #2 were also the highest of the three benchmark assessments. How do we continue the BOY to MOY momentum during the MOY to EOY time frame?

Another question lingers around the idea of core English language arts curriculum in the middle school. Currently, teachers work together in professional learning communities to plan their curriculum around their standards, and they find resources to use using the school library, book room, and online instructional websites (e.g., CommonLit, Flocabulary, Newsela, ReadTheory, and ReadWorks). This approach gives teachers a tremendous amount of autonomy, but as the number of lateral entry and beginning teachers and the number of students entering at lower achievement levels increase, this approach becomes exponentially harder to ensure a viable and guaranteed curriculum. Students enrolled in the same class, course, or grade should engage in the same rigorous curriculum regardless of the teacher to whom they are assigned. One part of i-Ready that I have found beneficial is the vertical and horizontal coherence of instruction and robust tracking of student data across classrooms, grades, and years. Even though it is a supplemental program, i-Ready has provided a consistency that we have not had with our core. I have heard that we cannot intervene our way out of ineffective core instruction. I wonder what would happen if we invested in the adoption of a core curriculum for Grades 6–8. Would we find

gains greater than what we found with i-Ready? Would i-Ready be able to push our student achievement even higher if we coupled it with a core curriculum adoption?

Lastly, within the supplementary resource realm, I wonder what we need to increase growth for all students. Do we need to purchase two supplementary resources? Should we keep i-Ready for our struggling middle school students but use another program for our on-grade level and above grade level students? During the program evaluation, it was apparent that we needed to have teacher and student buy in for any program to work, but it was also apparent that we need to do something in Grade 8 to increase student achievement as well as consider what we should do for students who are testing out of i-Ready Instruction in all three grade levels. Since the i-Ready program ends at Grade 8, and both teachers and administrators indicated that i-Ready appears more elementary, would changing to another program that appears more secondary result in increased teacher and student buy in and translate to increase student achievement?

Although it would be difficult to find out the answers to some of these lingering questions, sometimes there is power is asking the questions and having the discussions. We might not know until we tried something different. Because i-Ready was successful at increasing student reading achievement when implemented with fidelity in grades six and seven classrooms, I believe that the immediate next step for all three grade levels is to continue implementing i-Ready but increase the fidelity of district and school administrator monitoring to increase the fidelity of implementing i-Ready. I also believe that this program as well as any others the district uses should continue to be evaluated for fidelity of implementation, impact on student achievement, and perceptions of stakeholders. These steps are important for district and school administrators in creating a culture of high expectations for student achievement and for meeting

one of our greatest educational responsibilities as educators, parents, and a collective society to ensure that every child learns to read well.

In conclusion, the i-Ready program positively impacted student achievement when it was implemented with fidelity. Because it provided staff and students with timely, understandable, and helpful instruction and data that helped students grow, recover unfinished learning, and improve their achievement, the i-Ready program was valuable and worth its cost. It is recommended that the i-Ready program continue in its current form with a few adjustments to motivate and encourage teachers and students to result in increased usage and engagement.

REFERENCES

Adams, W. C. (2015). Conducting semi-structured interviews. In H. P. Hatry, J. S. Wholey, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (pp. 492-505). John Wiley & Sons. https://doi.org/10.1002/9781119171386.ch19

Alvermann, D. E. (2005). Literacy on the edge: How close are we to closing the literacy achievement gap? *Voices in the Middle, 13*(1), 8-14.

Andrade, H. L., & Heritage, M. (2018). *Using formative assessment to enhance learning, achievement, and academic self-regulation*. Taylor & Francis Group.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117-148. https://doi.org/10.1207/s15326985ep2802_3

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice, 25*(6), 551-575. https://doi.org/10.1080/0969594X.2018.1441807

Bloom, B. S. (1971). Mastery learning. In J. H. Block (Ed.), *Mastery learning: Theory and practice* (pp. 47-63). Holt, Rinehart & Winston.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Leadership, 41*(8), 4-17. https://doi.org/10.3102/0013189X013006004

Capin, P., Walker, M. A., Vaughn, S., & Wanzek, J. (2018). Examining how treatment fidelity is supported, measured, and reported in K–3 reading intervention research. *Educational Psychology Review, 30*, 885-919. https://doi.org/10.1007/s10648-017-9429-z

Carioti, D., Masia, M. F., Travellini, S., & Berlingeri, M. (2021). Orthographic depth and developmental dyslexia: A meta-analytic study. *Annals of Dyslexia, 71*(3), 399-438. https://doi.org/10.1007/s11881-021-00226-0

Carroll, V. (1997). *Learning to read, reading to learn* (ED413588). ERIC. https://files.eric.ed.gov/fulltext/ED413588.pdf

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5-51. https://doi.org/10.1177/1529100618772271

Center on Standards and Assessment Implementation. (2019). *Implementation of computerized adaptive assessments: CSAI brief* (ED603014). ERIC. https://files.eric.ed.gov/fulltext/ED603014.pdf

Chall, J. S. (1967). *Learning to read: The great debate; an inquiry into the science, art, and ideology of old and new methods of teaching children to read, 1910-1965*. McGraw-Hill.

Chall, J. S., & Jacobs, V. A. (2003). The classic study on poor children's fourth-grade slump. *American Educator, 27*, 14-15. https://www.aft.org/periodical/american-educator/spring-2003/classic-study-poor-childrens-fourth-grade-slump

Cheatham, J. P., & Allor, J. H. (2012). The influence of decodability in early reading text on reading achievement: A review of the evidence. *Reading & Writing, 25*(9), 2223-2246. https://doi.org/10.1007/s11145-011-9355-2

Cook, M., & Ross, S. M. (2022). *The impact of i-Ready Personalized Instruction with fidelity on 2021 MCAS ELA achievement: Final report* (ED622942). ERIC. https://eric.ed.gov/?id=ED622942

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.

Curriculum Associates. (n.d.). *The science behind i-ready's adaptive diagnostic*. https://www.setda.org/ls2013/wp-content/uploads/sites/8/2014/12/Adaptive-Diagnostic-Science-SETDA.pdf

Curriculum Associates. (2020). *Efficacy research for i-Ready personalized instruction: Reading and mathematics*. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-essa-l3-research-brief-2020.pdf

Curriculum Associates. (2021a). *i-Ready: Aligned with the science of reading.* https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-science-of-reading-overview-2021.pdf

Curriculum Associates. (2021b). *i-Ready diagnostic assessment.* https://www.curriculumassociates.com/programs/i-ready-assessment/diagnostic

Curriculum Associates. (2021c). *i-Ready efficacy research summary.* https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-efficacy-research-summary-2021.pdf

Curriculum Associates. (2021d). *Is i-Ready dangerous.* https://www.curriculumassociates.com/blog/is-iready-dangerous

Curriculum Associates. (2022a). *i-Ready diagnostic: Correlation study with North Carolina End-of-Grade (NC EOG) tests.* https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-diagnostic-assessments-correlation-study-north-carolina-2022.pdf

Curriculum Associates. (2022b). *Using i-ready personalized instruction with fidelity: Results from the 2018—2019 school year*. https://www.curriculumassociates.com/-/media/mainsite/files/i-ready/iready-efficacy-2018-2019-research-brief-2022.pdf

Curriculum Associates. (2023a). *i-Ready learning: Instructional resources for grades K-8.* https://www.curriculumassociates.com/programs/i-ready-learning

Curriculum Associates. (2023b). *i-Ready Diagnostic grades K-12 scale score placement tables.* *https://i-readycentral.com/pdfs/i-ready-scale-score-placement-tables/*

Curriculum Associates. (2023c). Get good data. *i-Ready Central*. https://i-readycentral.com/articles/get-good-data/

Davis, M. R. (2012, October 15). Adaptive testing evolves to assess common-core skills. *Education Week*. https://www.edweek.org/teaching-learning/adaptive-testing-evolves-to-assess-common-core-skills/2012/10

Decker, B. C. (1986). Aliteracy: What teachers can do to keep Johnny reading. *Journal of Teacher Education*, *37*(6), 55–58. https://doi.org/10.1177/002248718603700609

Dredger, K. (2013). Challenging rebellious adolescent aliteracy. *ALAN Review*, *40*(3), 117–122. https://doi.org/10.21061/alan.v40i3.a.25

Dumas, J. E., Lynch, A. M., Laughlin, J. E., Phillips Smith, E., & Prinz, R. J. (2001). Promoting intervention fidelity. conceptual issues, methods, and preliminary results from the EARLY ALLIANCE prevention trial. *American Journal of Preventive Medicine, 20*(1 Suppl), 38-47. https://doi.org/10.1016/s0749-3797(00)00272-5

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350. https://doi.org/10.1007/s10464-008-9165-0

Dusenbury, L., Brannigan, R., Falco, M., & Lake, A. (2004). An exploration of fidelity of implementation in drug abuse prevention among five professional groups. *Journal of Alcohol & Drug Education, 47*(3), 4-19. https://www.jstor.org/stable/45092333

EdWeek Research Center. (2020). *Early reading instruction: Results of a national survey of K-2 and elementary special education teachers and postsecondary instructors.* https://www.edweek.org/research-center/research-center-reports/early-reading-instruction-results-of-a-national-survey

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the national reading panel's meta-analysis. *Review of Educational Research, 71*(3), 393-447. https://doi.org/10.3102/00346543071003393

Fine, J., Zygouris-Coe, V., Senokossoff, G., & Fang, Z. (2011). Secondary teachers' knowledge, beliefs, and self-efficacy to teach reading in the content areas: Voices following professional development. In M. S. Plakhotnik, S. M. Nielsen, & D. M. Pane (Eds.), *Proceedings of the Tenth Annual College of Education & GSN Research Conference* (pp. 24-28). Florida International University. https://core.ac.uk/reader/46946440

Fisher, D., Frey, N., & Lapp, D. (2023). Veteran teachers' understanding of "Balanced literacy." *Journal of Education, 203*(1), 188-195. https://doi.org/10.1177/00220574211025980

Fixsen, D., Blase, K., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based programs. *Exceptional Children, 79*(2), 213-230. https://doi.org/10.1177/00144029130790020

Flesch, R. (1955). *Why Johnny can't read--and what you can do about it*. Harper.

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Heinemann.

Gambrell, L. B., Malloy, J. A., & Mazzoni, S. A. (2011). Evidence-based best practices in comprehensive literacy instruction. In L. M. Morrow, & L. B. Gambrell (Eds.), *Best practices in literacy instruction* (pp. 11-36). Guilford Press.

George, S. (2021). Literacy, illiteracy and the impending danger of aliteracy. *Access (Online), 35*(4), 23-28. https://search.informit.org/doi/10.3316/informit.181460839519928

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6*(4), 126-135. https://doi.org/10.1080/19388076709556976

Goodman, K. S. (1989). Whole-language research: Foundations and development. *Elementary School Journal, 90*(2), 207-221. https://doi.org/10.1086/461613

Goodwin, A. P., & Jiménez, R. T. (2020). The science of reading: Supports, critiques, and questions. *Reading Research Quarterly, 55*(S1), S7-S16. https://doi.org/10.1002/rrq.360

Gottfried, A. W. (1985). Measures of socioeconomic status in child development research: Data and recommendations. *Merrill-Palmer Quarterly, 31*(1), 85-92. https://www.jstor.org/stable/23086136

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6-10. https://doi.org/10.1177/074193258600700104

Great Schools Partnership. (2015). *Personalized learning.* Glossary of Educational Reform. https://www.edglossary.org/personalized-learning/

Guskey, T. R. (2005). *Formative classroom assessment and Benjamin S. Bloom: Theory, research, and implications.* https://files.eric.ed.gov/fulltext/ED490412.pdf

Guthrie, J. T., Klauda, S. L., & Ho, A. N. (2013). Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. *Reading Research Quarterly*, *48*(1), 9–26. https://doi.org/10.1002/rrq.035

Guthrie, J. T., Van Meter, P., McCann, A. D., Wigfield, A., Bennett, L., Poundstone, C. C., Rice, M. E., Faibisch, F. M., Hunt, B., & Mitchell, A. M. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly, 31*(3), 306-332. https://doi.org/10.1598/RRQ.31.3.5

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research, Vol. 3,* pp. 403-422. Erlbaum. https://doi.org/10.4324/9781315200613

Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools? *Exceptional Children, 79*(2), 181-193. https://doi.org/10.1177/001440291307900204

Harn, B. A., Parisi Damico, D., & Stoolmiller, M. (2017). Examining the variation of fidelity across an intervention: Implications for measuring and evaluating student learning. *Preventing School Failure: Alternative Education for Children and Youth, 61*(4), 289-302. https://doi.org/10.1080/1045988X.2016.1275504

Harter, S., Whitesell, N. R., & Kowalski, P. (1992). Individual differences in the effects of educational transition on young adolescent's perceptions of competence and motivational orientation. *American Educational Research Journal, 29*(4), 777-807. https://doi.org/10.3102/00028312029004777

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Heller, R., & Greenleaf, C. L. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement.* Carnegie. https://www.carnegie.org/publications/literacy-instruction-in-the-content-areas-getting-to-the-core-of-middle-and-high-school-improvement/

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "Educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest, 16*(1), 3-34. https://doi.org/10.1177/1529100615569721

Hollingsworth, H. (2023). *Why more U.S. schools are embracing a new 'science of reading'.* PBS News Hour. https://www.pbs.org/newshour/education/why-more-u-s-schools-are-embracing-a-new-science-of-reading

Hoover, W. A., & Tunmer, W. E. (2020). *The cognitive foundations of reading and its acquisition: A framework with applications connecting teaching and learning*. Springer.

Horn, M. B., & Staker, H. (2011). *The rise of K-12 blended learning.* https://www.christenseninstitute.org/wp-content/uploads/2013/04/The-rise-of-K-12-blended-learning.pdf

Ivey, G., & Fisher, D. (2006). *Creating literacy-rich schools for adolescents*. ASCD.

Jerald, C. D. (2001). *Dispelling the myth revisited: Preliminary findings from a nationwide analysis of "high-flying" schools* (ED462485). ERIC. https://eric.ed.gov/?id=ED462485

Johnson, D. R., Cushman, G. K., Borden, L. A., & McCune, M. S. (2013). Potentiating empathic growth. *Psychology of Aesthetics, Creativity, and the Arts, 7*(3), 306-312. https://doi.org/10.1037/a0033261

Kamil, M. L., Borman, G. D., Dole, J. A., Kral, C. C., Salinger, T. S. & Torgesen, J. K. (2008). *Improving adolescent literacy: Effective classroom and intervention practices. IES practice guide. NCEE 2008-4027.* National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/docs/practiceguide/adlit_pg_082608.pdf

Kazakoff, E. R. (2021). *Educational equity: The transformative impact of effective instruction and professional learning.* https://www.lexialearning.com/sites/default/files/Equitable Literacy Education White Paper.pdf

Kim, J. S. (2008). Research and the reading wars. *Phi Delta Kappan, 89*(5), 372-375. https://doi.org/10.1177/003172170808900514

Kroger, J., Martinussen, M., & Marcia, J. E. (2010). Identity status change during adolescence and young adulthood: A meta-analysis. *Journal of Adolescence, 33*(5), 683-698 https://pubmed.ncbi.nlm.nih.gov/20004962/

Krueger, R. A. (1998). *Developing questions for focus groups*. Sage.

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). *The alphabetic principle and learning to read* (ED427291). ERIC. https://files.eric.ed.gov/fulltext/ED427291.pdf

Lorimor-Easley, N. A., & Reed, D. K. (2019). *An explanation of structured literacy, and a comparison to balanced literacy.* Iowa Reading Research Center. https://iowareadingresearch.org/blog/structured-and-balanced-literacy#:~:text=Balanced%20Literacy%20is%20a%20%E2%80%9Cphilosophical,Fountas%20%26%20Pinnell%2C%201996).

Lotter, C. R., Thompson, S., Dickenson, T. S., Smiley, W. F., Blue, G., & Rea, M. (2018). The impact of a practice-teaching professional development model on teachers' inquiry instruction and inquiry efficacy beliefs. *International Journal of Science and Mathematics Education, 16*(2), 255-273. http://dx.doi.org/10.1007/s10763-016-9779-x

Łuniewska, M., Chyl, K., Dębska, A., Kacprzak, A., Plewko, J., Szczerbiński, M., Szewczyk, J., Grabowska, A., & Jednoróg, K. (2018). Neither action nor phonological video games make dyslexic children read better. *Scientific Reports, 8*(1), 549. https://doi.org/10.1038/s41598-017-18878-7

Lysaker, J. T., Tonge, C., Gauson, D., & Miller, A. (2011). Reading and social imagination: What relationally oriented reading instruction can do for children. *Reading Psychology, 32*(6), 520-566. https://doi.org/10.1080/02702711.2010.507589

Macaruso, P., Wilkes, S., & Prescott, J. E. (2020). An investigation of blended learning to support reading instruction in elementary schools. Educational Technology Research and *Development, 68*(6), 2839-2852. https://doi.org/10.1007/s11423-020-09785-2

McGlynn, K., & Kelly, J. (2017). Using formative assessments to differentiate instruction. *Science Scope, 41*(4), 22-25. https://doi.org/10.2505/4/ss17_041_04_22

Meisels, S. J., Atkins-Burnett, S., Xue, Y., Bickel, D. D., Son, S., & Nicholson, J. (2003). Creating a system of accountability: The impact of instructional assessment on elementary children's achievement test scores. *Education Policy Analysis Archives, 11*(9) https://www.researchgate.net/publication/256011325_Creating_a_system_of_accountability_The_impact_of_instructional_assessment_on_elementary_children's_achievement_test_scores

Mertens, D. M., & Wilson, A. T. (2019). *Program evaluation theory and practice: A comprehensive guide* (2nd ed.). Guilford Press.

Metsala, J. L., & David, M. D. (2016). Individual differences in word recognition: Reading acquisition and reading disabilities. In P. Afflerbach (Ed.), *Handbook of individual differences in reading*. Taylor & Francis.

Mikulecky, L. (1978). *Aliteracy and a changing view of reading goals* (ED157052). ERIC. https://files.eric.ed.gov/fulltext/ED157052.pdf

Moats, L. (2020). Teaching reading is rocket science: What expert teachers of reading should know and be able to do. *American Educator, 44*(2), 4-39. https://www.aft.org/ae/summer2020/moats

Moats, L. C., & Tolman, C. A. (2022). *LETRS: Volume 1: Units 1-4*. Lexia.

Molenaar, I., & Roda, C. (2008). Attention management for dynamic and adaptive scaffolding. *Pragmatics & Cognition, 16*(2), 224-271. https://doi.org/10.1075/p&c.16.2.04mol

National Center for Education Statistics. (2019). *Nation's report card.* https://www.nationsreportcard.gov/

National Center for Education Statistics. (2022). *NAEP report card: 2022 NAEP reading assessment.* https://www.nationsreportcard.gov/highlights/reading/2022/

National Center on Intensive Intervention. (n.d.). *i-Ready Diagnostic Reading*. https://charts.intensiveintervention.org/screening/tool/?id=7534542c4f422f85

National Center on Intensive Intervention. (2013). *Data-based individualization: A framework for intensive intervention.* https://intensiveintervention.org/sites/default/files/DBI_Framework.pdf

National Center on Intensive Intervention. (2021). *Academic screening tools chart.* https://charts.intensiveintervention.org/ascreening

National Early Literacy Panel. (2008). *Developing early literacy: Report of the national early literacy panel.* Washington, DC: National Institute for Literacy. https://lincs.ed.gov/publications/pdf/NELPReport09.pdf

National Institute of Child Health and Human Development. (2000). *Report of the national reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, DC: U.S. Government Printing Office. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Neal, J. C., & Kelly, P. R. (2002). Delivering the promise of academic success through late intervention. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 18*(2), 101-117. https://doi.org/10.1080/10573560252808503

Neugebauer, S., Hopkins, M., & Spillane, J. (2019). Social sources of teacher self-efficacy: The potency of teacher interactions and proximity to instruction. *Teachers College Record, 121*(4), 1-32. https://doi.org/10.1177/016146811912100405

North Carolina Department of Public Instruction. (2021). *The North Carolina Department of Public Instruction grades 3-8 reading End-of-Grade (EOG) and high school English II End-of-Course (EOC) edition 5 technical report 2020-21*. https://www.dpi.nc.gov/technical-report-eog-reading-and-eoc-english-ii-2020-21-edition-5/open

North Carolina Department of Public Instruction. (2022a). *NC check-ins grades 3 and 6 reading specifications.* https://www.dpi.nc.gov/documents/accountability/testing/nccheckin/nc-check-ins-reading-grades-3-and-6-test-specifications

North Carolina Department of Public Instruction. (2022b). *North Carolina end-of-grade reading tests at grades 3-8 test specifications*. https://www.dpi.nc.gov/documents/accountability/testing/eog/eog-reading-grades-3-8-test-specifications

North Carolina Department of Public Instruction. (2022c). *North Carolina personalized assessment tool (NCPAT) pilot: NC check-ins 2.0 grades 3–8 reading specifications.* https://www.dpi.nc.gov/documents/accountability/testing/nccheckin/nc-check-ins-20-reading-grades-3-8-specifications

Obama, B. (2005). Bound to the word. *American Libraries, 36*(7), 48-52. https://americanlibrariesmagazine.org/bound-to-the-word/

O'Connor, P. (2020). Visible learning and whole language: Revisiting the 'garbage in, garbage out' problem. *Australian Journal of Language & Literacy, 43*(2), 141-151. https://doi.org/10.1007/bf03652050

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78*(1), 33-84. https://doi.org/10.3102/0034654307313793

Pearson, P. D. (2004). The reading wars. *Educational Policy, 18*(1), 216-252. https://doi.org/10.1177/0895904803260041

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001) *Knowing what students know: The science and design of educational assessment.* The National Academies Press. https://doi.org/10.17226/10019

Peters, J. L., De Losa, L., Bavin, E. L., & Crewther, S. G. (2019). Efficacy of dynamic visuo-attentional interventions for reading in dyslexic and neurotypical children: A systematic review. *Neuroscience and Biobehavioral Reviews, 100*(May 2019), 58-76. https://doi.org/10.1016/j.neubiorev.2019.02.015

Petscher, Y., Cabell, S. Q., Catts, H. W., Compton, D. L., Foorman, B. R., Hart, S. A., Lonigan, C. J., Phillips, B. M., Schatschneider, C., Steacy, L. M., Terry, N. P., & Wagner, R. K. (2020). How the science of reading informs 21st-Century education. *Reading Research Quarterly, 55*(Suppl 1), S267-S282. https://doi.org/10.1002/rrq.352

Pressley, M. (2005). *Reading instruction that works: The case for balanced teaching* (3rd ed.). Guilford Press.

Pytash, K. E., & O'Byrne, W. I. (2018). Research on literacy instruction and learning in virtual, blended, and hybrid environments. In K. Kennedy, & R. E. Ferdig (Eds.), *Handbook of research on K-12 online and blending learning* (pp. 303-314). ETC Press. https://doi.org/10.1184/R1/6686813

Quinn, D. M., & Kim, J. S. (2017). Scaffolding fidelity and adaptation in educational program implementation: Experimental evidence from a literacy intervention. *American Educational Research Journal, 54*(6), 1187-1220. https://doi.org/10.3102/0002831217717692

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press.

Reutzel, D. R. (2022). *Putting scientific evidence to work in reading instruction.* https://www.curriculumassociates.com/topics/science-of-reading/ray-reutzel-reading-instruction

Reyhner, J. (2020). *Reading wars: Phonics versus whole language.* Northern Arizona University. https://jan.ucc.nau.edu/~jar/Reading_Wars.html

Rodriguez, M. C. (2004). The role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education, 17*(1), 1-24. https://doi.org/10.1207/s15324818ame1701_1

Ross, P. T., & Bibler Zaidi, N. L. (2019). Limited by our limitations. *Perspectives on Medical Education, 8*(4), 261-264. https://doi.org/10.1007/s40037-019-00530-x

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage.

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman, & D. K. Dickinson (Eds.), *Handbook for research in early literacy* (pp. 97-110). Guilford Press.

Shanahan, T. (2020). What constitutes a science of reading instruction? *Reading Research Quarterly, 55*(S1), S235-S247. https://doi.org/10.1002/rrq.349

Shepard, L. A., Penuel, W. R., & Davidson, K. L. (2017). Design principles for new systems of assessment. *Phi Delta Kappan, 98*(6), 47-52. https://doi.org/10.1177/0031721717696478

Smith, F. (1971). *Understanding reading: A psycholinguistic analysis of reading and learning to read*. Holt, Rinehart, & Winston.

Stiggins, R., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice, 44*(1), 11-18. https://doi.org/10.1207/s15430421tip4401_3

Strambler, M. J., & Weinstein, R. S. (2010). Psychological disengagement in elementary school among ethnic minority students. *Journal of Applied Developmental Psychology*, *31*(2), 155–165. https://doi.org/10.1016/j.appdev.2009.11.006

Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International handbook of educational evaluation* (Chapter 2). Springer Science+Business Media.

Swain, K., Whitley, R., McHugo, G. J., & Drake, R. E. (2010). The sustainability of evidence-based practices in routine mental health agencies. *Community Mental Health Journal, 46*(2), 119-129. https://doi.org/10.1007/s10597-009-9202-y

Swain, M., Randel, B., & Dvorak, R. N. (2020). *Impact evaluation of reading i-Ready instruction for middle school grades using 2018-19 data* (ED604745). ERIC. https://files.eric.ed.gov/fulltext/ED604745.pdf

Tesch, R. (1990). *Qualitative research: Analysis types and software tools*. Falmer.

Thomas B. Fordham Foundation. (2023). *Whole language lives on: The illusion of balanced reading instruction.* LD OnLine: All About Learning Disabilities and ADHD. https://www.ldonline.org/ld-topics/teaching-instruction/whole-language-lives-illusion-balanced-reading-instruction

Thomas, P. L. (2022). *The science of reading movement: The never-ending debate and the need for a different approach to reading instruction.* National Education Policy Center. https://nepc.colorado.edu/publication/science-of-reading

UNESCO Institute for Statistics. (2023). *Literacy.* UNESCO Glossary. https://uis.unesco.org/en/glossary-term/literacy

U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, What Works Clearinghouse. (n.d.). *Using the WWC to find ESSA tiers of evidence.* https://ies.ed.gov/ncee/wwc/essa

Verhoeven, L., & Snow, C. E. (2001). Literacy and motivation: Bridging cognitive and sociocultural viewpoints. In L. Verhoeven, & C. E. Snow (Eds.), *Literacy and motivation: Reading engagement in individuals and groups* (pp. 1-18). L. Erlbaum Associates.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping, teaching and learning* (pp. 53-82). Lawrence Erlbaum Associates.

World Literacy Foundation. (2018). *The economic and social cost of illiteracy* [White paper]. https://worldliteracyfoundation.org/wp-content/uploads/2021/07/TheEconomicSocialCostofIlliteracy-2.pdf

Yaden, D. B., Reinking, D., & Smagorinsky, P. (2021). The trouble with binaries: A perspective on the science of reading. *Reading Research Quarterly, 56*(S1), S119-S129. https://doi.org/10.1002/rrq.402

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Sage.

APPENDIX A

ASD STUDENT ACHIEVEMENT DATA

The tables and graphs included in Appendix A were presented at the ASD Board of Education meeting on November 16, 2021. The full presentation can be found online.

Aurelius School District Grades 3-8
5-Year Data Trend

Burke Elementary

Year	Grade Level	Math	Reading	Science
2016	Grade 3	46.7	50	
	Grade 4	42.4	55.6	
	Grade 5	49.5	36.1	55.7

Year	Grade Level	Math	Reading	Science
2017	Grade 3	42.6	43.4	
	Grade 4	48.5	45.5	
	Grade 5	39	24	57.1

Year	Grade Level	Math	Reading	Science
2018	Grade 3	39.3	32	
	Grade 4	41.5	48.3	
	Grade 5	52.6	43.9	76.3

Year	Grade Level	Math	Reading	Science
2019	Grade 3	41.3	44.6	
	Grade 4	45.1	36.1	
	Grade 5	49.6	37.4	76.4

Year	Grade Level	Math	Reading	Science
2021	Grade 3	21.8	25.9	
	Grade 4	23.1	26.4	
	Grade 5	21.7	25.3	34.9

2016-2021

Medoc Elementary

Year	Grade Level	Math	Reading	Science
2016	Grade 3	62.5	58.3	
	Grade 4	53.5	51.2	
	Grade 5	78.4	59.8	90.7

Year	Grade Level	Math	Reading	Science
2017	Grade 3	62.1	54.4	
	Grade 4	58.5	48.3	
	Grade 5	60.8	58.3	77.1

Year	Grade Level	Math	Reading	Science
2018	Grade 3	63.6	50.5	
	Grade 4	65.3	58.6	
	Grade 5	76.2	58.2	82

Year	Grade Level	Math	Reading	Science
2019	Grade 3	64.9	42.6	
	Grade 4	61.1	49.5	
	Grade 5	70	61.8	86.2

Year	Grade Level	Math	Reading	Science
2021	Grade 3	40.7		
	Grade 4	26.3	35.2	
	Grade 5	15.4	25.8	44

2016-2021

Cole Middle School

Year	Grade Level	Math	Reading	Science
2016	Grade 6	47.3	50.5	
	Grade 7	38.2	39	
	Grade 8	34.9	43.4	67.7

Year	Grade Level	Math	Reading	Science
2017	Grade 6	66.7	55.2	
	Grade 7	51.1	54.1	
	Grade 8	39.4	40.2	69

Year	Grade Level	Math	Reading	Science
2018	Grade 6	58.8	61.1	
	Grade 7	55.9	61.9	
	Grade 8	42.1	50	71

Year	Grade Level	Math	Reading	Science
2019	Grade 6	69.8	59.5	
	Grade 7	60.7	50.9	
	Grade 8	48.2	44.4	72.3

Year	Grade Level	Math	Reading	Science
2021	Grade 6	31.9	37.6	
	Grade 7	40.1	42.3	
	Grade 8	23	38.7	65.7

2016-2021

Table A1. ASD Grades 3-8 NC EOG data for 2016-2021 disaggregated by school. Reading and mathematics are assessed yearly in Grades 3-8, and science is assessed in Grades 5 and 8.

5-Year Data Trends - EOG Reading/Math/Science

Year	RRGSD Reading (3-8 exc. 20-21)	State Reading Average	RRGSD Math (Grades 3-8)	State Math Average	RRGSD Science (Grades 5 & 8)	State Science Average
2016	47.8	56.9	47.4	54.7	70.2	72.7
2017	47.3	57.5	51.5	55.4	68.4	72.8
2018	52.7	57.3	54.4	56.1	75.6	72.1
2019	48.7	57.2	57.9	58.6	77	75.5
2021	35.7	45.6	28.7	40	55.1	62.2

Table A2. ASD Grades 3-8 NC EOG data for 2016-2021 as compared to the NC state average.

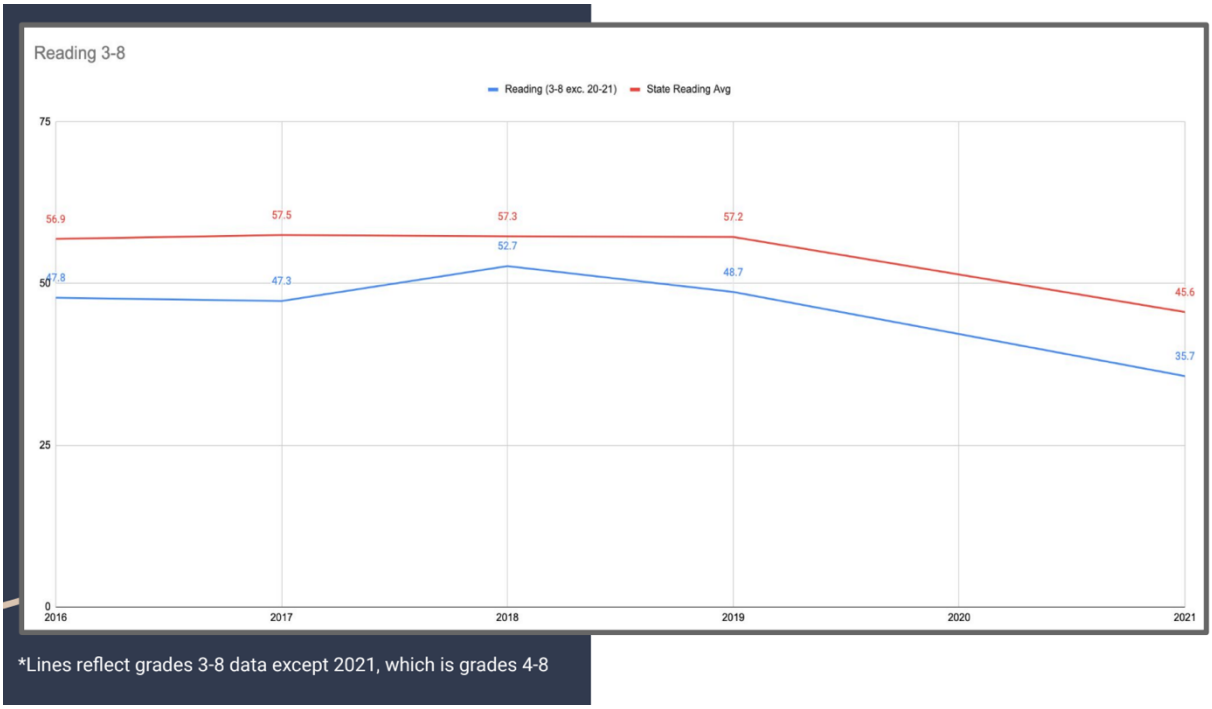


Figure A1. Reading Grades 3-8 NC EOG data for 2016-2021 for ASD (blue) as compared to the NC state average (red).

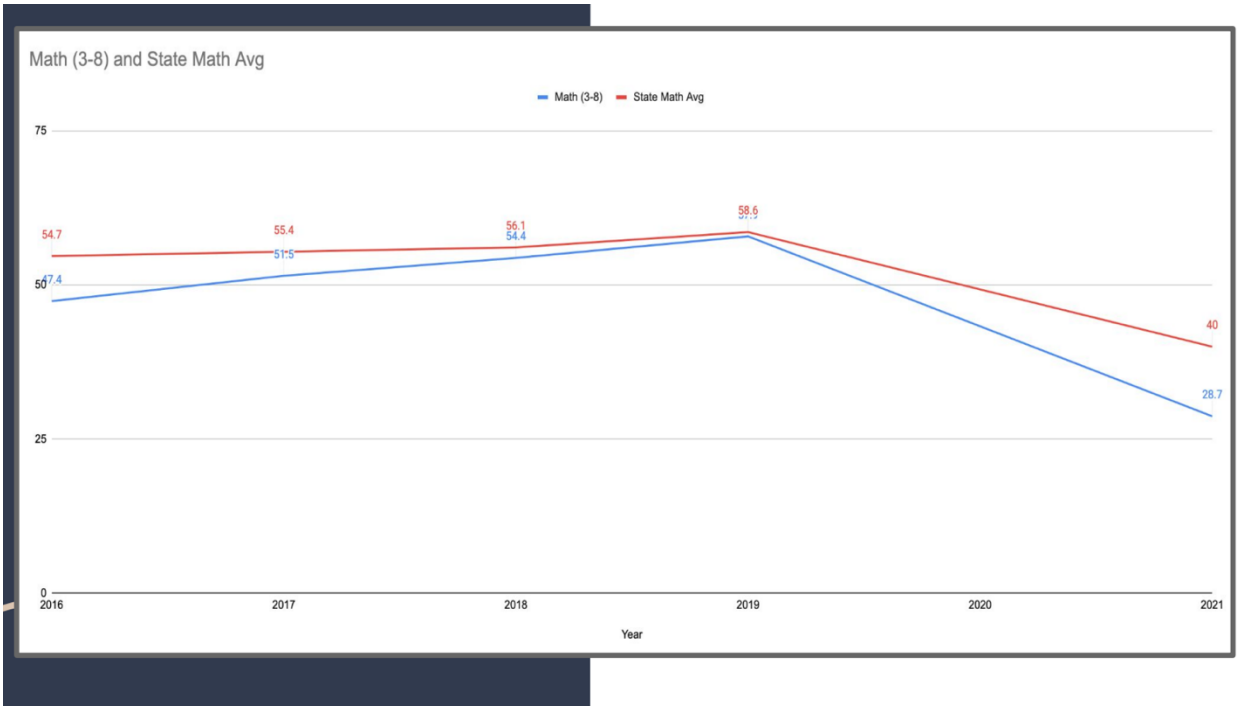


Figure A2. Math Grades 3-8 NC EOG data for 2016-2021 for ASD (blue) as compared to the NC state average (red).

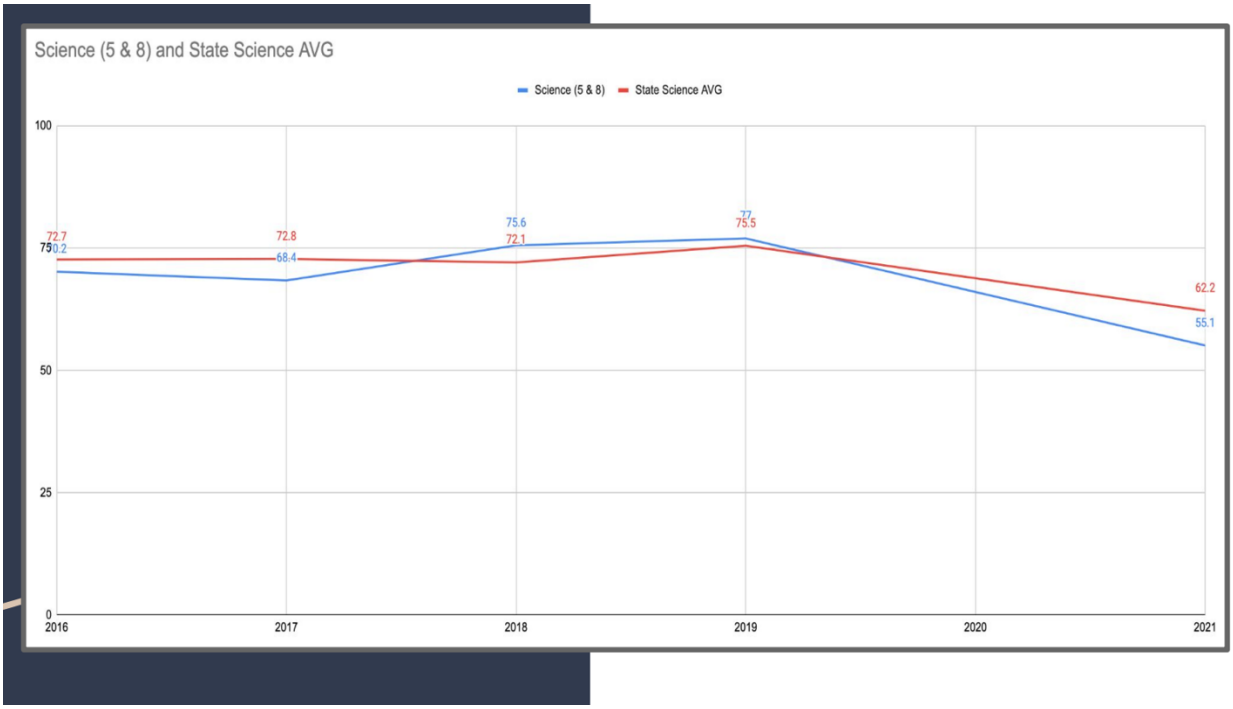


Figure A3. Science Grades 5 and 8 NC EOG data for 2016-2021 for ASD (blue) as compared to the NC state average (red).

APPENDIX B

Summary of Review of Related Literature

Philosophies of Teaching Reading

The importance of teaching students to read well (Moats, 2020), the "Reading Wars" (Castles et al., 2018; Chall, 1967; Kim, 2008; Yaden et al., 2021), and high number of non-proficient readers (NCES, 2022)

Phonics

Explicit, systematic instruction to teach students how to sound out words based on how they are spelled (Chall, 1967; Flesch, 1955), helps children learn to read better than all forms of control group instruction (Ehri et al., 2001) with a 0.60 effect size (Hattie, 2009), and should be an essential part of beginning reading and reading remediation programs (Ehri et al., 2001; NELP, 2008; NICHD, 2000)

Whole Language

Meaning-based process to create a literacy-rich, motivating, and engaging environment for children that emphasizes experiences of text and uses cueing strategies such as guessing (Goodman, 1967; Reyhner, 2020; Smith, 1971), but with only a 0.06 effect size, it has a "negligible effect on learning to read" (Hattie, 2009)

Balanced Literacy

Blends instruction and support on phonics, whole language, and captivating literature with student choice and engagement to get students reading books they enjoy as quickly as possible (Carroll, 1997; Fountas & Pinnell, 1996; Hollingsworth, 2023; Pressley, 2005) through shared, guided, and independent reading (Lorimor-Easley & Reed, 2019), but teachers' understanding and implementation have been vague and focus more on whole language (Fisher et al., 2023; TBFF, 2023)

Science of Reading

Combines objective investigation and reliable evidence about how humans learn to read and how reading should be taught (Goodwin & Jimenez, 2020) with instruction emphasizing phonics-rich, skill-driven, intentionally designed, and systematic approach to foundational word recognition and language comprehension (Gough & Tunmer, 1986; Hoover & Tunmer, 2020; Scarborough, 2001), and curriculum grounded in its five components is more effective than other approaches (NELP, 2008; NICHD, 2000)

- **Structured Literacy** - uses SVR (Gough & Tunmer, 1986) and Scarborough's Reading Rope (2001)
- **LETRS** – staff development program (Moats & Toulman, 2022)

Adolescent Literacy

According to the 2022 NAEP, only 31% of grade 8 students are reading proficiently and able to comprehend text at their grade level (NCES, 2022) and without continuous literacy instruction, non-proficient 6th graders will likely never catch up (Kamil et al., 2008)

Adolescent Literacy

Includes text comprehension, vocabulary, and communicating ideas (Kamil et al., 2008)

Adolescent Aliteracy

Being able to read, but choosing not to do so (Verhoeven & Snow, 2001)

Strategies for Addressing Adolescent Literacy

- 4 whole-school recommendations (Ivey & Fisher, 2006)
- 5 classroom recommendations (Kamil et al., 2008)

Secondary Teachers Self-Efficacy in Teaching Non-Proficient Adolescent Readers

Self-efficacy (Bandura, 1997); teachers feel unprepared (Heller & Greenleaf, 2007)

Coherent and Adaptive Assessment Systems

Assessment aims to determine how well students are learning and can be used to assist learning, measure individual achievement, and evaluate programs and ideally students use *assessments **as** learning* (Pellegrino et al., 2001)

Formative Assessments

*Assessment **for** learning*

Provide feedback to inform on skill progress for small groups, targeted instruction, and scaffolding (McGlynn & Kelly, 2017), which yields substantial learning gains and reduces achievement gaps (Black & Wiliam, 1998)

Summative Assessments

*Assessment **of** learning*

Occur at the end of an instructional period, evaluative, often high stakes, and have reputation for unintended consequences for learning and motivation (Andrade & Heritage, 2018)

Coherence in Educational Assessment

Uses multiple, aligned approaches to collect formative and summative information about student learning along with rich curriculum and equity-focused research on learning (Shepard et al., 2017) to drive systematic instruction (Reutzel, 2022), to track student growth continuously and consistently, and to provide meaningful student data (Curriculum Associates, n.d.)

Adaptivity in Educational Assessment

Uses multiple aspects of proficiency, chart progress over time, including multiple paths of performance, and model performance at various levels (Pellegrino et al, 2001) to create an in-depth, personalized assessment of each student

- **Fixed-Form Assessments** – all questions in the same order (CSAI, 2019)
- **Computer-Adaptive Assessments** (limitations and benefits) – different questions personalized for each student (CSAI, 2019) using the Rasch Model (1980)

Effective Instructional Interventions

Effective interventions are research-based and systematically delivered in small groups or individually by a trained professional based on a targeted area of need with a proactive, catch-up stance to enable acceleration to become on-grade level (Neal & Kelly, 2002)

Intensive Interventions

Individualized support for students when core or supplemental instruction is not working (NCII, 2013)

Blended Learning Interventions

Combine teacher-directed and student-centered, personalized, and scaffolded digital activities to close learning gaps (Hirsh-Pasek et al., 2015)

Mastery Learning

Uses instruction, formative assessment, feedback, and either individual corrective activities and second chance formative assessments or enrichment and extension activities (Bloom, 1971)

Motivation and Engagement

Important moderators for learning (Kamil et al., 2008) and key in adolescent literacy development (Alvermann, 2005; Guthrie & Wigfield, 2000; Kamil et al., 2008)

- **Strategies to Increase Motivation and Engagement** (Kamil et al., 2008)
- **Gamification** (Hirsh-Pasek et al., 2015)

Fidelity of Implementation

Implementation fidelity occurs when a program or intervention is delivered the way it was designed and in the way it was implemented during the research studies validating its effectiveness (NCII, 2013)

Fidelity of Implementation Measurement Criteria

(1) Adherence (2) Duration (3) Quality of Delivery (4) Participant Responsiveness (5) Program Differentiation (O'Donnell, 2008) but improving student outcomes (not perfect fidelity) is the most important goal (Harn et al., 2017)

i-Ready

A program that purports to combine effective reading instruction, coherent and adaptive assessment systems, and effective instructional interventions (if implemented with fidelity) to improve adolescent literacy is i-Ready Reading (Curriculum Associates, 2022a)

i-Ready Diagnostic (3x/year)

Leverages computer-adaptive technology (Rasch Model [1980]) to assess proficiency, track growth, and provide a personalized evaluation with feedback that informs teacher-led instruction through Teacher Toolbox and creates a personal pathway on i-Ready Instruction (Curriculum Associates, 2021a)

i-Ready Instruction (30-49 min/week)

Personalized lessons aligned to science of reading and structured literacy and designed to motivate with corrective feedback, student engagement, and student-owned progress monitoring (Curriculum Associates, 2023)

Effectiveness of i-Ready

- **Research Conducted by Curriculum Associates** (Curriculum Associates, 2020, 2021c, 2022a, 2022b)
- **Independent Research on i-Ready** (Cook & Ross, 2022; NCII, 2021; Swain et al., 2020)

Synthesizing the Literature:

When curriculum and instruction aligned to the science of reading are coupled with coherent assessment systems and personalized interventions, student achievement, especially for non-proficient adolescent readers, increases (Kazakoff, 2021).

APPENDIX C

SURVEY PARTICIPANT INFORMED CONSENT FORM

I,_____, agree to participate in a research study regarding my experiences with the i-Ready Program in _____ District middle school classrooms. The purpose of this study is to inform stakeholders who make decisions about reading curriculum, instruction, and assessment and to gain stakeholders' perspectives on the i-Ready usage.

As a participant, I understand that my participation in the study is purposeful and voluntary. All administrators and instructional staff in schools using the i-Ready will have the opportunity to participate voluntarily in the individual survey. I understand that I will be expected to participate in one (1) survey related to my knowledge and implementation of the i-Ready program in the middle school.

I understand that the interviewer has been trained in the research of human subjects, my responses will be confidential, and that my name will not be associated with any results of this study. Information from the survey will be safeguarded, so my identity will never be disclosed. My true identity will not be associated with the research findings.

I understand that there is no known risk or discomfort directly involved with this research and that I am free to withdraw my consent and discontinue participation at any time. I agree that should I choose to withdraw my consent and discontinue participation in the study that I will notify the researcher listed below, in writing. A decision not to participate in the study or to withdraw from the study will not affect my relationship with the researcher, the district, The College of William and Mary generally or the School of Education, specifically.

If I have any questions or problems that may arise as a result of my participation in the study, I understand that I should contact Juliana Thompson, the researcher, at 252-678-4749 or thompsonji.co@rrgsd.org; Dr. Margaret Constantino, dissertation chair, at constantinome@wm.edu; Dr. Tom Ward, chair of EDIRC, at 757-221-2358 or EDIRC-L@wm.edu; or Dr. Jennifer Stevens, Chair of the PHSC at 757-221-3862 or jastev@wm.edu.

My signature below signifies that I am at least 18 years of age, that I have received a copy of this consent form, and that I consent to participate in this research study.

_____ _____

Signature of Participant Date

_____ _____

Signature of Researcher Date

THIS PROJECT WAS FOUND TO COMPLY WITH APPROPRIATE ETHICAL STANDARDS AND WAS EXEMPTED FROM THE NEED FOR FORMAL REVIEW BY THE W&M PROTECTION OF HUMAN SUBJECTS COMMITTEE (Phone 757-221-3966) ON 2023-07-18 AND EXPIRES ON 2024-07-18.

APPENDIX D

SURVEY QUESTIONS

Survey Question	Question Type	Evaluation Question
What grade do you teach?	I	
In which academic years have you used i-Ready?	T	3
Respond to the following questions using this scale: (a) daily, (b) weekly, (c) monthly, (d) quarterly, (e) yearly, or (f) never. 1. I use i-Ready's Historical Reports to see my class's testing history. 2. I use i-Ready's Historical Reports to see an individual student's testing history. 3. I lead data chats with my students using i-Ready Diagnostic Results Report and Diagnostic Growth Report. 4. I use the Diagnostic Growth Report to create and discuss typical and stretch growth goals with students. 5. I use i-Ready's Diagnostic Results Report to create student groups to intervene on specific skill deficits. 6. I use the Personalized Instruction Summary Report to see students' weekly usage of i-Ready Instruction. 7. I use the Personalized Instruction Summary Report to see students' lessons passed in i-Ready Instruction. 8. I use the instructional usage spreadsheet to guide conversations with students about their i-Ready Instruction usage and progress. 9. I use Teacher Toolbox to provide specific interventions that are personalized to their needs to small groups or individual students. 10. I use Think Up!-ELA resources with my students to integrate critical thinking instruction on ELA standards. 11. I use Ready-NC resources with my students to build conceptual understanding through reasoning, practice, and productive discussion around real-world scenarios. 12. I use Ready Reading Discourse Cards with my students to promote richer, deeper dialogue to engage all students in meaningful ELA conversations. 13. I use PHONICS for Reading to help struggling students become fluent and independent readers. 14. I use the Grade Level Scaffolding Report to plan small groups and/or use reading buddies.	K	1

Respond to the following questions using this scale: (a) always, (b) often, (c) sometimes, or (d) never. 15. I prepare students to take the i-Ready Diagnostic by using the preparation materials found in i-Ready Central. 16. I administer and proctor the i-Ready Diagnostic for all students in my class three times per year. 17. I chunk the i-Ready Diagnostic over multiple days, so students do not test for more than 30 minutes per day. 18. I use i-Ready's Diagnostic Status Report to check student progress on the i-Ready Diagnostic. 19. I reset the i-Ready Diagnostic for individual students whose data on the Diagnostic Status Report indicates a red flag or yellow flag for rushing.	K	1
Respond to the following questions using this scale: (a) yes, (b) no, or (d) I am not sure. 20. Over 80% of students completed all three i-Ready Diagnostic assessments in 2022-23. 21. Over 80% of my students used i-Ready Instruction for 30-49 minutes per week for at least 18 weeks in 2022-23. 22. Over 80% of my students passed 70-100% of lessons in i-Ready Instruction in 2022-23.	K	1
Respond to the following questions using this scale: (a) strongly agree, (b) agree, (c) disagree, or (d) strongly disagree. 23. I have a scheduled weekly time for my students to use i-Ready Reading that ensures they access it 30-49 minutes per week.	K	1
Rate the value of each of the following i-Ready program elements using this scale: (a) very valuable, (b) somewhat valuable, or (c) not valuable. 24. Student data generated from the i-Ready Diagnostic assessment 25. Student instruction and practice in i-Ready Instruction 26. Instructional resources provided in Teacher Toolbox 27. Overall impact that i-Ready is having on student achievement in reading	K	3
Should the middle school continue, modify, or discontinue i-Ready use?	E	3
If you selected "modify," please share what should be modified.	E	3
Is there anything else you would like to share about the i-Ready program?	E	1, 2, 3

Note: I = Introductory question; T = Transitional question; K = Key question; E = Ending question

APPENDIX E

STAFF SEMI-STRUCTURED INDIVIDUAL INTERVIEW GUIDE

Project: A Program Evaluation of i-Ready Implementation in a North Carolina Middle School

Date of Interview:

Time of Interview:

Location of Interview:

Interviewee:

Position of Interviewee:

Interviewer: Juliana Thompson

Introduction

[Begin with customary pleasantries and introductions.]

"Thank you for making time to talk with me today about the i-Ready program. You were selected to participate because of your experiences as an instructional leader in the middle school that uses i-Ready. Your participation is important because I want to gain your perceptions about the impact of i-Ready on our students. There are no right or wrong answers to these questions. Please feel free to share your point of view, even if it may not be complimentary of the i-Ready program in some way. Your responses will become part of my doctoral research on i-Ready program outcomes for our middle school, and in aggregate, will be shared with district leadership. The information we discuss today will help our district team better understand our implementation of the i-Ready program and improve how we teach students how to read. Our conversation should last no longer than ninety minutes. To facilitate my note-taking, I would like to audio record our conversation today for transcription and analysis, and I will provide a copy of the transcription to you to verify accuracy. Please note that all of your responses will remain confidential, and I will redact identifying information from the transcript. Only I will be privy to the recordings, which will be eventually destroyed after they are transcribed. You may withdraw from this interview at any time without penalty.

Explained and Obtain Informed Consent

"In order to participate, you must sign a form devised to meet our human subject requirements. Essentially, this document states that (1) all information will be held confidential, (2) your participation is voluntary, and you may stop at any time if you feel uncomfortable, and (3) I do not intend to inflict any harm. Protecting your rights is essential to me. If you share any personally identifiable information, I will remove it from the responses that you give before reporting any findings. Nothing that you say in the course of this interview will be used in any personnel evaluation of you, and nothing that you say will be attributed to you to be shared with your supervisor. I want you to feel comfortable answering questions authentically. Are you willing to proceed with the interview? If you are, please sign the participant informed consent form.

Semi-Structured Interview Participant Informed Consent Form

I,_____, agree to participate in a research study regarding my experiences with the i-Ready Program in _____ District middle school classrooms. The purpose of this study is to inform stakeholders who make decisions about reading curriculum, instruction, and assessment and to gain stakeholders' perspectives on the i-Ready usage.

As a participant, I understand that my participation in the study is purposeful and voluntary. All administrators and instructional staff in schools using the i-Ready will have the opportunity to participate voluntarily in the individual interviews. I understand that I will be expected to participate in one (1) semi-structured interview related to my knowledge and implementation of the i-Ready program in the middle school.

I understand that the interviewer has been trained in the research of human subjects, my responses will be confidential, and that my name will not be associated with any results of this study. I understand that the data will be collected using an audio recording device and then transcribed for analysis. Information from the audio recording and transcription will be safeguarded, so my identity will never be disclosed. My true identity will not be associated with the research findings.

I understand that there is no known risk or discomfort directly involved with this research and that I am free to withdraw my consent and discontinue participation at any time. I agree that should I choose to withdraw my consent and discontinue participation in the study that I will notify the researcher listed below, in writing. A decision not to participate in the study or to withdraw from the study will not affect my relationship with the researcher, the district, The College of William and Mary generally or the School of Education, specifically.

If I have any questions or problems that may arise as a result of my participation in the study, I understand that I should contact Juliana Thompson, the researcher, at 252-678-4749 or thompsonji.co@rrgsd.org; Dr. Margaret Constantino, dissertation chair, at me constantino@wm.edu; Dr. Tom Ward, chair of EDIRC, at 757-221-2358 or EDIRC-L@wm.edu; or Dr. Jennifer Stevens, Chair of the PHSC at 757-221-3862 or jastev@wm.edu.

My signature below signifies that I am at least 18 years of age, that I have received a copy of this consent form, and that I consent to participate in this research study.

_____ _____

Signature of Participant Date

_____ _____

Signature of Researcher Date

THIS PROJECT WAS FOUND TO COMPLY WITH APPROPRIATE ETHICAL STANDARDS AND WAS EXEMPTED FROM THE NEED FOR FORMAL REVIEW BY THE W&M PROTECTION OF HUMAN SUBJECTS COMMITTEE (Phone 757-221-3966) ON 2023-07-18 AND EXPIRES ON 2024-07-18.

APPENDIX F

STAFF SEMI-STRUCTURED INDIVIDUAL INTERVIEW QUESTIONS

[Confirm that the participant has signed the consent form. Turn on the digital recorder and test it. Begin recording.]

"Thank you for agreeing to participate. I have planned this interview to last no longer than 90 minutes. During this time, I have several questions that I'd like to cover. If time begins to run short, it may be necessary to interrupt you in order to push ahead and complete this line of questioning. Do you have any questions before we begin?"

Semi-Structured Interview Question	**Question Type**	**Evaluation Question**
What do you find most rewarding about being an educator?	O	
What are some of the online curriculum, instruction, and assessment programs that you have observed teachers or students using in the middle school?	I	
When you were introduced to using i-Ready, what were your initial impressions?	T	3
What are the most valuable aspects of the i-Ready Diagnostic in reading? *(What aspects are most beneficial to teachers as they provide reading instruction? What aspects are most beneficial to students as they participate in reading instruction?)*	K	3
What are the most valuable aspects of the i-Ready Instruction in reading? *(What aspects are most beneficial to teachers as they provide reading instruction? What aspects are most beneficial to students as they participate in reading instruction?)*	K	3
What are the most valuable aspects of the i-Ready Teacher Toolbox in reading?	K	3
Describe how you use data generated from i-Ready Diagnostic to inform your lesson planning.	K	1
Describe how you have seen students' reading proficiency improve after participation in i-Ready Instruction. (*Can you tell me a story about a student whose reading proficiency or enthusiasm for reading increased?)*	K	2
How do you ensure that students are meeting the required amount of time using i-Ready Instruction each week?	K	1
What barriers exist to students meeting the required amount of time using i-Ready Instruction each week?	K	1
What other instructional strategies (or programs) are you using to supplement reading instruction? – *Listen for acceleration, personalization, gamification*	K	1, 3
What are the challenges of implementing the i-Ready program in reading as it is intended? What ideas do you have for overcoming these challenges?	K	1, 3
Describe the types of reading instruction you provide to students who test out of i-Ready Instruction.	K	1
In what ways has implementing the i-Ready program impacted student reading proficiency?	K	1, 2, 3
What ideas do you have to modify the i-Ready program, so it will be more effective in reading?	E	3
Is there anything else we should have talked about but did not?	E	1, 2, 3

[Thank the participant for his/her time, cooperation, and participation in sharing his/her experiences and perceptions in this interview. Assure him or her that you will provide an interview transcript and the final research product.]

APPENDIX G

DESCRIPTIVE STATISTICS FOR READING ACHIEVEMENT DATA

Grade 6 Reading Descriptive Statistics

Statistics

		C1	C2	C3	BOY	MOY	EOY	EOG
N	Valid	202	205	207	207	211	210	211
	Missing	17	14	12	12	8	9	8
Mean		13.16	15.35	13.35	547.51	564.23	564.60	549.74
Median		13.00	16.00	14.00	552.00	570.00	572.50	549.00
Mode		17	23	17	513[a]	622	567	549
Std. Deviation		4.993	5.568	5.463	62.352	60.230	64.942	9.297
Variance		24.933	31.002	29.840	3887.727	3627.712	4217.418	86.432
Skewness		-.050	-.344	-.258	-.698	-.939	-1.069	.172
Std. Error of Skewness		.171	.170	.169	.169	.167	.168	.167
Kurtosis		-.928	-.830	-1.184	.656	1.587	1.689	-.643
Std. Error of Kurtosis		.341	.338	.337	.337	.333	.334	.333
Range		20	23	20	338	360	405	42
Minimum		3	1	2	340	321	277	532
Maximum		23	24	22	678	681	682	574

a. Multiple modes exist. The smallest value is shown

Note. C1 = NC Check-In #1; C2 = NC Check-In #2; C3 = NC Check-In #3; BOY = beginning-of-year i-Ready Diagnostic; MOY = middle-of-year i-Ready Diagnostic; EOY = end-of-year i-Ready Diagnostic; EOG = North Carolina End of Grade Test.

Grade 7 Reading Descriptive Statistics

Statistics

		C1	C2	C3	BOY	MOY	EOY	EOG
N	Valid	184	185	185	189	189	184	185
	Missing	10	9	9	5	5	10	9
Mean		11.40	12.33	11.81	561.01	574.08	575.80	550.16
Median		11.00	12.00	12.00	573.00	590.00	587.00	550.00
Mode		6[a]	13	13[a]	538[a]	618	583[a]	550
Std. Deviation		5.426	4.973	4.862	65.031	66.927	68.013	9.368
Variance		29.444	24.733	23.643	4229.048	4479.265	4625.823	87.752
Skewness		.363	-.013	.110	-.886	-1.062	-1.127	-.001
Std. Error of Skewness		.179	.179	.179	.177	.177	.179	.179
Kurtosis		-.931	-.785	-.696	.557	1.392	1.622	-.818
Std. Error of Kurtosis		.356	.355	.355	.352	.352	.356	.355
Range		22	22	21	322	381	392	41
Minimum		2	2	2	360	317	330	531
Maximum		24	24	23	682	698	722	572

a. Multiple modes exist. The smallest value is shown

Note. C1 = NC Check-In #1; C2 = NC Check-In #2; C3 = NC Check-In #3; BOY = beginning-of-year i-Ready Diagnostic; MOY = middle-of-year i-Ready Diagnostic; EOY = end-of-year i-Ready Diagnostic; EOG = North Carolina End of Grade Test.

Grade 8 Reading Descriptive Statistics

Statistics

		C1	C2	C3	BOY	MOY	EOY	EOG
N	Valid	212	213	208	205	212	195	215
	Missing	16	15	20	23	16	33	13
Mean		11.28	11.85	11.29	574.09	575.10	586.55	553.81
Median		11.00	12.00	11.00	585.00	585.00	596.00	554.00
Mode		8	7[a]	8	570[a]	558[a]	581[a]	555
Std. Deviation		4.730	5.233	4.706	64.847	68.500	64.257	9.397
Variance		22.372	27.389	22.150	4205.130	4692.207	4128.939	88.305
Skewness		.230	.007	.508	-.833	-.770	-1.321	.313
Std. Error of Skewness		.167	.167	.169	.170	.167	.174	.166
Kurtosis		-.520	-.745	-.271	1.077	1.024	3.690	-.516
Std. Error of Kurtosis		.333	.332	.336	.338	.333	.346	.330
Range		23	23	22	375	415	451	41
Minimum		0	0	2	337	303	268	537
Maximum		23	23	24	712	718	719	578

a. Multiple modes exist. The smallest value is shown

Note. C1 = NC Check-In #1; C2 = NC Check-In #2; C3 = NC Check-In #3; BOY = beginning-of-year i-Ready Diagnostic; MOY = middle-of-year i-Ready Diagnostic; EOY = end-of-year i-Ready Diagnostic; EOG = North Carolina End of Grade Test.

VITA

Name: Juliana ("Julie") Irene Owen Thompson

Educational Background

Superintendent: Roanoke Rapids Graded School District (RRGSD), Roanoke Rapids, NC, September 2021-present
Assistant Superintendent, RRGSD, Roanoke Rapids, NC, July 2020 - August 2021
Executive Director of Instructional Services (Chief Academic Officer), RRGSD, Roanoke Rapids, NC, July 2017 - June 2020
Assistant Principal, RRGSD, Roanoke Rapids, NC, July 2016 - June 2017
Principal Resident, RRGSD, Roanoke Rapids, NC, August 2015 - June 2016
English Teacher, RRGSD, Roanoke Rapids, NC, August 2008 - July 2015
English Teacher, York County School Division, Yorktown, VA, August 2007 - July 2008

Degrees

Doctor of Education, Educational Policy, Planning, and Leadership, K-12 Administration: William & Mary, Williamsburg, VA, January 2024
Master of School Administration: Northeast Leadership Academy (NELA), North Carolina State University, Raleigh, NC, May 2016
Master of Arts in Education: Reading, Language, and Literacy: The College of William and Mary**,** Williamsburg, VA, August 2010
Bachelor of Arts: English, Monroe Scholar, The College of William and Mary, Williamsburg, VA, May 2007

Related Training and Experience

Crucial Conversations
Facilitative Leadership
Understanding by Design
Covey's 7 Habits/The Leader in Me
Digital Storytelling Retreat
Conflict Resolution
Flipped Classroom Training
Moving EVAAS beyond Teacher Evaluation
Ron Clark, Kim Bearden, and The Ron Clark Academy
Literacy and LETRS
Grant Writing and Community Engagement
School Culture Professional Development
EVAAS and Indistar/NCStar
MTSS/PBIS
Effective Use of Questioning
Distinguished Leadership in Practice
Social Justice
Adult Learning Theory
Non-Violent Crisis Intervention
Literacy Design Collaborative
Advanced Placement training in English Language and Composition and in English Literature and Composition