2019

The Effect of Comprehensive Instructional Program on Grades 3-8 Student Achievement

Marcia Shortt
Lincoln Memorial University, mshortt@wisek12.org

LMU | Carter & Moyers School *of* Education
LINCOLN MEMORIAL UNIVERSITY

Final Dissertation Approval
Form 11

The Effect of Comprehensive Instructional Program on Grades 3-8 Student Achievement

This is to certify that I have examined this copy of the doctoral dissertation by

Marcia L. Shortt

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final examining committee have been made.

Cherie Gaines	8/27/19
Committee Chair	Date
	8/27/19
Committee Member	Date
Shannon D. Collins	08/27/19
Committee Member	Date
	8/27/19
EdD Program Director	Date
	8/27/19
Dean, School of Education	Date

THE EFFECT OF COMPREHENSIVE INSTRUCTIONAL PROGRAM ON GRADES 3-8 STUDENT ACHIEVEMENT

Dissertation

Submitted in partial fulfillment
of the requirements for the degree of Doctor of Education
in the Carter and Moyers School of Education
at Lincoln Memorial University

by

Marcia L. Shortt

August 2019

Dedication

I would like to dedicate this dissertation to my husband, Ronnie, who graciously and unconditionally supported me along my career and educational advancement. Without the care and encouragement from him, I would not have had the strength nor fortitude to complete the program.

Acknowledgments

I would like to thank my professors, dissertation committee, and Chair from Lincoln Memorial University for their unwavering support, guidance, and dedication. My Chair, Dr. Cherie Gaines, provided words of encouragement at times when I felt I could not move forward, as well as the expertise required to complete such an enormous task. The talented leaders on my committee provided the foundation to build and complete this project. I would like to thank Dr. Andrew Courtner for agreeing to be a member of my committee and his unwavering support and assistance with the quantitative sections of my dissertation. Finally, I would like to thank Dr. Matt Hurt, Comprehensive Instructional Program director, who provided guidance, support, and conversation along the way that helped me to reason through much of my research and findings.

Abstract

Student achievement and the state Standards of Learning testing was a critical criterion for success for school districts, administrators, and teachers. The researcher in this study investigated the efficacy of the Comprehensive Instructional Program in improving student achievement in reading and mathematics and whether school configuration impacted student achievement. The population consisted of all students in District A in Virginia who took achievement tests in reading and math in 2012-2018. The researcher determined there was a significant difference in reading and math scores between pre-implementation and post-implementation in grades 3-8 as measured by the Virginia Standards of Learning testing program and that school configuration had no impact on the pass rate of the tests.

Table of Contents

List of Tables

Table **Page**

Chapter I: Introduction

The passage of the No Child Left Behind Act of 2001 (NCLB) marked a new era in American educational policy (Lee & Reeves, 2012; Meyers, 2012; Overbaugh & Lu, 2008; Wieczorek, 2017). NCLB created an accountability structure that endangered the funding of schools and generated anxiety for principals and teachers in public schools throughout the United States (Bautista & Wong, 2017; Eros, 2013; Glover, Reddy, Kettler, Kurz, & Lekwa, 2016; Kopcha, 2012; Robinson, Myran, Strauss, & Reed, 2014; Smith & Kovacs, 2011). Scholars argued that NCLB institutionalized the era of so-called technician teachers—that is, teachers responsible for teaching a centrally planned curriculum in increasingly standardized ways (DeMatthews, 2015; Grissom, Nicholson-Crotty, & Harrington, 2014; Ronfeldt, Farmer, McQueen, & Grissom, 2015). Additionally, scholars argued that limiting the autonomy of teachers to better achieve NCLB goals and priorities reduced teachers' self-concept, self-efficacy, and motivation (DeMatthews, 2015; Grissom et al., 2014; Ronfeldt et al., 2015). NCLB mandated that students take achievement tests in reading and math in each grade 3-8 and once more in high school between grades 10-12 (Lee & Reeves, 2012; Meyers, 2012; Overbaugh & Lu, 2008; Wieczorek, 2017). NCLB also mandated tests in other subjects such as science, but the main substance of the law was reading and mathematics testing (Lee & Reeves, 2012; Meyers, 2012; Overbaugh & Lu, 2008; Wieczorek, 2017).

School districts took up the challenge of creating an educational process that would meet NCLB standards by raising the student achievement levels in subjects such as reading and mathematics. One school district in southwestern Virginia, in Region 7, referred to as District A, an economically disadvantaged county in that state, created a program that demonstrated student achievement gains (Hurt, 2015). After that school

district experienced success using that locally developed instructional program, a consortium of educators from across Virginia modified the program, called the Comprehensive Instructional Program (CIP), and made it available to all districts in the state. Five principles formed the basis of the CIP, including expectations and standards for academic excellence with no excuses; communications of those expectations and standards to all stakeholders and ensure buy-in by all stakeholders; creation of a practical course of action designed specifically to achieve the academic goals; constant measurements and review of data to ensure ongoing progress and course adjustments defined where needed; and a focus on improving the people involved, especially the teachers, so they could provide the best instruction possible (Hurt, 2015). The program development group provided the resources, lesson plans, and assessments needed to implement the CIP (CIP, 2016). The CIP included such materials for grades 3-8 and for reading, math, science, and history, though schools had the option of not implementing completely the all-inclusive CIP in all subjects and all grades.

Due to decreasing student Standards of Learning (SOL) test scores that might result in removing state accreditation for some schools, in the 2015-16 school year, public schools in 19 of the 132 districts in the state of Virginia public school system began implementation of the CIP, designed to both standardize student instruction and improve student achievement (CIP, 2016). In a benchmark study after one year of broad-scale implementation in regions across the state, CIP administrators found there were strong correlations (i.e., 0.67 or higher) between achieving all the pre-defined benchmarks in the CIP and student achievement scores in the annual SOL tests (CIP, 2016). That study was short term (i.e., one academic year only), however, and only included schools that partially implemented CIP in only some subjects or some grade

levels. Thus, while the benchmark results were generally positive, the results did not clearly delineate which student achievement improvements were specific to CIP and which might have been the result of other factors.

Twenty-seven school districts within all eight regions of the Virginia Department of Education (VDOE) used the CIP because of the 2017-2018 school year data (Hurt, 2017). The educators who formalized the CIP guaranteed that the costs to implement the program would be minimal to ensure that schools in economically disadvantaged areas of the state could take advantage of the CIP if they chose to do so (Hurt, 2017). Since 2015, those schools that have chosen to implement completely the full CIP instructional process have shown the greatest student achievement gains as measured by the Virginia SOL testing requirements for each grade level 3-8 (Hurt, 2017). Proponents of the CIP have thus declared this instructional program to be a great success.

Statement of the Problem

Student achievement, as measured by the results of the state SOL testing, was a critical criterion for success for school districts, administrators, and teachers. The federal Race to the Top program assessed schools and districts and encouraged teacher performance assessments that weighed student achievement scores as a major measurement factor; this program controlled $4.3 billion in 2015 (Chingos, Whitehurst, & Gallaher, 2015). The results of these performance measures affected teacher tenure and promotions, while the assessment processes were determined at the local district level, often using weak or unsupported theories (Chingos et al., 2015). Furthermore, schools and districts could lose both federal funding and state accreditation with inadequate student achievement scores (Chingos et al., 2015). In addition to federal dollars, private philanthropical foundations also influenced school district funding, often

basing funding decisions on student achievement test scores (Chingos et al., 2015). Unfortunately, the research on effectiveness of various instructional pedagogies, such as the CIP and other programs, commonly addressed only successful schools, making the results of their studies of questionable validity. For example, Chingos et al. (2015) found that two-thirds of studies of the effect of various instructional programs included only high-performing districts, and fewer than 15% included districts with a variety of instructional approaches. Chingos et al. (2015) also found that nearly all such studies were case studies, with very few using quantitative measures of success.

Region 7, the Southwest region of Virginia, began implementation of the CIP in the 2015-2016 school year across all grades and all subjects. Figure 1 illustrates the regional organization of schools in Virginia. Region 7 is the blue region on the map in the southwestern portion of the state.

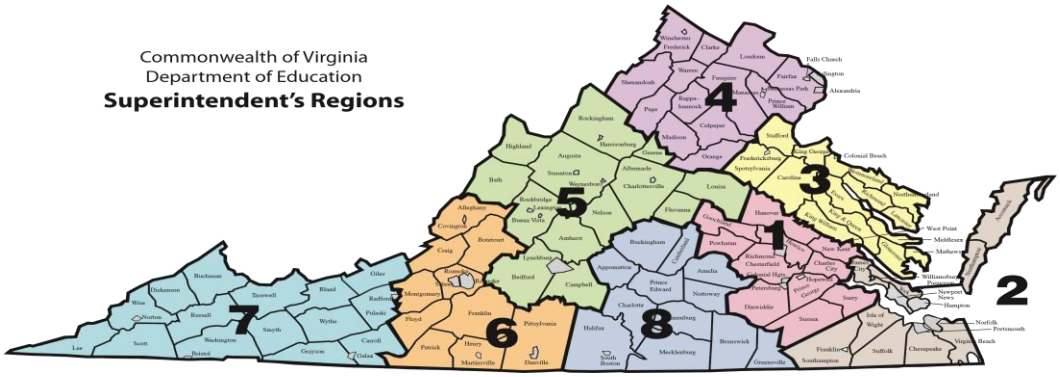

Figure 1. Map of Superintendent's Regions for Schools in Virginia (VDOE, 2018).

Region 7 was the birthplace of the CIP in large part due to its underachievement in educational test scores. Region 7 traditionally had the second-lowest per-pupil expenditures in the state and the lowest starting teacher salaries in the state. Furthermore, Region 7 had the second highest rate of students living in poverty in the state, with more

than 57% of students in need, and Region 7 had the highest number of students with disabilities, with approximately 15% of students having at least one disability. Prior to the implementation of the CIP, Region 7 educational systems experienced such negative effects such as declining student test scores, declining enrollments, declining budgets, little help from the state government to improve education, and overall instructional materials that did not align with the state's SOL standards (VDOE, 2018). By all economic measures, Region 7 should be one of the lowest-performing regions in terms of student academic achievement (VDOE, 2018).

By the conclusion of the 2016-2017 school year, the region had two full academic years of experience with the CIP. During this time, Region 7 became the top-performing region in reading and mathematics as well as other subjects, despite the region being economically disadvantaged (CIP, 2018). Proponents credited the implementation of the CIP in this region for the improvement in Region 7 SOL achievement scores (CIP, 2018). When compared to the aggregate scores of all the other regions in the state of Virginia, Region 7's aggregate scores in math, reading, writing, history, and science demonstrated the most growth (Hurt, 2017). With that said, those results met or exceeded the goals of the SOL, including the 15 counties and 4 cities of Region 7 (VDOE, 2018). While that success was notable on a district basis rather than the overall Region 7, no statistical study has confirmed whether the district's CIP implementation could be associated with student improvement or whether the improvements were statistical outliers or due to other unspecified causes.

The district studied, District A, unlike other districts in the region, chose to do a complete implementation of the CIP across all grades and all subjects in the 2015-2016 school year. District A also experienced some improvements in SOL achievement test

scores since implementing CIP district-wide. The problem that arose from the implementation of the CIP in District A was that it was unknown whether the CIP program was a statistically significant factor in the observed improvement of student Virginia SOL test scores. While the SOL test scores in Region 7 improved in general, no studies have determined if the improvements were statistically significant. Without regard to whether that program generated a significant improvement in student achievement, all student SOL test score improvements credited the success to the CIP. Thus, conducting a statistical study was vital to understand if CIP implementation may have generated the changes in student achievement scores in District A or whether other factors may have led to the district student achievement improvements.

The goal of this researcher was to define the efficacy of the CIP in improving student achievement in reading and mathematics as well as to determine if school configuration (i.e., K-4, 5-8, K-8) made a difference in student achievement. Although Hurt (2017) ascribed credit for student achievement improvements to the implementation of CIP in those regions that implemented the program, there existed a lack of statistical studies to confirm the relationship. Other factors may have influenced student achievement in both reading and math as described in the literature review.

Another issue of concern within District A was how the school configuration of elementary grades within the district may affect student achievement. Elementary schools in District A fall under one of K-4, 5-8, or K-8 school grade level configurations. Of concern was whether the presence or absence of both early elementary and late elementary grades in a single school affected student achievement in reading and mathematics. No prior study has considered this aspect of school configuration in determining student achievement.

The intent of this study was to provide data that helped explain the impact of the CIP program on Virginia school District A by comparing pre-CIP student achievement from the three years prior to CIP implementation and student achievement in the three years post-CIP implementation. This researcher's goal in the study was to determine the efficacy of CIP implementation on grades 3-8 reading and math SOL scores. Reading and math were tests that all students in all grades took each year, while students did not take tests in other subjects such as science and history every year. In addition, reading and math were fundamental skills that affected achievement in many other subjects. The study also included analysis of data comparing achievement in reading and math for schools in each of the K-4, 5-8, and K-8 school configurations to determine if such grade level inclusion or exclusion in the school configuration affected student achievement scores. The purpose of this study was to determine if improvements in student SOL test scores since the implementation of CIP in District A were statistically significant and thus reflective of genuine improvements in student achievement as well as to determine which school configuration was most conducive to improvements in student achievement.

Research Questions

The researcher developed the following questions to guide this study.

Research question 1. What differences existed in student achievement scores (pass percentages) in reading in grades 3-8 as measured by the Virginia Standards of Learning testing program between pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

Research question 2. What differences existed in student achievement scores (pass percentages) in math in grades 3-8 as measured by the Virginia Standards of Learning testing program comparing pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

Research question 3. What differences, if any, existed among K-4, 5-8, and K-8 school configurations in student achievement scores (pass percentages) in reading in grades 3-8 as measured by the Virginia Standards of Learning testing program comparing pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

Research question 4. What differences, if any, existed among K-4, 5-8, and K-8 school configurations in student achievement scores (pass percentages) in math in grades 3-8 as measured by the Virginia Standards of Learning testing program comparing pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

Theoretical Framework

Illeris (2015) developed the social learning theory, which was the theoretical foundation for this study. Social learning theory represented a general theory of behavior with an emphasis on learning from the social environment (Mowrer, 1960). Illeris (2015) examined each dimension of cognitive, emotional, and social learning and later integrated the three separate dimensions to explain the learning process as a whole, the social

learning theory. This theory included a learning triangle as its central element in which the three legs of content, interaction, and incentive were located within a circle of society (i.e., the social context of the school) (see Figure 2) (Illeris, 2015). Three key elements in Illeris's dimensions of learning included instructional content, interactions teachers have with students, and incentives students have to learn the instructional content. In the context of this study, the measure of instructional content consisted of the changes in the student SOL achievement test scores.

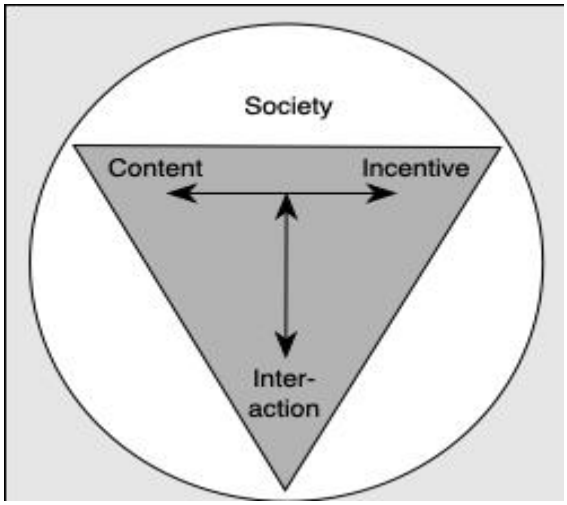

Figure 2. Three dimensions of learning (Illeris, 2015).

In the context of elementary education, the learner's acquisition of knowledge included both content and incentive, which operated in an integrated manner. Illeris (2015) posited that all learning took place in all three of the learning legs. In the social learning theory, Illeris (2015) also supposed that learning could be either an addition to the learner's existing knowledge or a reconstruction of existing knowledge to reorganize knowledge and generate new understanding. The variations in individual students' prior knowledge and their methods of accommodating and adapting knowledge into new forms

provided an explanation for why different students had different preferred learning styles (Illeris, 2015).

Illeris's social learning theory included the topic of learning barriers, which appeared as a lack of understanding of the information presented or an adaptation of the information in a distorted or inaccurate manner (Illeris, 2015). Other learning barriers included resistance the learner possessed to the knowledge presented due to knowledge antithetical to the learner's beliefs or that was unwanted for any reason (Illeris, 2015). Critical to this study was that the theory differentiated between the content teachers presented as opposed to the content the students learned (Illeris, 2015). The competency measures of the Virginia SOL addressed the information that students absorbed, while the CIP defined the knowledge that the teachers taught. Potential discrepancies between knowledge taught and knowledge learned described the fundamental dichotomy that described student achievement issues.

This social learning theory served as the basis of this study since the CIP implemented in the district incorporated the theory's interactions between learner and environment. In particular, the CIP concept was one of linking student incentive and interest into the content of the class and emphasizing the interaction between the student and the material presented. In addition, the Illeris (2015) social learning theory addressed the issue of learning barriers, which was an issue in District A. Region 7 as a whole was an economically disadvantaged region, and District A was not an exception. The CIP program had a focus on student achievement and addressed only those aspects of teaching and schools that influenced student success. By getting administrators, teacher leaders, and district leaders all focused on helping students learn and achieve more, the CIP program provided an exemplar of the Illeris (2015) social learning theory in

implementation. In particular, the CIP keys to success in Region 7 included that the focus of attention of everyone from individual teachers to school superintendents was on the task of helping students learn, overcoming all barriers to learning, and that all other aspects of education—other than student safety—were subordinate to the goal of achievement (VDOE, 2018). The implementation of the CIP used the Illeris (2015) social learning theory by having a strong focus on removing barriers to learning.

Significance of the Project

Continual monitoring of the successes and shortcomings of the CIP program was an important element of the implementation of the program. The district personnel used formative benchmark test scores and summative SOL test scores to adjust the implementation and further improve student achievement (Hurt, 2015). The development of CIP thus incorporated a process of continual improvement and progress. To achieve success, it was essential to understand the specifics of how well the program improved student achievement by grade levels, by schools, by subject, and by three dimensions of learning—content, incentive, and interaction.

Region 7 consisted of 15 counties, 4 cities, and 19 school districts, thus the achievement of the region as a whole may not reflect explicitly on the achievement of individual school districts within the region. An exploration of the student achievement improvement in the district in this study determined the link between student achievement and CIP implementation. In addition, while most school districts that implemented the CIP had done so on a piecemeal basis (i.e., one school or one grade or even one teacher at time) in District A, all grades and all schools implemented CIP simultaneously in a single academic year. This region offered a chance to decide if the CIP had a factually and noteworthy effect on achievement with an examination that could

plainly characterize achievement based on CIP usage. Training backed usage in four different ways: leading needs evaluations; utilizing the standards of both human-focused outline and enhancement science to plan and test thoughts, refine, and consistently enhance execution; co-planning point-by-point activity intended to encourage the arrangement or program implementers execute the huge thoughts; and working hand-in-glove to actualize the plans. This study included data from three years prior to CIP implementation to three years after CIP implementation in both reading and math achievement scores for all students and all schools in grades 3-8 of one Region 7, District A, in Virginia.

In this study, the researcher provided stakeholders in District A with the information needed to understand the progress in student achievement. These data helped the researcher to identify grade levels and subjects that required changes to the CIP implementation to improve student achievement. These data provided the researcher direction to assist all stakeholders with implementation of any necessary improvements to the educational process and to observe the workings of the program.

The significance of this project thus lay in defining the efficacy of the CIP in improving student achievement in reading and mathematics in grades 3-8 eight in District A within Region 7 in Virginia. Although Hurt (2015) ascribed credit for student achievement improvements to the implementation of CIP, there existed a lack of statistical studies to confirm that relationship. By conducting this study, the researcher closed that gap by providing statistical analysis of the student achievement data to identify the relationship between CIP and student achievement improvement in District A within Region 7, as well as to determine whether CIP created statistically significant improvements in student achievement. The researcher also determined if school

configuration within District A made a difference in improvements in student achievement based on reading and math SOL test scores (pass percentages).

Description of the Terms

Comprehensive Instructional Program (CIP). The CIP provided teachers with lesson plans, classroom activities, and assessments, all closely aligned with Virginia SOL and with the goal of improving student achievement on the annual SOL examinations (CIP, 2018). One school district in Virginia originally developed this curriculum program for Virginia public schools (Hurt, 2015). After implementation of CIP in the developing district resulted in improved student achievement test scores, other school districts across the state adopted CIP into their curricula.

Math achievement. As used in this study, math achievement referred to changes in the student's score on the annual Virginia SOL math test for that student's grade level. Students received a score from 0-600 on the test (i.e., one test each year in grades 3-8) with 400 representing minimal proficiency and 500 or more representing advanced proficiency (VDOE, 2018).

Reading achievement. As used in this study, reading achievement referred to changes in the student's score on the annual Virginia SOL reading test for that student's grade level. Students received a score from 0-600 on the test (i.e., one test each year in grades 3-8) with 400 representing minimal proficiency and 500 or more representing advanced proficiency (VDOE, 2018).

Standards of Learning (SOL). The VDOE (2018) defined a core curriculum for students in each grade level for public schools in the state. At the end of each academic year, students took SOL achievement tests on the core subjects. While some grade levels included tests on other subjects in the core curriculum such as history or science, reading

and math appeared in the annual SOL tests for all grades (i.e., grades 3-8). These tests provided the math achievement and reading achievement measures used in this study to measure student achievement and learning (VDOE, 2018).

Chapter II: Review of the Literature

Student achievement served as the basis for Virginia school accreditation (VDOE, 2018). Researchers in prior studies attempted to identify interventions that improved student achievement outcomes. In this literature review, the researcher summarized the relevant studies within themes applicable to the research questions. The topics of this literature review concentrated around the issue of how other researchers attempted to access and improve educational achievement in students in public schools and the factors that affected student achievement with a focus on students in elementary and middle school (i.e., excluding high school and college or university-level research). With the current study focused on student achievement in reading and mathematics for children in grades 3-8, those concepts provided guidance in delving into studies to identify relevant literature. In particular, District A in this project was located in Region 7 in Virginia, where more than half the students were economically disadvantaged (VDOE, 2016). While this researcher did not directly measure the effect of economic issues on student achievement, other researchers identified economics as an important issue correlated with lower student achievement (Morrissey & Vinopal, 2018). Such economic challenges also played a role in student incentive to learn, an important aspect of the theoretical foundation of this study.

The researcher conducted this review by consulting peer-reviewed journals using online academic databases. Search terms included student achievement, reading achievement, math achievement, school leadership, teachers and student achievement, communication in schools including administration to school board, administration to community and parents, administration to teachers, and teacher professional development. These terms reflected the literature relevant to this study. In addition, the

literature review included information on the CIP as implemented in Virginia and on the VDOE SOL testing program. Several themes provided background and insight into the study such as academic achievement in both reading and math in elementary schools, teacher professional development, and student achievement. The researcher included the theme of teacher-student interactions of the Illeris (2015) social learning theory, as well as the theme of student incentives to learn, which when combined with content of instruction constituted the three aspects in that theory of learning.

Standards of Learning (SOL) Testing in Virginia

The Commonwealth of Virginia established SOL for all public schools in the state. According to the VDOE (2018), these standards established "minimum expectations for what students should know and be able to do at the end of each grade or course in English, mathematics, science, history/social science and other subjects" (para. 1). SOL achievement tests measured student achievement in each subject to determine learning and achievement. The VDOE used classroom teachers to develop the tests and to confirm that they were both accurate and fair. In addition, classroom teachers assisted the state Board of Education to determine appropriate proficiency standards.

The SOLs for each subject and each grade level were available for reference on the VDOE SOL website (VDOE, 2018). The blueprints for each SOL assessment included the category of learning, the SOL for that grade, and the number of passage-based test items (i.e., questions) about each standard of learning. As one example, the 2016-2017 grade 4 reading blueprint specified three learning skills: using word analysis strategies and word reference materials; demonstrating comprehension of fictional texts; and demonstrating comprehension of nonfiction texts (VDOE, 2018). Each of those learning categories referenced specific mandated SOL as defined in the

state's learning standards. The number of items (questions) in the computer adaptive test form for word analysis strategies was five, for comprehension of fiction was 12, and for comprehension of nonfiction was 1l (VDOE, 2018). Similarly, for the 2017 grade 4 math SOL test, the learning skills included number and number sense (9 questions); computation and estimation (9 questions); measurement and geometry (9 questions); and probability, statistics, patterns, functions, and algebra (8 questions) (VDOE, 2018).

Students took SOL tests in reading and mathematics using a computer adaptive test delivery platform. The VDOE used this platform because it customized the test based on how the student responded to test questions. Students who took reading and math tests in grades 3-8 took the test in this computer adaptive platform (VDOE, 2018). All students in grades 3-8 taking the math test took computer adaptive versions of the test starting in the 2016-2017 academic year. Prior to that, students in grades 3, 6, 7, and 8 took the computer adaptive version of the math test in the 2015-2016 academic year (VDOE, 2018). According to the VDOE, the use of computer adaptive testing both increased the engagement students had with the testing process and reduced security risks in the state-mandated testing system (VDOE, 2018).

In reading tests, the SOL test was a passage-based reading computer adaptive test. In comparison with a traditional paper-and-pencil test, the VDOE (2018) noted that some differences existed between the two test forms. In the computer adaptive test, students answered different numbers of questions and read different numbers of passages than in a paper-based test, but all students at a given grade level read the same number of passages and answered the same number of test items. The specific passages and questions about the passages differed from student to student based on responses to prior questions (VDOE, 2018). Students could not skip questions in the computer-based test as they

could in a paper-based exam, but they had as much time as needed to complete the entire test, though they had to complete the test in a single test day (VDOE, 2018). In addition, while students could back up and revisit questions on the passage they were currently working on, they could not back up and revisit questions or passages answered before the current passage (VDOE, 2018). The computer adaptive test began with medium-difficulty passages, then transitioned to either more- or less-difficult passages depending on how well the student responded to the first passage. Individual questions asked about each passage also varied in difficulty (VDOE, 2018).

The mathematics computer adaptive testing system related more to the traditional paper-and-pencil tests (VDOE, 2018). While the number of test items on a particular grade level test differed from the number on a paper-and-pencil test, all students at each grade level answered the same total number of questions, and the computer customized the specific test items for each student based on responses to prior questions (VDOE, 2018). As with the reading SOL tests, the math test had no time limits, but instructors asked students to complete the test within a single test day (VDOE, 2018). The mathematics test format required students to complete each question before moving on to the next question. The student could not skip questions or answer them out of order, as was possible in a paper-and-pencil test. Once the student answered a question, the student could not backtrack to a previously answered question (VDOE, 2018).

Comprehensive Instructional Program (CIP)

The CIP in Virginia consisted of a consortium of public schools in Virginia that provided annual detailed lesson plans for teachers to use in all tested subjects as part of the Virginia SOL testing program (CIP, 2018). These materials included information on classroom activities, lesson plans, and assessments of student progress that teachers could

use. YouTube videos were available for teachers under the TeacherTube educational channel online to assist teachers in transitioning to the CIP curriculum (CIP, 2018). A panel of teachers who demonstrated success in the classroom by the high scores their students achieved on the Virginia SOL tests created the materials for the CIP. The subjects included in the elementary and middle school CIP programs were reading (grades 3-8), math (grades 3-8), science (grades 5 and 8), history (grades 6-7), and algebra (grade 8).

Implementation of CIP in a school or district included training teachers and school administrators in the proper use of CIP instructional resources, implementation of the benchmarks defined by CIP, and use of CIP data and assessments throughout the academic year to improve student outcomes (CIP, 2018). Each region in Virginia that implemented CIP used a prescriptive approach that avoided insisting on identical instruction in all schools. This design allowed local teachers and administrators to adapt the CIP concepts to the needs of their students (CIP, 2018).

A key aspect of the CIP was the ongoing collaboration between teachers and administrators through a process of teacher meetings, curriculum team meeting, and school visits (VDOE, 2018). In addition, the CIP focused on the data from the classrooms, including district, school, and teacher-level reports to both identify successful approaches and outperforming trends and remove excuses for lower performance in the classroom (VDOE, 2018). The effect of the CIP in District A as a whole was to reduce the impact of poverty and minority or disadvantaged children from student achievement scores (VDOE, 2018).

CIP provided overall achievement test comparisons yearly that determined how well CIP benchmarks and assessments aligned with the Virginia SOL achievement tests

(CIP, 2018). Results from Region 7 between 2016 and 2017 showed a 0.38% improvement in reading achievement (from 81.58% to 81.96%) and a 0.34% improvement in math achievement (from 82.84% to 83.18%), the greatest improvement of any district that fully implemented CIP in the 2015-2016 school year in each of those subjects.

The goal of CIP was to help with continual improvement in teaching methods and materials to positively impact student achievement. To determine the effectiveness of CIP, each year the CIP panel compared student achievement scores on the CIP assessments to the SOL scores students took to determine how well the material presented in the CIP matched the SOL (CIP, 2018). In the results for 2018, the CIP correlated with the SOL for reading with Pearson's correlations of between 0.732 (grade 3) to 0.779 (grade 8). Similar correlations for math ranged from 0.773 (grade 4) to 0.863 (grade 6) (CIP, 2018). Comparing these results to the previous year (2017), the correlations dropped for one grade in reading (grade 8, dropping from 0.799 in 2017 to 0.779 in 2018) and for one grade in math (grade 4, dropping from 0.795 in 2017 to 0.773 in 2018). All other correlations improved in 2018 from 2017 by 0.73% (reading, grade 5) to 20.01% (math, grade 6).

While the measures of CIP improvement for student achievement were entirely positive for Region 7, that was less true for some other districts. The CIP website provided facts for changes in student achievement scores between 2016 and 2017 SOL testing dates (CIP, 2018). In particular, while all regions that fully implemented CIP in 2015-2016 school year saw reading achievement improvements, in math only Region 7 saw improved achievement between 2016 and 2017 SOL testing. All other regions that fully implemented CIP—as well as all regions that only partially implemented CIP—saw

math achievement decline by amounts ranging from –0.45% to –1.25%. This data provided the researcher the implication that other regions did not appropriately implement CIP, or CIP as a program was less successful in improving math achievement scores than proponents claimed.

The instructional activities, materials, and assessments made available to teachers was both extensive and available to teachers throughout the state via the online TeacherTube channel. For example, to help students understand irony, a short video from the movie *Frozen*, a 2013 3D American animated fantasy film produced by Walt Disney Animation Studios, illustrated an act of true love when Anna saved Elsa (CIP, 2018). Other videos explained math topics ranging from elementary arithmetic problems through advanced algebra (CIP, 2018). The TeacherTube channel provided teachers with instructional materials including videos, audios, documents, specific playlists, and other instructional material designed for CIP use (CIP, 2018).

Factors Affecting Reading Achievement

Bowman-Perrott and Lewis (2008) identified a number of factors that contributed to low reading achievement, particularly for African American and low-income students. Some of the factors Bowman-Perrott and Lewis (2008) identified included students moving frequently from district to district, lack of participation in early childhood programs, and the prevalence of underprepared, underqualified teachers for schools in poor urban and rural districts. Bowman-Perrott and Lewis (2008) compared standardized reading achievement scores for 4,135 African American students in grades 3-10 in an urban Midwestern school district with a total student population of more than 30,000 students; African American students constituted just over 20.5% of the total student population. In their study, Bowman-Perrott and Lewis (2008) targeted at-risk students

identified as disabled and African American. Elementary grade interventions focused on increasing the reading performance of students lagging behind grade level and increasing the number of advanced readers from the at-risk groups. Teachers received specialized literacy instruction training. Teachers also implemented a positive behavior support system to improve student discipline, a program that focused on three key principles: be responsible, be courteous, and be safe (Bowman-Perrott & Lewis, 2008). Bowman-Perrott and Lewis (2008) compared reading achievement and disciplinary referrals for African American students to that of other ethnic groups and found that African American students performed below all other ethnic groups other than Hispanic students in each of grades 3-10. In addition, African American students received more discipline referrals and received harsher penalties than other ethnic groups (Bowman-Perrott & Lewis, 2008). Based on these results, Bowman-Perrott and Lewis (2008) concluded that both Hispanic and African American students were at-risk students. Bowman-Perrot and Lewis (2008) also concluded that students in these at-risk groups required early and effective reading interventions as soon as teachers identified potential difficulties for these students.

Moon and Hofferth (2016) similarly studied factors that contributed to reading achievement in immigrant children. In particular, Moon and Hofferth (2016) probed the contribution that parental involvement had for both boys and girls in reading and math achievement in grades K-5. The researchers considered both the parental efforts and the child's own independent efforts and reading behaviors. The data Moon and Hofferth (2016) used came from the Early Childhood Longitudinal Study Kindergarten Class of 1998-1999 (ECLS-K). The study cohort included 21,260 kindergarten children nationwide in 1998 through the cohort that completed grade 8 in 2007. Moon and

Hofferth (2016) sampled the cohort's data for fall kindergarten, spring grade 1, spring grade 3, and spring grade 5, restricting the sample to 2,613 children from immigrant families. As a part of the ECLS-K, Moon and Hofferth (2016) used reading and mathematics achievement scores from nine activities (e.g., telling stories to the child, interaction with the child during play, helping with arts and crafts, doing puzzles or games with the child), as well as child-effort measured by parent's assessment of the frequency of the child's reading outside of school as a part of the data. Moon and Hofferth (2016) considered the family socioeconomic status and the family structure, including the number of siblings and used the mother's country of origin as determiners of the ethnic status of the child. Overall, Moon and Hofferth (2016) found that boys benefitted more than girls did from parental involvement. Girls showed no improvement in achievement scores based on parental involvement. Both boys and girls showed significant reading achievement improvement with more reading activity at home in early grades (grade 3 or lower). Only boys benefitted from improvements in math achievement scores for grades 3-5. Higher socioeconomic status had a greater impact on girls' reading and mathematics achievement scores for grades 1-3 than it had for boys (Moon & Hofferth, 2016). Moon and Hofferth (2016) found that a two-parent family positively affected boys' achievement scores but was less important for girls.

Hoy, Tarter, and Woolfolk Hoy (2006) conducted a similar study on student achievement and identified three factors that directly influenced student achievement beyond individual innate capacity and socioeconomic status: collective teacher efficacy, faculty trust in clients (i.e., parents and students), and an academic emphasis on student achievement, no matter what socioeconomic status of the student. Hoy et al. (2006) termed the combination of these three factors academic optimism. The researchers

concluded the three factors had reciprocal causal relationships among the factors that affected student achievement. In Hoy et al.'s (2006) framework, those three factors, combined with socioeconomic status, previous student achievement success, and the degree of urbanicity (e.g., population density, a measure from state statistical data) provided a model of student achievement. Of these, urbanicity was the least meaningful, with greater population density associated with lower student achievement (Hoy et al., 2006). Overall, Hoy et al. (2006) found that their tri-fold construct of academic optimism was second only to prior student achievement in its impact on predictions of student achievement.

Bevel and Mitchell (2012) tested Hoy et al.'s (2006) academic optimism as a factor in grade 5 reading achievement across 29 Alabama elementary schools using the School Academic Optimism Scale, the Collective Teacher Efficacy Scale, the Faculty Trust in Students and Parents Scale, and the Academic Emphasis Scale. These latter three instruments had previously measured reliability ranging from 0.83 to 0.94. After measuring the three components of academic optimism, Bevel and Mitchell (2012) found that the components created academic optimism as a second-order construct and that academic optimism in turn correlated strongly ($r = 0.78$, $p < 0.001$) with reading achievement. These results indicated that Hoy et al.'s (2006) academic optimism construct was predictive of reading achievement and that it accounted for about 18% of the variance in student reading achievement scores. One limitation of Bevel and Mitchell's (2012) study was it was a correlational study that lacked a way of identifying how to change the academic optimism within a school; thus, it lacked explicit guidance on improving student achievement.

Harlaar, Deater-Deckard, Thompson, DeThorne, and Petrill (2011) conducted a study of 10-11 year-old twins to study whether an association existed between reading achievement and independent (i.e., non-school-based) reading. The goal of the study was to determine if programs that challenged elementary school children to read 1,000,000 words (or similar reading challenges) as a way of developing good reading habits had a scientific basis or simply was intuitive guesses. Using 436 pairs of 10 year-old identical, same-sex twins, the researchers combined genetic marker measurements with reading performance measures (i.e., the Word Identification and Passage Comprehension subtest from the Woodcock Reading Mastery Test) and family and child reports of how often each child read at home for pleasure using a five-point Likert scale (1=almost never; 5=more than 3 times a day). The child participants filled out the Motivation for Reading Questionnaire (MRQ), with each twin filling out the survey independently in separate rooms. The MRQ included a measure of participants' self-efficacy as well as the participants' willingness to take on challenging material. Harlaar et al. (2011) conducted a follow-up set of measures at age 11. Harlaar et al. (2011) found that 10 year-olds' reading achievement accounted for about 8% of the variance in independent reading at age 11. Independent reading at age 10 did not predict reading achievement at age 11. Harlaar et al. (2011) found that individual reading achievement and independent reading were both inheritable (i.e., associated with genetics) to a significant degree. When the researchers extended the study to include non-identical twins (i.e., dizygotic twins), the correlation between twins' behaviors was no stronger than it would have been for randomly paired children. Key limitations of this study were that it relied on parent and child reports of independent reading and that it had no measurement of the types or difficulties of the books the twins read for pleasure.

Passage of NCLB also encouraged parental involvement in the schools (Park, Stone, & Holloway, 2017). Park et al. (2017) tested the effect of parental involvement in the schools using information from the Early Childhood Longitudinal Study, a nationwide study of student academic and behavioral development following one cohort of K-5 students in public schools. Park et al. (2017) measured parental participation in schools at the level of support for general school improvements, support for the parents' own children, and social networking among parents. Of these factors, schools with a high degree of parental involvement in general school support and improvements and parental networking had students with higher national testing scores on both reading and math and were more likely to have good learning environments (Park et al., 2017). Park et al. (2017) found that differences did exist based on socioeconomic factors. Park et al. (2017) found in lower socioeconomic status schools, the greater impact occurred from parents aggregately supporting their own children's schools and networking, while higher socioeconomic status schools experienced greater impact from more general public school support and parental networking. Once again, Park et al. (2017) used a correlational study that lacked a means of changing the factors identified as affecting student achievement.

Interventions to Improve Reading Achievement

Bakosh, Snow, Tobias, Houlihan, and Barbosa-Leiker (2015) reported results from a study of audio-guided mindful awareness training as an intervention to improve student grades at the elementary school level, including improving math and reading grades. The Bakosh et al. (2015) study was quasi-experimental, using a 10-minute-per-day automated mindfulness training program to determine if student grades improved compared to a similar control group. The student participants were in

grade 3 at two schools (Bakosh et al., 2015). Ninety-three students in four grade 3 classrooms, with two classes from each of two schools, participated in the intervention group, and 98 students in four other grade 3 classrooms from the same schools constituted the control group (Bakosh et al., 2015). Across 8 weeks, the intervention group listened each day to 10-minute mindfulness training recordings from 35 different recordings; the content of the recordings focused on stress reduction and instruction on how to sit, why to practice mindfulness, and what the mindfulness practice could help (Bakosh et al., 2015). The mindfulness awareness program included teaching awareness of senses, thoughts, and emotions, as well as provided breathing, relaxation, and silence in the recordings. Bakosh et al. (2015) measured student grades at the end of the 8-week quarter, daily measures of overall (not individual) classroom behavior, and the impact the program had on the classroom operation. Bakosh et al. (2015) found that the daily mindfulness training was a significant predictor of improved classroom grades in science and reading, as well as producing noticeably improved classroom behavior. Bakosh et al. (2015) also noted that the intervention was explicitly teacher-independent using only pre-recorded mindfulness training. Although Bakosh et al. (2015) found that student grades did improve in the intervention group compared to the control group, the effect of the 8-week program on grades was modest, leaving unknown whether a longer or more intensive mindfulness training would produce greater achievement improvements.

Lee et al. (2017) reported on an intervention study using an after-school EdVenture program aimed at underserved, ethnic minority, and low-income students in northern California. The study included 28 elementary children in grades 1-6, with 75% female, 79% ethnic Latino, 7% African American, 7% Russian/Ukrainian, and 7% Pakistani or Indian. The measure Lee et al. (2017) used in the study was a 15

question *The Me and My World Survey*, a measure of the students' developmental assets rated on a five-point Likert scale. Lee et al. (2017) also included students' school progress reports, which contained proficiency assessments on language arts, reading, writing, math, science/health, history/social studies, homework, classroom, and quality of work. These proficiency reports used a four-point Likert scale: below basic, basic, proficient, and advanced (Lee et al., 2017). Teachers also rated students' homework completion, classwork completion, and quality of work on a poor-good-excellent scale (Lee et al., 2017). Lee et al. (2017) reported that the EdVenture after-school program demonstrated significant improvement in students' self-efficacy. The EdVenture program was not a significant predictor of student reading achievement (Lee et al., 2017). Lee et al. (2017) speculated that teacher expectations and perceptions had a stronger influence on student reading achievement.

Independent reading was part of the Virginia SOL for elementary students, including specifying that teachers should provide opportunities for independent reading (VDOE, 2018). Fisher and Frey (2018) identified evidence-based interventions that encouraged elementary students to read outside the classroom. Fisher and Frey (2018) first identified four factors that encouraged students to read more: greater access to books outside the classroom, giving students greater choice of what they read, encouraging classroom discussion of texts read, and talks by peers and trusted adults about books students might enjoy. Fisher and Frey (2018) then created a Reading Volume Program (RVP) incorporating those four factors and arranged for six elementary schools to test the program. The six schools had more than 450 students each in K-6. At least 50% of the students in each school qualified for reduced- or free-lunch programs. More than 36% of the 3,846 students in the schools were new learners of English, 10.5% had identified

learning disabilities, and 2,384 were Latino. Odd-numbered grades (i.e., grades 1, 3, and 5) participated in the RVP intervention, while even-numbered grades (i.e., grades 2, 4, and 6) did not. All but nine of the 53 teachers in the odd-numbered grades chose to participate in the RVP program (Fisher & Frey, 2018). The RVP program lasted 12 weeks, beginning three weeks into the school year. After the 12-week program, teachers provided specific data points that demonstrated the effect of the program on the students (Fisher & Frey, 2018):

- Teachers reported that students checked out 9% more books from the school library than the same students had done the previous year;
- Teachers reported 4% higher benchmark writing scores compared to other schools in the district;
- Teachers reported 2% higher fluency rates compared to the students' previous reading records or with other schools in the same district; and
- Teachers reported more students and more parents anecdotally claiming students read more books.

Fisher and Frey (2018) also noted that when the teachers in the even-numbered grades heard of the successes the odd-grade students achieved, the teachers in grades 2, 4, and 6 started implementing similar strategies in the classroom. Fisher and Frey (2018) reported that one teacher of an even-numbered grade threatened to go to the union representative if denied access to the RVP program training and additional books for students in her classroom. Fisher and Frey (2018) had specific classroom recommendations but noted in particular that deep reading of classroom texts did not sacrifice broad reading out of the

classroom, nor could broad reading sacrifice deep classroom text reading; both were important.

Faust and Kandelshine-Waldman (2011) investigated three different reading instruction approaches (i.e., phonic/synthetic, whole language/global, and eclectic) to determine the effect on low-achieving and normally achieving readers in elementary school (specifically, grades 1-6). The phonics approach was a bottom-up process focused on learning sounds represented by letters and letter combinations (Faust & Kandelshine-Waldman, 2011). Faust and Kandelshine-Waldman (2011) described the whole language/global approach as a top-down process that emphasized extracting meaning of words from context. The eclectic method incorporated both types of approaches, teaching both bottom-up processes while also focusing on textual extraction based on context (Faust & Kandelshine-Waldman, 2011). The goal of the study was to determine if low-achieving readers who tend to rely on top-down processes to recognize words would achieve better reading comprehension if taught using the whole-language approach (Faust & Kandelshine-Waldman, 2011).

Faust and Kandelshine-Waldman (2011) conducted two experiments using 1,505 grades 1-6 students in four elementary schools. Of these students, 451 received phonics-based reading instruction, 492 received whole language/global reading instruction, and 562 received eclectic instruction, with the instructional approach varying by the school (Faust & Kandelshine-Waldman, 2011). The instrument used was a skill-appropriate text the participant read at normal reading speed, while circling every instance of a particular letter (Faust & Kandelshine-Waldman, 2011). After reading, the students answered three comprehension questions with no text provided to confirm their understanding of the story (Faust & Kandelshine-Waldman, 2011). Faust and

Kandelshine-Waldman (2011) found the instructional approach did not affect the number of correct comprehension questions the participants answered. The participants' teachers assessed reading ability based on a six-point Likert scale (1 = poor; 6 = excellent) (Faust & Kandelshine-Waldman, 2011). Faust and Kandelshine-Waldman (2011) classified participants who scored in the range from 1 to 3 as low-achieving, a classification assigned to 21% of the participants. Faust and Kandelshine-Waldman (2011) identified differences in performance on the measuring tasks among the three reading instruction approaches but did not find support for the hypothesis that the whole language/global reading instruction approach would narrow the gap for low-achieving readers. With that said, Faust and Kandelshine-Waldman (2011) found the whole language/global approach resulted in overall higher omissions from all readers taught with that approach across all word types, implying that such an approach might not facilitate top-down approaches that were related to reading proficiency.

Kirnan, Siminerio, and Wong (2016) investigated the impact of therapy dogs on student reading achievement. In this mixed methods study, Kirnan et al. (2016) analyzed reading test scores of 169 students in grade K-4. The therapy dogs visited intervention classrooms at least once a week for about an hour for the duration of the school year (Kirnan et al., 2016). During this visit, students read to the dog in groups of four to six students, based on student reading level (Kirnan et al., 2016). The classes also included a writing component where grade 4 students created a newspaper with dog-themed stories (Kirnan et al., 2016). Grade 2, 3, and 5 students kept written journals they could illustrate (Kirnan et al., 2016). Grades K-1 students had dogs more fully incorporated into the language arts curriculum, with reading, writing, and vocabulary games with dog themes (Kirnan et al., 2016). Students with severe allergies participated remotely via

iPad; students afraid of dogs began at the periphery of the reading groups, but by the end of the program these students fully participated in the program, reading to, petting, and working with the dog (Kirnan et al., 2016). In addition to reading achievement scores for participants from standard achievement tests, Kirnan et al. (2016) conducted semi-structured interviews of the dog owners and the teachers to note observations about the sessions with the students. Kirnan et al. (2016) found that reading skills only varied at a statistically significant level for grade K students. Implementation of the program was school-wide, but the total number of students participating at each grade level caused difficulty in establishing statistical significance (Kirnan et al., 2016). Kirnan et al. (2016) hypothesized that the incorporation of the dog in the broader language arts programs in grades K-1resulted in a stronger effect than in later grades. The qualitative data Kirnan et al. (2016) collected indicated increased confidence, greater self-esteem, and increased interest in reading by the students. These observations included ones from teachers who initially expressed great skepticism toward the program but who later also observed the self-esteem, confidence, and reading interest improvements (Kirnan et al., 2016)

Mokhtari, Thoma, and Edwards (2009) reported on a suburban school in the Midwest that used standardized test data to direct the creation of a professional development program for teachers intended to improve reading instruction in the school. According to Mokhtari et al. (2009), the program included explicit goals of increasing reading performance of all students in one Midwestern elementary school. After the program implementation, approximately 90% or more students in grade K-5 tested as either proficient or advanced in the state reading achievement tests (Mokhtari et al., 2009). This constituted an improvement of 5% to 27%, depending on grade level, from fall testing to spring testing after implementation of the reading achievement program

(Mokhtari et al., 2009). All but one grade had 92%-96% of the students achieving proficient or advanced proficient levels, and the single outlier (grade 2) had 88% achieving that level of reading success (Mokhtari et al., 2009). Mokhtari et al. (2009) noted three elements to this successful program: employing reading professionals and teachers credentialed in their area of expertise; establishing a professional learning community that focused on improving student achievement in reading; and establishing programs that supported the student, the teachers, and overall school performance. The study was an interventional study with pre-/post-test results that identified specific strategies the program used to improve students' reading achievement (Mokhtari et al., 2009). According to Mokhtari et al. (2009), limitations to the study were that it involved only a single elementary school with 638 students (grades K-5) and a culturally homogeneous population (96.8% Caucasian and 7.8% economically disadvantaged) that participated in the free-lunch program). Mokhtari et al. (2009) also had only two years of project implementation to report.

Another example of all-school commitment to improving literacy in elementary students came from Fisher and Frey (2007), who reported on a literacy program in a heavily Hispanic urban school in San Diego, California. The school location was in the highest crime-rate area of that city. According to Fisher and Frey (2007), the school implemented a school-wide, all-grades focus on improving student literacy that created a coordinated educational program across all teachers and all grades. When compared to other schools in the city in nearby areas, this school generated higher academic performance than the other schools, raising the academic achievement rating generated from California achievement testing from a score of 455 to one of 746 over a six-year period from 1999 through 2005 and increasing the percentage of students labeled

proficient or advanced on the state tests from approximately 10-15% to approximately 36% in that period (Fisher & Frey, 2007). Fisher and Frey (2007) noted that the program centered on foundational principles: learning was a social activity; conversations were important to learning; integrating reading, writing, and oral instruction was essential; and learning required a gradual increase in the responsibilities of the student. As with Mokhtari et al. (2009), Fisher and Frey's (2007) was a single-school intervention study that used a pre-/post-test approach to determine the effect of the intervention. Also as with Mokhtari et al. (2009), Fisher and Frey (2007) had no control group, although the San Diego study did cover the years from 1999 through 2005 rather than only two years as with Mokhtari et al. (2009).

Factors Affecting Math Achievement in Elementary School

An important aspect of Illeris's (2015) social learning theory was the interactions teachers have with students. The teachers' own beliefs in their self-efficacy and their attitudes to the subjects they teach mediate those interactions with students (Illeris, 2015). The National Science Foundation coined the acronym STEM to refer to science, technology, engineering, and mathematics (Madden, Beyers, & O'Brien, 2016). This term originated to mean a more integrative concept in which two or more of the STEM fields combined in the classroom to improve student understanding (Madden et al., 2016). The attitudes and self-efficacy beliefs of teachers, particularly in elementary grades, influenced their attitudes toward teaching math and science subjects. Madden et al. (2016) noted that teachers who admitted to an affinity to math and science were more likely to use innovative teaching methods in these subjects than the teachers who claimed to dislike those subjects.

Burns et al. (2015) noted that achievement in math in elementary school was an important predictor of overall school achievement in middle school and high school and that math competency had multiple dimensions. Researchers indicated that factors—including the specific math teacher, student gender, parental attitudes and their math anxiety levels—and student self-efficacy in math all had correlations to student math achievement in elementary school levels (Crosnoe et al., 2010; Gottfried & Graves, 2014; Soni & Kumari, 2015; Weidinger, Steinmayr, & Spinath, 2018). Researchers also reported results on various interventions to improve math scores. Burns et al. (2015) found that individually tailored interventions based on student deficiencies were important to improve math achievement scores, while Carr, Taasoobshirazi, Stroud, and Royer (2011) tested computer-based computational fluency instruction interventions. Other researchers reported on a variety of classroom interventions to improve math scores (Heatly, Bachman, & Votruba-Drzal, 2015; Ing et al., 2015). The researcher began this section with individual information about the studies that identified factors affecting math achievement followed by information about those studies that tested various interventions to improve math achievement scores.

Xu and Jang (2017) investigated the effect of student math self-efficacy and non-school use of technology-related activities (e.g., internet usage, video games, television viewing) on grade 6 students' math achievement scores. Xu and Jang (2017) conducted the study in Ontario, Canada, with a sample of 26,767 English monolingual grade 6 students who took Ontario's standard Education Quality and Accountability Office math achievement test in 2013. As part of this test, students also filled in a background questionnaire detailing their use of technology outside of school and their math self-efficacy (Xu & Jang, 2017). Six questions in the background questionnaire

addressed student math self-efficacy (e.g., questions such as *How much do you agree with the statement, 'I like math'?*; *How much do you agree with the statement, 'I am good at math'?*; etc.). In addition, six background technology usage questions asked about the number of hours per day spent watching television, using the internet, and playing video games before and after school (Xu & Jang, 2017). Xu and Jang (2017) found similar results to prior studies in that greater use of technology was associated with lower student math achievement, but the researchers also found that when the students had positive math self-efficacy there was a positive mediating effect on math achievement (Xu & Jang, 2017).

Crosnoe et al. (2010) conducted a longitudinal study investigating the effects of different instructional styles for students with low, medium, and high math aptitude to determine what style would most improve student achievement. The teaching styles included focusing on basic math skills, providing students with higher-level inference-based training, and providing socioeconomic support for the students (Crosnoe et al., 2010). According to Crosnoe et al. (2010), inference-based training focused on word problems in which students had to infer solutions or expectations about the situation based on the information they had about the situation. Crosnoe et al. (2010) found children at all levels of math skills responded most positively to inference-based instruction techniques; students with the least math skills initially received inference-based training and closed the gap in math skills. Crosnoe et al. (2010) noted an exception occurred if there was a conflict in the relationship between student and teacher. In that event, Crosnoe et al. (2010) found no narrowing of the gap between least skilled and most skilled students. Crosnoe et al. (2010) thus concluded that the math teachers'

relationship with the students combined with specific instructional styles was most effective at raising elementary students' math achievement.

In comparing math achievement in girls and boys, Gottfried and Graves (2014) noted that researchers targeted gender-preferenced classrooms (i.e., concentrating more of one gender) and found improved student outcomes. The researchers found that classrooms with more girls tended to improve outcomes for all students, but those results varied by grade level, with more girls in classrooms associated with better student achievement in early elementary grades, but the effect diminished by grade 3 (Gottfried & Graves, 2014). To determine the validity of such results, Gottfried and Graves (2014) investigated the effect of having more or fewer girls than boys on student achievement with a specific focus on subject-by-subject results rather than more general measures of student achievement. While Gottfried and Graves's (2014) results confirmed more girls are better in student achievement for most subjects, for math in particular Gottfried and Graves (2014) found that gender segregation improved student performance for both boys and girls. While differences arose as soon as grade 1, Gottfried and Graves (2014) found that having a higher percentage of girls in the classroom tended to result in greater student achievement in most subjects for both boys and girls. Specifically for math, Gottfried and Graves (2014) found that girls performed better when there were 30% or fewer boys in the classroom, and grade 3 girls performed better in math when placed in an all-girls classroom, while boys' performance showed no difference in performance whether girls were in the classroom or not.

Weidinger et al. (2018) studied changes in self-belief of competence in elementary school children in Germany between grade 2 (ending at approximately age 8) and grade 4 (ending at approximately age 10). Weidinger et al.'s (2018) goal was to

understand how changes in children's beliefs in their abilities linked to their achievement in math using standardized tests for that measure and student math grades. Weidinger et al. (2018) found that children tended to become more negative in their beliefs of their own competency between grades 2-4. Because negative or lower belief in competency was often associated with lower levels of effort on the part of the student, Weidinger et al. (2018) suggested that maintaining math competency beliefs starting in very early grades were crucial to increasing math competency in later grades.

Soni and Kumari (2015) found that math anxiety in parents tended to be a precursor to their children experiencing math anxiety. Similarly, the attitudes parents held toward math also presaged the attitudes their children displayed toward math (Soni & Kumari, 2015). Ramirez, Gunderson, Levine, and Beilock (2013) found that even students in grades 1-2 in a large urban school district experienced math anxiety. Ramirez et al. (2013) found that children were strongly reliant on working memory solution strategies (i.e., memorizing specific solution methods) for problem solving experienced difficulty when they also had math anxiety. Ramirez et al. (2013) urged an early focus on identifying elementary students with math anxiety and treating it early because students with strong working memories would have greater potential for success in math. Hirvonen, Tolvanen, Aunola, and Nurmi (2012) studied math performance in elementary students between grade K and grade 4. Hirvonen et al. (2012) measured task avoidance behavior in the students as rated by their teachers and compared that to the students' growth in math performance. Hirvonen et al. (2012) found that as task avoidance behavior increased, growth in math performance also decreased.

The above researchers provided important clues as to the causes and solutions to lower math achievement in elementary students. One theme that emerged from these

studies in math achievement was the importance of the teachers' instructional styles and the relationships between teachers and their students. Poor student-teacher relationships resulted in lower academic achievement, but Crosnoe et al. (2010) also found that students with the poorest math skills benefitted from using an inferential teaching approach. This was in line with Ramirez et al. (2013) who noted that even students with substantial levels of math anxiety benefitted from learning to use no-memorization-based math solution processes and relied on more inferencing and analysis approaches. Gottfried and Graves (2014) noted that the gender balance of the math classroom affected math achievement, with an all-girls classroom being best for both boys and girls starting in grade 3. Weidinger et al. (2018) found a critical period in students' self-efficacy with respect to math occurred in the period from grades 2-4, while Soni and Kumari (2015) noted that parents frequently transmitted their own math anxiety to their children, negatively affecting their children's math achievement. This evidence supported the inferences that math instruction, even in very early elementary grades, requires sensitivity to the children's beliefs and attitudes, careful attention to the gender balance in the classroom, attention to how teachers present math as a subject (e.g., not in a rote style), and the quality of the relationships between teachers and students.

Intervention Studies on Math Achievement

Burns et al. (2015) conducted small-scale interventions based on measured deficiencies in early elementary school students. The researchers tested student participants for three skill clusters: fluency with whole numbers, grasping basic arithmetic operations (i.e., addition, subtraction, multiplication, and district), and logical problem solving (Burns et al., 2015). Burns et al. (2015) designed interventions to address identified deficiencies and created a conceptual intervention and a procedural

intervention designed to address those deficiencies and tailored for the students. The specifics of the interventions varied based on the needs of the student. Conceptual interventions followed a model-lead-test format in which the tutor demonstrated the solution process and then supported the student until the student could solve the problems independently (Burns et al., 2015). Procedural interventions used incremental rehearsals in which tutors read math facts (such as *2 + 3 = ?*) identified as unknown to that student. The tutor read the fact to the student, gave the student the correct answer orally, then asked the student to repeat the answer. After the repetitive practice, students answered a set of nine questions, with four previously known facts and five previously unknown facts. When the student answered all nine questions correctly, the tutor removed one known fact, added a new unknown fact, and repeated the rehearsal of the facts (Burns et al., 2015). Burns et al. (2015) provided an example of the success achievable using individually tailored interventions, with each of the participating children demonstrating achievement gains when using as few as four intervention sessions.

Uribe-Flórez and Wilkins (2016) probed the use of manipulative objects to teach elementary math and the relationship of the use of manipulative objects to math achievement and math learning. In Uribe-Flórez and Wilkins's (2016) study, the manipulative objects included any of a variety of objects students could handle, including geometric shapes, base-ten blocks, and pattern blocks. Uribe-Flórez and Wilkins (2016) measured math achievement using the ECLS-K, a longitudinal study of students who were in grade K in the 1998-1999 school year and followed that cohort through grade 8. Uribe-Flórez and Wilkins (2016) used data from a baseline kindergarten measurement (1999), grade 1 spring (2000), grade 3 spring (2002), and grade 5 spring measurements (2004). The study included 10,673 students, of which 57.6% were Caucasians, 16%

African American, 18.8% Hispanic, 2.8% Asian, and the remainder from either no identified race or other ethnicities (Uribe-Flórez & Wilkins, 2016). To measure student math learning (as opposed to achievement test scores), Uribe-Flórez and Wilkins (2016) used item response theory techniques to represent the math knowledge each student had at the various grade levels. The independent variable in this study was the use of manipulative objects in the classroom, which Uribe-Flórez and Wilkins (2016) measured by asking teachers how often grade-appropriate manipulative objects appeared as part of the classroom training on a class group basis for grades K-1 and on an individual basis for grades 3-5. Uribe-Flórez and Wilkins (2016) found that the use of manipulative objects decreased in later grades and that the use of the manipulative objects had no correlation to student achievement scores in math. Despite the lack of impact on student achievement scores in math, Uribe-Flórez and Wilkins (2016) found a significant association between math learning and the use of manipulative objects in class; more specifically, students who used manipulative objects between two and eight times a month learned at a faster rate than those who rarely or never used manipulatives. In addition, Uribe-Flórez and Wilkins (2016) found that those students who used manipulative objects nearly every day learned math faster than those who used the objects only once or twice a week.

Researchers also tested different types of math instruction in different grade levels, based on the presumption that a child is growing and maturity might alter the most effective instructional style. Heatly et al. (2015) investigated how various instructional modes affected growth in math achievement at grade levels from K-5. Heatly et al. (2015) compared conceptual instruction (i.e., focused on problem solving and reasoning skills) and procedural instruction (i.e., focused on specific calculation processes such as

counting, addition, subtraction, and chalkboard activities). The researchers found that teachers of kindergarten who spent more time on procedural instruction had students who achieved greater gains in math test scores, while teachers of grade 5 students who spent more time on conceptual instruction saw their students gain more in math achievement scores (Heatly et al., 2015).

Ing et al. (2015) investigated teacher instructional style in effective math classes in elementary school. The researchers noted specific teacher practices and linked them to student participation and student achievement (Ing et al., 2015). Ing et al. (2015) considered whether student participation in class activities was a mediating factor in how teacher practices affected student achievement. The researchers measured student participation by observing the degree and completeness of students explaining a problem-solving solution and by how students engaged with other students about math problems (Ing et al., 2015). The researchers also measured teacher instructional style by how often teachers persuaded students to share their thinking processes about a problem and by how frequently teachers encouraged students to engage with other students in class discussions (Ing et al., 2015). Ing et al. (2015) found that greater teacher support for student participation was a positive predictor of student math achievement but had only an indirect predictor of student achievement, while teacher encouragement increased student participation, which was the direct predictor of student math achievement.

Carr et al. (2011) identified student deficiencies and created a computer-based tutoring program designed to address those deficiencies and improve standardized test scores in elementary students. Carr et al. (2011) identified computational fluency (i.e., the speed with which students could perform computational tasks, specifically single-digit and double-digit arithmetic problems, as important in overall student

achievement scores on standardized tests. Carr et al. (2011) related such fluency to the generation of a sense of numbers that would enable students to solve more complex math problems. Carr et al. (2011) noted that as soon as grade 2, children displayed differences in problem solving strategies according to gender, with girls more likely to use finger-counting strategies, while boys used more cognitive approaches to deconstruct a problem for faster computation. Carr et al. (2011) tested a variety of computer-based instructional programs and found those that combined teaching computational fluency and cognitive strategies produced greater math achievement than programs focused on only one approach. Furthermore, Carr et al. (2011) found that the use of such a combined instructional approach eliminated any gender gap in math achievement for the grade 2 children.

School Configuration

According to Herman (2004), the school grade setting during the 1800s was predominantly grade K-8 and was largely one room across all grades. Limited resources and economic hurdles attributed to the reason. The introduction of elementary schools and middle schools only came to fruition during 1900s when educational programs and instructional materials started to expand, as was observed by Juvonen, Le, Kaganoff, Augustine, and Constant (2004). Enrollment in schools was low before school configurations largely due to the traditional concepts, which even left a number of students out of schools. Enrollment in schools increased, especially after World War I when school configuration became more recognized. Further reorganization focused on K-8 school configuration and proportionate alignment of student cognition to the learning materials needed to meet the needs of students. Due to this mass school configuration from grades K-8, the establishment of junior high schools (especially grades 7-9) were

customary to prepare adolescents for the rigorous high school years and improve their performance in all grade levels. Researchers reported a downfall in student performance, especially at high school levels, and attributed the downfall to improper instructional materials used to enhance students' cognitive understanding (Juvonen et al., 2004).

School configuration included structure of the school—such as grades K-4, 5-8, or K-8—and how students transition from one level or school to the next in the education system. Primary education lasted for six years and children who had attained the age of five were eligible to enter kindergarten. In addition, national primary schools were often public schools while at the same time they were not co-educational. According to the CIP program, reading and math dominated most studies within the primary grades (grades 1-8) as well as weekly class periods while other subjects cumulatively averaged in weekly class periods. These were subjects such as science, social studies, and physical education. This resulted in the impression that primary schools targeted reading and math. Lower primary grades were often between grades K-4. Grades K-4 emphasized learning reading skills and mathematics as a means of improving students' learning outcomes. The middle grades were comprised of grades 5-8 and were schools where students started to prepare for their transition to high school. After completion of primary education, students had the opportunity to attend general, specialized, or technical secondary institutes where they acquired training in specific fields. It was important to improve learning outcomes of students within the primary levels such that they acquired the knowledge and skills to engage in other trainings in line with their career of choice.

The establishment of middle schools in District A in the early 1970s was a means of reducing congestions in other configurations within the schools and a way of meeting

the learning requirements of students at the adolescent stage. It was important to note that learning requirements for early-aged students (such as those between grades K-4) were distinct from the learning requirements for students who were early adolescents (such as those in grades 5-8). District A established middle schools to meet this requirement as well as respond to studies for students who were at their early adolescent development stage. In this case, middle schools supported providing educational studies for the whole child as well as meeting children's psychological developmental needs. Smaller towns within District A did not build grades 5-8 schools due to smaller enrollment numbers. District A established K-8 schools in those towns and communities that did not have student enrollment numbers to justify both a grades K-4 school and a grades 5-8 school.

Wren (2003) reported that transitioning from the early maturing environment such as the elementary school to middle school presented a challenge to schools in America because there was a high level of absenteeism among children as well as instructional and discipline problems that affected children's academic achievement. For instance, Alspaugh (1998) realized that there were significant losses of achievement among students in grade 6 in accordance to poor transitioning from elementary to middle school. Additionally, school specialists articulated that schools were social systems characterized by independent parts of clearly defined populations (Edwards, 2011), and school configuration defined the district of schools into sections with each section having its own populations with branded characteristics. This revealed how school configuration helped designate populations and environmental structures that determined what instruction to provide at which grade (Edwards, 2011). This was why instructional materials for grades 3-8 were distinct and different as they addressed multiple

populations and individuals of diverse characteristics. Similarly, the CIP also provided instructional materials that were different for each class and grade within schools in District A.

Researchers focused efforts on the need to improve learning outcomes among students from grades K-8 in American schools to determine how improvements through these grades enhanced the effectiveness of the learning outcomes within these grades. Consequently, research focused on K-8 in American schools to understand the importance to children and academic achievement. This brought about a subsequent configuration in schools among districts and anticipated a positive effect on student learning outcomes. Researchers conducted a number of grade level configurations in schools, the earliest being the schools such as Baltimore Peninsula and Cincinnati as described by Look (2002) and Yecke (2006). Despite the grade level configurations in these schools and student performance differences, there remained a literature gap of whether or not there was a change in student performance; however, other school and educational specialists shared their thoughts that school configuration had an effect on student performance especially when it came to transition from primary schools to upper levels. For this reason, many intellectuals believed that the success of K-8 models were still sparse, leading to varying conclusions (Look, 2002), which meant that the success of K-8 models were based on student requirements and learning capacities. Essentially, as students progressed, they developed psychological progress in their cognitive and social development; hence, the instructional materials to use in instructing these children varied in accordance to their psychological expansions. This enabled student to transition smoothly from one level to the next with limited complications as compared to the use of

instructional materials that did not conform to their cognitive development as they transitioned to next grades in their learning progress.

Observation provided that the effectiveness of instructional learning materials from grades K-8 in some schools within America were not effective due to the failure of most materials not conforming to the students' psychological constructs and needs. For instance, Hoy and Miskel (2005) observed that most K-8 schools in Georgia were still not effective despite school configuration. Attribution given to the lack of aligning student cognitive needs with the instructional materials at each grade (grades K-8) resulted. Despite this menace, it was still articulated that the failure at hand was yet to be probed but with no significant reason.

In line with the school configuration chronology, it was important to note that it was not the configuration system that had a problem but the alignment to students' cognitive and psychological capabilities to the learning materials that was a danger. In essence, students needed a reasonable amount of course and academic work, not more than they could handle, and at the same time, giving students lower than their cognitive capability declined their performance (Juvonen et al., 2004). This meant there was a need to develop a program that would enhance student learning capabilities and understanding in accordance to students' cognitive developments from one grade to the next, thus the introduction of CIP in District A, which focused on student reading and math cognition from grades 3-8. Understandably, those two variables had an effect on improving student learning outcomes from one grade to the next if appropriately utilized in schools within America.

Summary of Literature Review

In the literature review, this researcher covered aspects of student achievement in elementary reading and math from several perspectives. The results from this review indicated that a number of factors influenced student achievement. Community and family factors such as poverty directly affected student achievement, with neighborhood and community-level poverty having greater impact than individual family poverty. In reading achievement, factors such as teacher efficacy, trust in parents and students, and a strong school academic emphasis on student achievement were powerful indicators of success in reading achievement growth. In addition, a twin study found that reading achievement had a significant genetic factor. Other factors found in the literature that positively influenced student reading achievement included strong parental support for the school, parental expectations, confidence in their students, and effective teacher development programs. Innovative instructional programs such as a Fisher and Frey's (2018) RVP program to increase student independent reading and Kirnan et al.'s (2016) use of therapy dogs for reading achievement were also successes.

Researchers found math anxiety present as early as kindergarten and derived at least in part from parental math anxiety and parental attitudes toward the subject. Specific instructional styles contributed to greater math achievement, with the appropriate instructional style varying by grade level, but emphasis on inferencing and problem-solving skills were more effective than rote or memorized solution procedures. Class gender distribution played a part in math achievement as well. Girls performed better in math classes consisting only of girls beginning in grade 3, though in other subjects, having mixed genders (but weighted toward more girls) was the best mix for both boys and girls. Student self-efficacy and self-confidence were important in math

achievement. When students had high levels of math self-efficacy, evidence indicated that such self-efficacy mediated the negative effects of out-of-school technology use. In addition, evidence showed that the regular use of manipulative objects as part of math instruction in elementary grades improved the rate of student math learning but had no significant effect on student math achievement scores.

Teacher attitudes and beliefs about the NCLB standardized testing mandates impacted student achievement, as did the degree to which students perceived the teacher as maintaining an orderly classroom and choosing an appropriate instructional style. Teacher participation in professional development programs positively associated with increased student test scores. The lack of understanding of math concepts in elementary education program trainees contributed to inadequate math achievement scores for students taught by such teachers. One recommendation was to increase the required math training of prospective elementary teachers in elementary education programs to ensure that elementary teachers had a good grasp of the foundations of math. CIPs offered a novel approach to teacher preparation (Hurt, 2015). By having all of the resources necessary for sound instruction, teachers had more time available for whole-group instruction, small-group instruction, and remediation (Hurt, 2015).

Although CIPs offered the possibility to increase student achievement, strict mandates to follow these programs exactly adversely affected student achievement (Katz & Shahar, 2015; Sparks, 2014). Teacher autonomy, self-efficacy, and self-direction linked to increased student achievement (Katz & Shahar, 2015; Sparks, 2014). The key to the success of CIPs centered in the applied leadership approach (Hurt, 2015). Hurt (2015) stated the components of the CIP consisted of ingredients, not a completed dish. The teacher had the autonomy of presentation and preparation (Hurt, 2015). Engaged

students and motivated teachers led to substantial student achievement growth (Wagner & Dintersmith, 2015). High-stakes test results linked to accountability mandates served as judges for educators, schools, and students (Wagner & Dintersmith, 2015). Dutiful preparation for these assessments and strategic planning could lead to increased student achievement (Wagner & Dintersmith, 2015).

Policies aimed at reforming education actually stifled student learning and disheartened teachers (Wagner & Dintersmith, 2015). Leading experts felt reform measures in the form of accountability structures actually did harm (Wagner & Dintersmith, 2015). Wagner and Dintersmith (2015) shared that mandates led to standardized tests that only prepared individuals for routine tasks. Employers needed problem solvers, yet the educational system stifled the creativity needed for this essential skill (Wagner & Dintersmith, 2015). While comprehensive instructional programming increased student achievement on high-stakes tests, its long-term use linked to the inclusion of project-based learning (Wagner & Dintersmith, 2015).

Comprehensive instructional programming offered an innovative approach to increase student achievement, but the use of such a program must include a support structure and balance (Hurt, 2015). There was a need to guard teacher autonomy and efficacy (Katz & Shahar, 2015; Sparks, 2014). It was essential that data gathered from formative assessments aligned with the comprehensive curriculum to determine future instruction (Bancroft, 2010). The inclusion of instruction that promoted collaboration and innovation was essential (Wagner & Dintersmith, 2015).

.

Chapter III: Methodology

The purpose of this chapter was to provide the methodology used to conduct the study. The researcher sought to analyze the differences in mathematics and reading SOL test scores (grades 3-8) and the differences in SOL scores based on school configuration in District A within Region 7 in the Commonwealth of Virginia after the implementation of the CIP. In this chapter, the researcher provided information about the research design, methods, and tests used to collect and analyze the data. Within this chapter, the researcher described the methodology used to explore the four research questions presented earlier.

Research question 1. What differences existed in student achievement scores (pass percentages) in reading in grades 3-8 as measured by the Virginia Standards of Learning testing program between pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

Research question 2. What differences existed in student achievement scores (pass percentages) in math in grades 3-8 as measured by the Virginia Standards of Learning testing program comparing pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

Research question 3. What differences, if any, existed among K-4, 5-8, and K-8 school configurations in student achievement scores (pass percentages) in reading in grades 3-8 as measured by the Virginia Standards of Learning testing program comparing

pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

Research question 4. What differences, if any, existed among K-4, 5-8, and K-8 school configurations in student achievement scores (pass percentages) in math in grades 3-8 as measured by the Virginia Standards of Learning testing program comparing pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

In the first section of the chapter, the researcher described an overview of the research design, followed by a section explaining the population of the study and the sampling method used in the conduct of the study. Next, the researcher described the procedure used to collect data, followed by a section explaining the analytical methods used on the study findings. In the final three sections, the researcher described the reliability and validity of the measures used in this study, the limitations and delimitations of the study, and the assumptions and biases of the researcher in the development of this study.

Research Design

This study was a retrospective, non-experimental, quantitative study that compared the SOL test scores (pass percentages) in reading and math in grades 3-8 in District A in Region 7 of Virginia's public school system over the course of six academic years, beginning in the 2012-2013 academic school year and concluding in the 2017-2018 academic school year. The study was retrospective in that the researcher used historical data. The first three academic school years occurred before the district

implemented the CIP program (2012-2013, 2013-2014, 2014-2015) and the final three academic school years (2015-2016, 2016-2017, 2017-2018) were post-implementation of the CIP program.

This study included data from three years prior to CIP implementation to three years after CIP implementation in both reading and math achievement scores for all students and all schools in grades 3-8 of District A. The differences in student achievement post-CIP implementation compared to pre-CIP implementation showed the impact of CIP implementation and could identify specific schools within the district that required assistance in proper CIP implementation. In addition, the analysis in this study included school configuration comparisons in both reading and math to address the issue of whether CIP was more successful in some school configurations than others. Thus, in this analysis, the researcher identified areas where the CIP implementation needed improvement in specific school configurations.

The decision to use a quantitative study rather than a qualitative study arose from the need to understand how the implementation of CIP affected the reading and math achievement scores on the annual SOL assessments required by the state of Virginia. In particular, the researcher developed research questions for this study asked about numerical relationships to determine the presence or absence of statistically significant differences among various measures. Creswell (2014) noted that qualitative study designs were more appropriate for research questions that investigated the lives of individuals (narrative studies), those that asked about the life experiences of individuals experiencing a specific phenomenon (phenomenological studies) or sought to identify key variables to create a potential theory about a situation (grounded theory), those that explored the ethnology of a connected group (ethnological studies), or those that explored

a specific case in extreme detail (case studies). None of those approaches addressed the research questions posed in this study.

The choice of a non-experimental study design derived from the researcher's access to student achievement test scores from both before and after the district implemented the CIP across grades 3-8 and all schools in a single academic school year. A non-experimental design referred to a study where the researcher did not manipulate any variables in the study (Creswell, 2014). In the case of this study, students fell into different groups based on the schools they attended and their grade levels rather than via random assignments. Though the researcher was analyzing differences based on the implementation of CIP as the teaching paradigm, the researcher did not manipulate any variables or implement the CIP program. The researcher obtained archival data and performed statistical analyses after the district had already implemented CIP. With the data organized by the year—the participants took the achievement tests each year—by grade level, and by the school attended during that school year, the researcher analyzed student achievement scores in reading and math to directly address the four research questions. The reading achievement scores for all students and all grades both before and after the implementation of the CIP addressed research question one. The math achievement scores for all students in all grades and all schools both before and after the implementation of the CIP addressed research question two. A comparison of changes in both reading and math achievement scores before and after the implementation of the CIP provided a means to determine if different school configurations affected student achievement scores in reading or math in a significant way.

Population of the Study

The population of this study consisted of all students in the district under study who took the reading and math SOL achievement tests for grades 3-8 during the school years from 2012-2013 through 2017-2018 and all schools housing grades 3-8 in the years 2012-2013 through 2017-2018. The sample selected for this study for student achievement test scores was identical to the population, since the researcher included all student reading and math achievement test scores from this population for the six academic years noted. The sample of schools was also complete, including all schools in the relevant grades in District A. District A served six unique communities and has three school configurations. The school district had a total of eight elementary and middle schools. The school configurations of these schools varied. Of these eight schools, three were grades K-4, three were grades 5-8, and two were grades K-8. Two of the communities were small and have a grades K-8 configuration in which one school serves the entire community. One of the grades K-8 schools had approximately 445 students and the other had approximately 275 students. Three of the communities were larger and had both a grades K-4 school and a grades 5-8 school. The grades K-4 schools had student populations ranging from 425 to 785 and the grades 5-8 schools had student populations ranging from 345 to 600.

Data Collection

This study included the collection of records from student achievement tests derived from SOL test score pass percentages in publicly available databases. The researcher extracted data for this study from a publicly available database of SOL achievement test results. Data extracted included the following:

- Grade level, in this case grades 3-8, extracted individually;
- Reading and math test scores;
- Academic school years 2012-2013, 2013-2014, 2014-2015, 2015-2016, 2016-2017, and 2017-2018; and
- School configuration.

After collected, the researcher transferred the data into a spreadsheet and formatted the data in preparation for statistical analyses.

Analytical Methods

According to Schenker and Rumrill (2004), the statistical analysis techniques for inferential statistics included *t*-tests and analysis of variance (ANOVA) methods. For the current study, the researcher used independent samples *t*-tests and two-way ANOVA to answer the research questions.

To answer research question one, the researcher used an independent samples *t*-test to determine if the two populations of students, those measured pre-intervention (pre-CIP implementation) and those measured post-intervention (post-CIP implementation), differed in pass percentages in reading. The independent variable was measurement period with two levels—those measured pre-CIP implementation and post-CIP implementation. The dependent variable was pass percentages in reading. The researcher assessed the assumptions of normality and equal variances needed for an independent samples *t*-test prior to interpreting test results. To answer research question two, the researcher used an independent samples *t*-test to determine if the two populations of students, those measured pre-intervention (pre-CIP implementation) and those measured post-intervention (post-CIP implementation), differed in pass percentages in

math. The researcher assessed the assumptions of normality and equal variances needed for an independent samples *t*-test prior to interpreting test results.

To answer research question three, the researcher used a two-way ANOVA to determine statistical differences in K-4, 5-8, and K-8 school configurations in student achievement scores (pass percentages) in reading in grades 3-8 pre-CIP implementation and post-CIP implementation. The independent variable was school configuration with three levels (i.e., K-4, 5-8, and K-8) and measurement period with two levels (pre-CIP implementation and post-CIP implementation). The dependent variable was pass percentages in reading. Prior to data interpretation, the researcher inspected the data to determine if the assumptions of normality and equal variances were met. To answer research question four, the researcher used a two-way ANOVA to determine statistical differences in K-4, 5-8, and K-8 school configurations in student achievement scores (pass percentages) in math in grades 3-8 pre-CIP implementation and post-CIP implementation. The independent variable was school configuration with three levels (K-4, 5-8, and K-8) and measurement period with two levels (pre-CIP implementation and post-CIP implementation). The dependent variable was pass percentages in reading. Prior to data interpretation, the researcher inspected the data to determine if the assumptions of normality and equal variances were met.

A significance, α, of 0.05 or smaller determined statistical significance of the results. The researcher used an open-source statistical data analysis software package, Jeffreys's Amazing Statistics Program (JASP), version 0.9.0.1, to conduct the statistical study of the data. Released under a free and open source license, JASP was named after the Bayesian pioneer Sir Harold Jeffreys.

Reliability and Validity

In a non-experimental study, it was essential to establish external validity because internal validity of the research design was challenging to determine (Schenker & Rumrill, 2004). External validity relied on the degree to which the test sample was representative of the overall population. For this study, the test sample was identical to the overall population of students in grades 3-8 over the years of the study period. Thus, the results of the analysis in this study were representative of the student population in this district over that time.

Reliability and validity of the measures used in this study provided another means of determining the reliability of the results of the study. The measures for student achievement in this study derived from the achievement test scores from the Virginia SOL. In the context of the SOL achievement test program, reliability referred to a high correspondence between the achievement test score and the student's proficiency in the subject tested. Thus, a highly reliable score implied test/re-test consistency.

In addition, validity was a measure of four key elements: did the test content cover items from the Virginia SOL without including extraneous content; did the test measure the knowledge expected based on how the student answered the questions; were the test questions consistent internally and across ethnic groups (a statistical measure using coefficient alpha measures); and was the test results consistent with other measures such as student grades. A technical report on the Virginia SOL (VDOE, 2016) stated that the "direct relationship of the SOL curriculum framework with the SOL test blueprint and the SOL assessments lended support to the content validity of the SOL assessment" (p. 38). Further validity assessments of the test measuring desired curriculum proficiency came from the same report, "The items on the Virginia SOL assessments are measuring

the content standards and not measuring other, unintended constructs or disadvantaging particular student subgroups" (VDOE, 2016, p. 40).

Limitations and Delimitations

Limitations referred to possible weaknesses the researcher had no control over and thus could not remove (Creswell, 2014). For the current study, a weakness was the lack of ability to measure the changes on students as they progressed from grade level to grade level across the six academic years of the study duration. Some students moved out of the district; some students moved into the district; other students stayed in the district but moved from school to school. The researcher lacked a means to measure the number or impact of those student variations. This limitation forced an assumption that the overall impact of such student movements was negligible compared to the overall effect from the main variables. Thus, the researcher assumed that any changes in specific cohort membership from year to year had no significant effect on the overall student achievement scores for individual grades and schools. This assumption could be false if a number of students who were remarkably above grade level or remarkably below grade level either entered or left the population during the six-year study period. There was no way to determine if this was the case since the researcher received data blinded with respect to student identity.

Delimitations referred to the study limits or boundaries that the researcher placed on the study and which narrowed the focus (Creswell, 2014). The delimitations of this study derived from the choice of limits on the district venue for this study and the limits placed on the achievement tests and grade levels included. To be as inclusive as possible, the researcher chose to include the six grade levels (grades 3-8) in K-4, 5-8, and K-8 schools with state-mandated SOL reading and math achievement tests for all students. In

addition, the choice of district reflected the fact that this district and only this district implemented CIP across all schools and all grade levels in a single school year, the 2015-2016 academic school year. Such a clear pre-intervention and post-intervention process provided less complex evidence to address the research questions under consideration. Finally, the choice to include reading and math as the two subjects investigated addressed two fundamental skill sets that students require for success in both their upper level grades and in life in general.

Assumptions and Biases of the Study

The researcher assumed schools administered these assessments in a uniform manner aligned with protocols set forth in the School Test Coordinator's Manual. The researcher also assumed that each person who administered the assessments did so in a uniform manner aligned with the specific instructions in the School Test Coordinator's Manual, which very specific step-by-step instructions for testing administration. Another assumption in this study was that few students in the various study cohorts (i.e., grade levels of students who tested within the six-year period of the study) changed schools during the course of this study period. The researcher assumed that any changes in specific cohort membership from year to year had no effect on the overall student achievement scores for individual grades and schools.

A final assumption was that the SOL tests were approximately of equivalent difficulty each year. The VDOE asserted that SOL tests, while different every year, were also statistically consistent from year to year in terms of difficulty (VDOE, 2018). Extensive statistical analysis of the SOL tests, available from VDOE (2018), provided support for this assumption.

Chapter IV: Analyses and Results

The purpose of this study was to determine if improvements in student SOL test scores since the implementation of CIP in District A were statistically significant and thus reflective of genuine improvements in student achievement as well as to determine which school configuration was most conducive to improvements in student achievement. The retrospective, non-experimental quantitative study compared the SOL test scores (pass percentages) in reading and math in grades 3-8 in District A in Region 7 of Virginia's public school system over the course of six academic years, beginning in the 2012-2013 academic school year and concluding the 2017-2018 academic school year. This study included data from three years prior to CIP implementation to three years after CIP implementation in both reading and math achievement scores for all students and all schools in grades 3-8 of District A. In addition, the analyses in this study included school configuration comparisons in both reading and math to address the issue of whether CIP was more successful in some school configurations than others. The researcher ran independent samples *t*-tests to determine if a significant difference in reading scores and math scores existed between pre-CIP and post-CIP implementation and a two-way ANOVA to determine if there was a statistical difference in student scores among the three school configurations.

Data Analysis

The researcher utilized the JASP data analysis software package to determine post-CIP implementation differences in SOL scores as well as post-CIP implementation differences in SOL scores by school configuration. JASP was an open-source statistical analysis software package. Data entry and analysis occurred with the use of the JASP

software. A collection of student data from all reading and all math SOL achievement test scores for grades 3-8 within District A during school years 2012-2013 through 2017-2018 made up the data set for the study. During data collection, the researcher ensured that sample participants (specifically for student achievement test scores) were identical to the population as the researcher sought to include all students that underwent reading and math achievement testing for the six academic years of interest. Data collection involved three areas that aided to have potential in achieving the study purpose. The data extracted from the publicly available database of SOL achievement test results served useful to help answer research question one about reading SOL scores pre-CIP implementation and post-CIP implementation. These data also played an important role in determining students' SOL achievements for the six academic years of concern. These data also allowed the researcher to compare learning achievement of students between pre-CIP and post-CIP implementation within District A. Additionally, the data served as essential for three school configuration levels within the district. This included K-4, 5-8, and K-8 in District A. Through this data, the researcher was able to differentiate the student achievement scores in reading in grades 3-8 as measured by the Virginia SOL testing program between pre-CIP implementation and post-CIP implementation. Consequently, findings under this section were to answer research question one as outlined in the introduction and methodology section.

The student achievement data obtained from these databases of interest also enabled the researcher to determine the difference that existed in student achievement scores in math in grades 3-8 as measured by the Virginia SOL testing program comparing pre-CIP implementation and post-CIP implementation in the Region 7 District A under

study. Subsequently, this enabled the researcher to answer research question two. Since data on students' achievement also engaged segregation based on school level, the researcher was also able to explore descriptively the difference that existed among K-4, 5-8, and K-8 schools during three years prior to CIP implementation and three years after CIP implementation. This enabled the researcher to answer research questions three and four.

Research Questions

Research question 1. What differences existed in student achievement scores (pass percentages) in reading in grades 3-8 as measured by the Virginia Standards of Learning testing program between pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

To answer research question one, the researcher ran an independent samples *t*-test to determine if a significant difference in reading scores existed between pre-CIP and post-CIP implementation. An independent samples *t*-test showed whether or not there was a significant difference in reading scores between pre-CIP and post-CIP in grades 3-8 as measured by the Virginia SOL testing program. Levine's test for equality of variances showed that the assumption of variance was not met ($F = 5.001$, $p = .026$). Since the assumption of equal variances was violated, the researcher used the *t* statistic and *p* value for equal variances not assumed. The researcher determined that there was a significant difference in reading scores between pre-CIP and post-CIP in grades 3-8 as measured by the Virginia SOL testing program ($t = -5.131$, $p = .000$) (see Table 1). The

reading pass rates post-CIP implementation (*M* = 87.67%) were significantly higher compared to reading pass rates prior to CIP implementation (*M* = 82.47%).

Table 1

Independent sampless t-test between reading achievement scores and student performance

	t	df	p	Std. Error Dif.	95% Confidence Interval of the Difference	
					Lower	Upper
Pass Rate	-5.131	200.007	.000	1.013	-7.194	-3.200

Research question 2. What differences existed in student achievement scores (pass percentages) in math in grades 3-8 as measured by the Virginia Standards of Learning testing program comparing pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

To answer research question two, the researcher ran an independent sampless *t*-test to determine if a significant difference existed between pre-CIP and post-CIP on math passing rates. An independent sampless *t*-test showed whether there was a significant difference in math scores between pre-CIP and post-CIP in grades 3-8 as measured by the Virginia SOL testing program. Levene's test for equality of variances showed that the assumption of variance was not met ($F = 8.348$, $p = .004$). Since the assumption of equal variances was violated, the researcher used the t statistic and p value

for equal variances not assumed. The researcher determined that there was a significant difference in math scores between pre-CIP and post-CIP in grades 3-8 as measured by the Virginia SOL testing program (t = -4.376, p = .000) (see Table 2). The math passing rates post-CIP implementation (M = 90.94%) were significantly higher compared to math passing rates prior to CIP implementation (M = 85.54%).

Table 2

Independent sampless t-test between math achievement scores and student performance

					95% Confidence Interval of the Difference	
	t	df	p	Std. Error Dif.	Lower	Upper
Pass Rate	-4.376	201.429	.000	1.234	-7.833	-2.967

Research question 3. What differences, if any, existed among K-4, 5-8, and K-8 school configurations in student achievement scores (pass percentages) in reading in grades 3-8 as measured by the Virginia Standards of Learning testing program comparing pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

To answer research question three, the researcher conducted a two-way factorial ANOVA to determine if a significant difference existed regarding reading SOL pass rates pre-CIP implementation and post-CIP implementation based on school configuration. The researcher conducted the ANOVA to compare the main effects of CIP implementation and school configurations on reading SOL scores (see Table 3).

Table 3

Two-way ANOVA reading scores

Tests of Between-Subjects Effects

Dependent Variable: Reading Pass Rate

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
School Con	1184.903	2	592.451	12.264	.000
Implementation	1217.091	1	1217.091	25.195	.000
School Con Implementation	226.792	2	113.396	2.347	.098

Levene's test for equality of variances showed that the assumption of variance was met ($F = 4.545$, $p = .091$). There was a statistically significant difference in the main effect of school configuration on reading pass rates ($F = 12.26$, $p = .000$). Tukey post hoc tests revealed a significant difference between K-4 and K-8 school configurations ($p = .000$) and between 5-8 and K-8 school configurations ($p = .000$). K-8 school configuration ($M = 88.35$) had significantly higher reading pass rates than K-4 school configuration ($M = 82.93$) and 5-8 school configuration ($M = 83.55$). A significant difference also existed between the main effect of CIP implementation on reading pass scores ($F = 25.19$, $p = .000$). Reading pass rates were significantly higher after CIP implementation ($M = 87.45$) compared to pre-CIP implementation ($M = 82.43$). There was not a significant interaction between school configuration and implementation of CIP on reading pass rates ($F = 2.35$, $p = .098$).

Research question 4. What differences, if any, existed among K-4, 5-8, and K-8 school configurations in student achievement scores (pass percentages) in math in grades

3-8 as measured by the Virginia Standards of Learning testing program comparing pre-Comprehensive Instructional Program implementation (2012-2013, 2013-2014, 2014-2015) and post-Comprehensive Instructional Program implementation (2015-2016, 2016-2017, 2017-2018) in the Region 7 District A?

To answer research question four, the researcher conducted a two-way factorial ANOVA to determine if a significant difference existed regarding math SOL pass rates pre-CIP implementation and post-CIP implementation based on school configuration (see Table 4).

Table 4

Two-way ANOVA math scores

Tests of Between-Subjects Effects

Dependent Variable: Math Pass Rate

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
School Con	200.111	2	100.055	1.247	.298
Implementation	1319.261	1	1319.261	16.448	.000
School Con Implementation	151.932	2	75.966	.947	.390

The researcher conducted the ANOVA to compare the main effects of CIP implementation and school configurations on math SOL scores. Levene's test for equality of variances showed that the assumption of variance was met ($F = 3.007$, $p = .092$). There was no statistically significant difference in the main effect of school configuration on math passing rates ($F = 1.25$, $p = .289$). There was a significant difference in the main effect of CIP implementation on math passing rates ($F = 16.45$,

p = .000). Math pass rates were significantly higher after CIP implementation (M = 90.61) compared to pre-CIP implementation (M = 85.37). There was not a significant interaction between school configuration and implementation of CIP on math pass rates (F = .947, p = .390).

Summary of Results

The researcher obtained data for grades 3-8 in reading and math for pre-CIP and post-CIP implementation in the Region 7 District A being studied, enabling all four research questions to be answered. To answer the first research question, the researcher ran an independent sampless *t*-test to determine if a significant difference in reading scores existed between pre-CIP and post-CIP implementation. The researcher determined that there was a significant difference in reading scores between pre-CIP and post-CIP in grades 3-8 as measured by the Virginia SOL testing program. To answer the second research question, the researcher ran an independent sampless *t*-test to determine if a significant difference existed between pre-CIP and post-CIP on math passing rates. The researcher determined there was a significant difference in math scores between pre-CIP and post-CIP in grades 3-8 as measured by the Virginia SOL testing program. To answer the third research question, the research conducted a two-way ANOVA to determine if a significant difference existed in reading pass rates pre-CIP and post-CIP based on school configuration. Based on the two-way ANOVA results, the researcher concluded that there was a statistically significant difference of reading pass rates pre-CIP and post-CIP based on school configuration. Tukey post hoc tests revealed a significant difference between K-4 and K-8 school configurations and between 5-8 and K-8 school configurations. The K-8 school configurations had significantly higher

reading pass scores than K-4 and 5-8 configurations. Reading pass rates were significantly higher after CIP implementation compared to pre-CIP implementation. There was not a significant interaction between school configuration and implementation of CIP on reading pass rates. To answer the fourth research question, the researcher ran a two-way ANOVA test to determine if a significant difference existed in math pass rates pre-CIP and post-CIP based on school configuration. Based on the two-way ANOVA results, the researcher concluded that there was no statistically significant difference of math pass rates pre-CIP and post-CIP based on school configuration; however, there was a significant difference in the main effect of CIP implementation on math passing rates. Math pass rates were significantly higher after CIP implementation compared to pre-CIP implementation. There was not a significant interaction between school configuration and implementation of CIP on math pass rates.

Chapter V: Conclusions and Recommendations

The researcher used the information gathered throughout the study to summarize the purpose and provide implications for further study. The purpose of this study was to determine if improvements in student SOL test scores since the implementation of CIP in District A were statistically significant and thus reflective of genuine improvements in student achievement as well as to determine which school configuration was most conducive to improvements in student achievement. The researcher used the evidence and sought to determine if the CIP program had an impact on student and school performance within District A. Similarly, the researcher outlined the significance of the study results as well as how those related to the literature gaps. The researcher also included further recommendations on practicable strategies for CIP implementation in District A as well as other school districts in Region 7 and across the state of Virginia.

Discussion and Conclusions of the Study

It was evident in the study findings that there was a significant difference in pre-CIP implementation and post-CIP implementation. Post-CIP implementation SOL scores were higher than pre-CIP implementation SOL scores (within the period of 2012-2013, 2013-2014, and 2014-2015). Consequently, after the introduction of the CIP program (between 2015-2016, 2016-2017, and 2017-2018), student performance on SOL tests started to rise. The researcher determined that there was a significant difference in reading scores between pre-CIP and post-CIP in grades 3-8 as measured by the Virginia SOL testing program. The reading pass rates post-CIP implementation (M = 87.67) were significantly higher compared to reading passing rates prior to CIP implementation (M = 82.47). This increase in performance related to the ability of the students to read

and understand questions presented in the SOL assessment. It was important to note that all classroom/academic subjects required students be able to read. The literacy level of students was important for them to be able to improve their learning outcome just as was observed in the literature section.

Based on the study findings, the conclusion was that the CIP program in mathematics had an influence on improving student performance within District A. The researcher determined that there was a significant difference in math scores between pre-CIP and post-CIP in grades 3-8 as measured by the Virginia SOL testing program. The math passing rates post-CIP implementation (M = 90.94) were significantly higher compared to math passing rates prior to CIP implementation (M = 85.54). It was viable to recommend increased implementation of the program within schools in other districts in Region 7 and in the state to improve student performance within the schools as well as improve school performance.

The researcher found there were statistically significant differences in the main effect of school configuration on reading pass rates. K-8 school configuration had significantly higher reading pass rates than K-4 school configuration and 5-8 school configuration. A significant difference also existed between the main effect of CIP implementation on reading pass scores. Reading pass rates were significantly higher after CIP implementation (M = 87.45) compared to pre-CIP implementation (M = 82.43). There was not a significant interaction between school configuration and implementation of CIP on reading pass rates (F = 2.35, p = .098).

The researcher found there were no statistically significant differences in the main effect of school configuration on math pass rates. There was a significant difference in

the main effect of CIP implementation on math passing rates. Math pass rates were significantly higher after CIP implementation (M = 90.61) compared to pre-CIP implementation (M = 85.37). There was not a significant interaction between school configuration and implementation of CIP on math pass rates (F = .947, p = .390).

These findings were beneficial in understanding the topic under investigation because they provided evidence regarding the impact of reading and math testing on performance concerning CIP implementation, thus, confirming that factors of school configuration affected performance (Fisher & Frey, 2018). The researcher noted that the school configuration affected student achievement in reading, which confirmed Hurt's (2015) finding that time, ability, and availability of necessary resources affected student achievement. The finding regarding school configuration was important, as it was often difficult to obtain resources in schools with wider grade ranges, such as K-8, as opposed to K-4 or 5-8 (Hurt, 2015). The findings of the present study somewhat contradicted the findings of Katz and Shahar (2015) and Sparks (2014), who found that there was an adverse effect on student achievement when schools were mandated to strictly follow CIP programs.

Implications for Practice

The CIP program had an impact on student performance in the context of the present study. Prior researchers, such as Hurt (2015), suggested that other factors were influential—such as student-teacher interactions and student motivation. The researcher did not study these aspects in the present study but it is important to remember that part of an effective CIP program includes effective interaction in the classroom and adequate motivation (Hurt, 2015). The fact that the CIP program impacted student performance,

especially at the K-8 level in reading, revealed that school configuration may be an influencer as well. In District A the CIP program was found to be most influential at the K-8 level in reading as compared to the K-4 or 5-8 level; therefore, it would be beneficial for administrators of other school configurations (K-4 and 5-8) to assess what differences in implementation existed in the CIP program and emulate the more effective implementation.

The researcher found that this study had implications for the Commonwealth of Virginia and its educational leadership by providing evidence that there was substantial statistical improvement in reading and math SOL test scores in grades 3-8 following CIP implementation. The researcher also concluded there was a statistically significant difference in the main effect of school configuration on reading pass rates. K-8 school configurations had significantly higher reading pass rates than K-4 and 5-8 school configurations. This finding could have implications for school divisions that may be considering a K-8 school configuration as they discuss possible consolidation options or even the construction of new buildings.

State and district leaders should consider the findings and influence of the CIP program on student performance when drafting policy and implementing programs to increase student pass rates on SOL assessments. Specifically, leaders in Region 7, District A, should consider the results as they prepare to make informed decisions about the continuation of the CIP and potential implementation of similar programs in the future. School leaders could use the findings of this study to arrive at a decision related to the continued use, discontinuation, or implementation of the CIP. Furthermore, the study may help to determine how much emphasis school leaders place on the CIP

implementation. That is, this study could assist in the decision whether to make the CIP the absolute curriculum or simply an available resource given the fact that District A used the CIP as the absolute curriculum during post-CIP implementation.

The results of this study may serve as a tool for school district leadership as they consider the implementation of programs such as the CIP. Although the findings of this study are limited to District A in Region 7 of the Commonwealth of Virginia, districts that are considering the implementation of similar programs should consider the findings from this study due to the significant increase in SOL pass rates in grades 3-8 following CIP implementation. In addition, leaders may use this study to initiate similar studies within districts both inside and outside of Virginia. Based on the significant statistical findings of the this study, school districts searching for ways to increase reading and math SOL scores in grades 3-8 to avoid losing state accreditation or federal funding should definitely consider the use of the CIP. Based on the significant statistical findings in increased SOL scores, this study could serve as a model for any school district considering the use of a CIP aimed to increase student achievement on end-of-course state assessments in both reading and math in grades 3-8, especially for reading improvement in a K-8 school.

Recommendations for Future Research

This study contributed to the effectiveness of CIPs. Student-teacher interactions and student motivation considerations must take place as well to lead to improved outcomes in testing. Future studies should address student success rates on Virginia SOL assessments in school districts with consideration given for both student-teacher interactions and student motivation. School improvement is more than a pre-scripted

recipe and should include adaptive areas (such as rapport, motivation, interactions, and relationships), as well as technical areas like instructional resources and tools. School leaders in charge of drafting and implementing policy regarding the implementation of programs similar to the CIP used in Region 7 District A need more information based on further research considering these adaptive areas. School leaders must make prudent purchase decisions of resources that may lead to improved student achievement due to limited availability of funding. Teachers value such resources as the CIP that are meaningful, relevant, and purposeful. Although this study contributed to this area of research, related areas need more exploration. Future studies should focus on growth in mathematics and reading as opposed to the pass/fail model defined in this study.

The researcher found positive differences in student achievement in mathematics and reading following the implementation of the CIP in Region 7 District A. Those findings were results from pre-CIP and post-CIP test scores on the Virginia SOL assessments in mathematics and reading in District A. This researcher recommends to study demographic characteristics of local communities and see which of those characteristics, if any, influenced student achievement. This may enable schools to develop programs to combat problems related to those characteristics and/or offer support to affected students and families. It would be beneficial to have a study that included a control group of students and an experimental group within each classroom to determine if there would be a significant difference in reading or math achievement. It would also be beneficial to have a population from numerous school districts within Virginia to see if the results would differ or be similar to those in District A.

The study needs replication in more educational regions of Virginia as this program and similar programs become topics of discussion for adoption. Similarly, the administration of studies of comprehensive instructional programs outside of the state of Virginia need to be administered to assess the validity of such programs in states with different educational standards and testing models. Finally, replication of this exact study based upon a longer duration after the implementation of the CIP in Region 7 District A would be beneficial. Future studies could examine the effect after five years and again after ten years of data post-CIP implementation become available to chart historical trends. These studies would eliminate any error related to training and the associated learning curve. All of these recommended studies would broaden research related to the effectiveness of comprehensive instructional programs.

Conclusion to the Study

Increased student achievement is a primary focus for school districts. The methods for measuring student performance varies among states but are similar in the basic framework. Federal and state mandates create target pass rates for overall and subgroup populations on state assessments (Bancroft, 2010). Given this scenario, it is imperative that school districts make informed decisions regarding policy and practice implementation aimed to increase student achievement. Further, limited educational funding makes data-driven decision making even more critical.

The accountability era requires alignment; products and programs like the CIP are becoming readily available. These programs offer the possibility to increase student achievement by supplying teachers with all needed resources; therefore, more time is available for direct contact with students.

The researcher examined the implementation of the CIP in Region 7, District A of the Virginia Public School System. Student achievement scores (pass percentages) in mathematics and reading were compared for pre-CIP and post-CIP implementation. Data showed a positive difference in student achievement in both mathematics and reading after the implementation of the CIP. The K-8 school configurations had significantly higher reading pass scores than K-4 and 5-8 configurations. There was no statistically significant difference of math pass rates pre-CIP and post-CIP based on school configuration. These findings align with related existing research. As school leaders consider implementation and draft policies related to the use of a comprehensive instructional programs, review of the findings of this study is crucial.

References

Alspaugh, J. W. (1998). Achievement loss associated with the transition to middle school and high school. *The Journal of Educational Research, 92*(1), 20–25.

Bakosh, L., Snow, R., Tobias, J., Houlihan, J., & Barbosa-Leiker, C. (2015). Maximizing mindful learning: Mindful awareness intervention improves elementary school students' quarterly grades. *Mindfulness, 7*(1), 59-67. doi: 10.1007/s12671-015-0387-6

Bancroft, K. (2010). Implementing the mandate: The limitations of benchmark tests. *Educational Assessment, Evaluation and Accountability, 22*(1), 53.

Bautista, A., & Wong, J. (2017). Music teachers' perceptions of the features of most and least helpful professional development. *Arts Education Policy Review, 3*(2), 1-14. doi: 10.1080/10632913.2017.1328379

Bevel, K. R., & Mitchell, M. R. (2012). The effects of academic optimism on elementary reading achievement. *Journal of Educational Administration, 50*(6), 773-787. doi: 10.1108/09578231211264685

Bowman-Perrott, L., & Lewis, C. (2008). An examination of reading and discipline data for elementary and secondary African American students: Implications for special education. *Multicultural Learning and Teaching, 3*(2), 73-98. doi: 10.2202/2161-2412.1036

Burns, M., Walick, C., Simonson, G., Dominguez, L., Harelstad, L., Kincaid, A., & Nelson, G. (2015). Using a conceptual understanding and procedural fluency heuristic to target math interventions with students in early elementary. *Learning Disabilities Research & Practice, 30*(2), 52-60. doi: 10.1111/ldrp.12056

Carr, M., Taasoobshirazi, G., Stroud, R., & Royer, J. (2011). Combined fluency and cognitive strategies instruction improves mathematics achievement in early elementary school. *Contemporary Educational Psychology, 36*(4), 323-333. doi: 10.1016/j.cedpsych.2011.04.002

Chingos, M. M., Whitehurst, G. J., & Gallaher, M. R. (2015). School districts and student achievement. *Education Finance and Policy, 10*(3), 378-398. doi: 10.1162/EDFP_a_00167

Comprehensive Instructional Program. (2016). *Comprehensive Instructional Program benchmark assessment analysis.* Charlottesville, VA: Author. Retrieved from http://cip.education/Portals/5/Resources/CIP%202015-2016%20Benchmark-SOL%20Correlations.pdf

Comprehensive Instructional Program. (2018). *Comprehensive Instructional Program Materials.* Charlottesville, VA*:* The Comprehensive Instructional Program. Retrieved from http://www.cip.education/Home.aspx

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: SAGE.

Crosnoe, R., Morrison, F., Burchinal, M., Pianta, R., Keating, D., Friedman, S., & Clarke-Stewart, K. (2010). Instruction, teacher–student relations, and math achievement trajectories in elementary school. *Journal of Educational Psychology, 102*(2), 407-417. doi: 10.1037/a0017762

DeMatthews, D. (2015). Making sense of social justice leadership: A case study of a principal's experiences to create a more inclusive school. *Leadership and Policy in Schools, 14*(2), 139-166. doi: 10.1080/15700763.2014.997939

Edwards, H. (2011). *A quantitative examination of grade configuration and achievement of eighth-grade students in Georgia on the mathematics and reading criterion-referenced competency tests* (Doctoral dissertation). Retrieved from ProQuest LLC. (ED550014)

Eros, J. (2013). Second-stage music teachers' perceptions of their professional development. *Journal of Music Teacher Education, 22*(2), 20-33. doi: 10.1177/1057083712548771

Faust, M., & Kandelshine-Waldman, O. (2011). The effects of different approaches to reading instruction on letter detection tasks in normally achieving and low achieving readers. *Reading and Writing, 24*(5), 545-566. doi: 10.1007/s11145-009-9219-1

Fisher, D., & Frey, N. (2007). Implementing a schoolwide literacy framework: Improving achievement in an urban elementary school. *The Reading Teacher, 61*(1), 32-43. doi: 10.1598/RT.61.1.4

Fisher, D., & Frey, N. (2018). Raise reading volume through access, choice, discussion, and book talks. *The Reading Teacher, 72*(1), 89-97. doi: 10.1002/trtr.1691

Glover, T. A., Reddy, L. A., Kettler, R. J., Kurz, A., & Lekwa, A. J. (2016). Improving high-stakes decisions via formative assessment, professional development, and comprehensive educator evaluation: The school system improvement project. *Teachers College Record, 118*(14), 1-26.

Gottfried, M., & Graves, J. (2014). Peer effects and policy: The relationship between classroom gender composition and student achievement in early elementary school. *The B. E. Journal of Economic Analysis & Policy*, *14*(3), 937-977. doi: 10.1515/bejeap-2013-0123

Grissom, J. A., Nicholson-Crotty, S., & Harrington, J. R. (2014). Estimating the effects of No Child Left Behind on teachers' work environments and job attitudes. *Educational Evaluation and Policy Analysis, 36*(4), 417-436. doi: 10.3102/0162373714533817

Harlaar, N., Deater-Deckard, K., Thompson, L., DeThorne, L., & Petrill, S. (2011). Associations between reading achievement and independent reading in early elementary school: A genetically informative cross-lagged study. *Child Development, 82*(6), 2123-2137. doi: 10.1111/j.1467-8624-2011.01658.x

Heatly, M., Bachman, H., & Votruba-Drzal, E. (2015). Developmental patterns in the associations between instructional practices and children's math trajectories in elementary school. *Journal of Applied Developmental Psychology*, *41*, 46-59. doi: 10/1016/.appdev.2015.06.002

Herman, B. E. (2004). The revival of K-8 schools. *Phi Delta Kappan Fastbacks, 5*(19), 7-37.

Hirvonen, R., Tolvanen, A., Aunola, K., & Nurmi, J. (2012). The developmental dynamics of task-avoidant behavior and math performance in kindergarten and elementary school. *Learning and Individual Differences, 22*(6), 715-723. doi: 10.1016/j.lindif.2012.05.014

Hoy, W. K., & Miskel, C. G. (2005). *Educational administration: Theory, research, and practice.* Boston, MA: McGraw-Hill.

Hoy, W. K., Tarter, C. J., & Woolfolk Hoy, A. (2006). Academic optimism of schools: A force for student achievement. *American Educational Research Journal, 43*(3), 425-446. doi: 10.3102/00028312043003425

Hurt, M. (2015). *Comprehensive school improvement: An argument against traditional educational reform.* Charlottesville, VA: Virginia School Boards Association. Retrieved from http://cip.education/Portals/5/Resources/Comprehensive%20School%20Improvement.pdf

Hurt, M. (2017, Spring). *Showcases for success: VSBA special edition newsletter*. Charlottesville, VA: Virginia School Boards Association. Retrieved from http://www.vsba.org/images/uploads/2017ShowcasesforSuccess_WebVersion.pdf

Illeris, K. (2015). The development of a comprehensive and coherent theory of learning. *European Journal of Education, 50*(1), 29-40. doi: 10.1111/ejed.12103

Ing, M., Webb, N., Franke, M., Turrou, A., Wong, J., Shin, N., & Fernandez, C. (2015). Student participation in elementary mathematics classrooms: The missing link between teacher practices and student achievement? *Educational Studies in Mathematics, 90*(3), 341-356. doi: 10.1007/s10649-015-9625-z

Juvonen, J., Le, V., Kaganoff, T., Augustine, C., & Constant, L. (2004). *Focus on the wonder years: Challenges facing the American middle school.* Santa Monica, CA: RAND.

Katz, I., & Shahar, B. (2015). What makes a motivating teacher? Teachers' motivation and beliefs as predicators of their autonomy-supportive style. *School Psychology International*, *36*(6), 575-588.

Kirnan, J., Siminerio, S., & Wong, Z. (2015). The impact of a therapy dog program on children's reading skills and attitudes toward reading. *Early Childhood Education Journal, 44*(6), 637-651. doi: 10.1007/s10643-015-0747-9

Kopcha, T. J. (2012). Teachers' perceptions of the barriers to technology integration and practices with technology under situated professional development. *Computers & Education, 59*(4), 1109-1121. doi: 10.1016/j.compedu.2012.05.014

Lee, D., Dang, T., Ulibas-Pascual, J., Gordon Biddle, K., Heller de Leon, B., Elliott, D., & Gorter, J. (2017). Exploring the influence of efficacy beliefs and homework help in predicting reading achievement among underserved children in an afterschool program. *The Urban Review, 49*(5), 707-728. doi: 10.1007/s11256-017-0418-9

Lee, J., & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990–2009 reading and math achievement gaps and trends. *Educational Evaluation and Policy Analysis*, *34*(2), 209-231. doi: 10.3102/0162373711431604

Look, K. (2002). *K-8 schooling: A retreat? A mandate? An intervention?* (Doctoral dissertation). Philadelphia, PA: University of Pennsylvania.

Madden, L., Beyers, J., & O'Brien, S. (2016). The importance of STEM education in the elementary grades: Learning from pre-service and novice teachers' perspectives. *Electronic Journal of Science Education, 20*(5), 1-18. Retrieved from http://ejse.southwestern.edu/article/viewFile/15871/10324

Meyers, C. (2012). The centralizing role of terminology: A consideration of achievement gap, NCLB, and school turnaround. *Peabody Journal of Education, 87*(4), 468. doi: 10.1080/061956X.2012.705149

Mokhtari, K., Thoma, J., & Edwards, P. (2009). How one elementary school uses data to help raise students' reading achievement. *The Reading Teacher, 63*(4), 334-337. doi: 10.1598/RT.63.4.10

Moon, U., & Hofferth, S. (2016). Parental involvement, child effort, and the development of immigrant boys' and girls' reading and mathematics skills: A latent difference score growth model. *Learning and Individual Differences, 47,* 136-144. doi: 10.1016/j.lindif.2016.01.001

Morrissey, T., & Vinopal, K. (2018). Neighborhood poverty and children's academic skills and behavior in early elementary school. *Journal of Marriage and Family, 80*(1), 182-197. doi: 10.1111/jomf.12430

Mowrer, O.H. (1960). *Learning theory and the symbolic processes.* Hoboken, NJ, US: John Wiley & Sons Inc.

Overbaugh, R., & Lu, R. (2008). The impact of a NCLB-EETT funded professional development program on teacher self-efficacy and resultant implementation. *Journal of Research on Technology in Education, 41*(1), 43-61. doi: 10.1080/15391523.2008.10782522

Park, S., Stone, S., & Holloway, S. (2017). School-based parental involvement as a predictor of achievement and school learning environment: An elementary school-level analysis. *Children and Youth Services Review, 82,* 195-206. doi: 10.1016/j.childyouth.2017.09.012

Ramirez, G., Gunderson, E., Levine, S., & Beilock, S. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development, 14*(2), 187-202. doi: 10.1080/15248372.2012.664593

Robinson, J., Myran, S., Strauss, R., & Reed, W. (2014). The impact of an alternative professional development model on teacher practices in formative assessment and student learning. *Teacher Development, 18*(2), 141-162. doi: 10.1080/13664530.2014.900516

Ronfeldt, M., Farmer, S. O., McQueen, K., & Grissom, J. A. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal, 52*(3), 475-514. doi: 10.3102/0002831215585562

Schenker, D. J., & Rumrill, P. D. Jr. (2004). Causal-comparative research designs. *Journal of Vocational Rehabilitation, 21*(3), 117-121.

Smith, J. M., & Kovacs, P. E. (2011). The impact of standards-based reform on teachers: The case of 'No Child Left Behind.' *Teachers and Teaching: Theory and Practice, 17*(2), 201-225. doi: 10.1080/13540602.2011.539802

Soni, A., & Kumari, S. (2015). The role of parental math anxiety and math attitude in their children's math achievement. *International Journal of Science and Mathematics Education, 15*(2), 331-347. doi: 10.1007/s10763-015-9687-5

Sparks, S. D. (2014). California school draws lessons from failure. *Education Week, 6*(4), 16-19.

Uribe-Flórez, L., & Wilkins, J. (2016). Manipulative use and elementary school students' mathematics learning. *International Journal of Science and Mathematics Education, 15*(8), 1541-1557. doi: 10.1007/s10763-016-9757-3

Virginia Department of Education. (2016). *Virginia Standards of Learning assessments.* Richmond, VA: Author. Retrieved from http://www.doe.virginia.gov/testing/test_administration/technical_reports/sol_technical_report_2014-15_administration_cycle.pdf

Virginia Department of Education. (2018). *The standards and SOL-based instructional resources.* Richmond, VA: Author. Retrieved from http://www.doe.virginia.gov/testing/sol/standards_docs/index.shtml

Wagner, T., & Dintersmith T. (2015). *Most likely to succeed: Preparing our kids for the innovation era.* New York City, NY: Scribner.

Weidinger, A., Steinmayr, R., & Spinath, B. (2018). Changes in the relation between competence beliefs and achievement in math across elementary school years. *Child Development, 89*(2), e138-e156. doi: 10.1111/cdev.12806

Wieczorek, D. (2017). Principals' perceptions of public schools' professional development changes during NCLB. *Education Policy Analysis Archives, 25*, 1-49. Retrieved from http://www.redalyc.org/pdf/2750/275050047008.pdf

Wren, S. D. (2003). *The effect of grade span configuration and school-to-school transition on student achievement.* Detroit, MI: Wayne State University. (ED 479332)

Xu, Z., & Jang, E. (2017). The role of math self-efficacy in the structural model of extracurricular technology-related activities and junior elementary school students' mathematics ability. *Computers in Human Behavior, 68,* 547-555. doi: 10.1016/j.chb.2016.11.063

Yecke, C. P. (2006). Mayhem in the middle: Why we should shift middle to K-8. *Educational Leadership, 63*(7), 20–25.